ANATOMY OF A JOINT SOUND – USING JOINT ACOUSTIC EMISSIONS TO DIAGNOSE AND GRADE MUSCULOSKELETAL DISEASE AND INJURY

A Dissertation
Presented to
The Academic Faculty

by

Daniel C. Whittingslow

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Wallace H. Coulter Department of Biomedical Engineering

Emory University
Georgia Institute of Technology
December 2019

ANATOMY OF A JOINT SOUND – USING JOINT ACOUSTIC EMISSIONS TO DIAGNOSE AND GRADE MUSCULOSKELETAL DISEASE AND INJURY

Approved by:

Dr. Omer Inan, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Shelly Abramowicz
Division of Oral and Maxillofacial Surgery, Department of Surgery,
Emory University School of Medicine

Dr. Robert Butera
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Sampath Prahalad
Division of Pediatric Rheumatology,
Emory University School of Medicine

Dr. Young-Hui Chang
School of Biological Sciences
Georgia Institute of Technology

Date Approved: October 22, 2019

To my Friends, Family, and Fiancée. Tetelestai.

ACKNOWLEDGEMENTS

Before we begin this journey into the world of joint sounds, I would like to thank the many people that made it possible for me to perform this research.

First, I would like to thank my advisor, Dr. Omer Inan. From our first meeting to our latest conversation, he has served as a constant source of inspiration and support. The first time we met he saw potential in me that I did not yet realize. I told him I was interested in his lab but knew nothing of electrical engineering. He said, “I’ve got a whole lab of electrical engineers. You’ll bring something new to the table.” After rotating for three weeks in his lab, I knew it was the one for me. I’ll always remember our phone calls to brainstorm new ideas and directions on our projects. Traveling together to D.C. and Minnesota were two highlights of my PhD experience. It was during those times that I got to know him betterand see what life as a PhD researcher was like outside of the lab. Thank you for everything, Omer!

I would not be here today without the help of Dr. Sampath Prahalad and Dr. Shelly Abramowicz. You two helped give a clinical focus to my work, and taught me so much about what it takes to be a great physician and how to balance a clinical workload with research. Thanks as well to Dr. Robert Butera and Dr. Young-Hui Chang – your perspectives while serving on the committee helped balance out the clinical focus and ensured that I remembered my engineering roots.

Next, I would like to thank all my lab mates in the Inan Research Lab. The lab has grown tremendously since I first joined, and each new addition has brought a new friend

into my life. Beren Semiz and Dr. Sinan Hersek, thank you for helping me establish the JIA project in my first year. Venu Ganti, Hyeon-Ki Jeong thank you for all the hours of fun in the Cadaver lab. We were driven a little crazy in there, but I would not trade those days for anything! Samer Mabrouk, thank you for joining us in there (if only briefly) and always calling me "Doc". Nick Bolus and Andrew Carek, thanks for being my friends from day one. Our work did not overlap too heavily, but our lives certainly did. Our trip to the land and fight night took away from the drudgery of constantly listening to knees popping. Jon Zia, thank you for the machine learning lessons and taking up the mantle of MD/PhD student, without your joining I don't know that I ever could have graduated. It's up to you to teach everyone the physiology now! To Caitlin Teague, Md Mobashir Hasan Shandhi, Nil Gurel, Hewon Jung, Sevda Gharehbaghi, Jacob Kimball, Nordine Sebkhi, Brandi Nevius, Kristy Scott, and Asim Gazi, thank you for the many conversations, laughs, and lessons on the basics of electrical engineering for me. You have all helped me with everything from soldering to signal processing. Thanks to the awesome undergraduates that helped me do the things I least wanted to do: Nazli Goller, Adith Srivatsa, Hannah Kim, and Elizabeth Ham, if you all ever need anything don't hesitate to reach out. I'm excited to see your lives and careers develop. I'd like to give a special thanks to Lara Orlandic. As an undergraduate she knew more about everything I was trying to do than I did, and always taught me with kindness and humility. Good luck earning that PhD across the pond – you're going to do great things!

My acknowledgments would not be complete without me mentioning my friends, family and fiancée. To my MD/PhD cohort (notably Dr. Mariko Peterson, Dr. Myles McCrary, Dr. Davic Trac, Dr. Saumya Gurbani, Steve Yarmoska, Matt Stern), thank you

for sticking with me through all these years and for all the encouragement. We've been in the trenches for 6 years and are finally starting to climb out! I cannot wait to see where you all end up. You're the most interesting and inspiring cohort I could have imagined. To my lifelong buddies, Luke Donohue and Charlie Spalding, thank you for listening to me talk about things you may not have understood and always being there when I needed you. Thank you to Dr. Nathan Evans, without you I would have never gotten involved in research. To my mom, dad, Susan, Meme, Yaya, John Drew, Jessica, Drew, Sarah, Cathryn, Ayden, Regan, and Taylor. You all are my family. I knew that even at my lowest times I could lean on you and at my highs I could celebrate with you. Thank you for believing in me. And dear Dr. Jamie Puente, my fiancée, thank you for encouraging me to pursue my dreams while reminding me to take some time to breathe. I love the life we are building together and there is no one else I'd rather have at my side.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ACL	Anterior Cruciate Ligament
AE	AE
AUC	Area Under the Curve
Bagging	Bootstrap Aggregating
CT	Computed Tomography
DAQ	Data Acquisition Unit
FU	Follow-Up
HCW	Health Care Worker
IMU	Inertial Measurement Unit
IRB	Institutional Review Board
JIA	Juvenile Idiopathic Arthritis
LCL	Lateral Collateral Ligament
LOSO-CV	Leave-One-Subject-Out Cross Validation
MCL	Medial Collateral Ligament
MRI	Magnetic Resonance Imaging
MSK	Musculoskeletal
OMS	Oral Maxillofacial Surgeon
PCL	Posterior Cruciate Ligament
ROC	Receiver Operating Characteristic
RoM	Range of motion
TMJ	Temporomandibular Joint

XROMM	x-ray reconstruction of moving morphology

SUMMARY

Knee injuries and chronic disorders, such as arthritis, affect millions of Americans. Currently, diagnosis of these conditions relies on expensive, time-consuming imaging studies and physical examination by a health care professional. After diagnosis, there are few quantitative technologies available to provide feedback to patients regarding their rehabilitation or efficacy of their prescribed treatments. Most assessments are qualitative, relying on patient-reported symptoms, functional performance, or physical examinations.

To address that gap in the management of musculoskeletal (MSK) conditions in this work, I describe my progress in developing a joint health sensing technology focused on measuring the sounds – or acoustic emissions (AEs) - produced during a joint's movement. The goal of this research is to provide an easily interpretable, quantitative metric of joint health status that could be measured using affordable hardware. To develop such a metric, I had to first build a system for accurately and repeatably recording AEs, better understand the nature of AEs using a cadaver model, build a database of AEs from different clinical cohorts, and define a technique for the accurate analysis of AEs that would yield clinically meaningful results. All these steps led to my proposed AE analysis algorithm, which takes low-level AE signal and feature data and fuses it using machine learning to present a joint health score or index. This score is shown to closely track with the clinical condition the patient groups studied as well as the injuries in the cadaver model.

During the development of this metric, my work on a cadaveric model of joint sounds helped us better understand the underlying physiology and anatomy that contributes to the production of AEs. Building off of the cadaver AE work, I translated AE analysis

into the clinic and performed cross-sectional and longitudinal recordings in a pediatric patient population with juvenile idiopathic arthritis (JIA). In JIA, AEs of the knee and temporomandibular joint (TMJ) are explored – with an emphasis placed on longitudinally monitoring patients' responses to treatment. I envision that one day, AE monitoring could serve as a much-needed objective marker of joint health status. It could be a valuable tool for quantifying clinical studies, personalizing rehabilitation efforts after injury, tracking therapeutic efficacy, and ultimately reducing the burden of musculoskeletal injury and disease for both the patients and the health care system.

CHAPTER 1. INTRODUCTION

1.1 Motivation

Musculoskeletal injuries and chronic conditions affecting the joints are prevalent [1], and the knee is one of the most frequently injured body parts with nearly 18 million patient visits occurring per year in the United States [2]. Often, these injuries occur due to the repetitive stress and high loads experienced during movement [3], [4]. The knee is also one of the most commonly affected joints by chronic degenerative joint diseases, such as arthritis. Nearly 23% of the U.S. population is diagnosed with some form of arthritis. The high number of patients with chronic joint conditions and acute injuries imposes a severe burden on patients and the healthcare systems [1].

The gold standard for diagnosing and managing knee health is a combination of physical exams and medical imaging. Physical exams rely on health care worker expertise, and imaging is not always feasible due to its high cost, restriction to a clinical setting, and contraindications [5], [6]. After diagnosis and initial intervention, patients enter a period of recuperation and rehabilitation. Ideally, during this period medical recommendations could be individualized to each patient's own rate of recovery. However, imaging studies provide limited information during rehabilitation, and repeat imaging is often prohibitively expensive. Instead, rehabilitation decisions are based on subjective analysis, patient-reported symptoms, and assessment of functional activity levels [7].There is a need for a more sensitive and objective system for monitoring joint conditions that could complement these existing approaches.

One promising avenue for addressing this lack of objective, actionable data during the rehabilitation period is by using wearable sensors. Many researchers have developed technologies for assessing knee health using wearable sensors. Most often, inertial sensors have been used to measure gait and knee kinematics [8]–[12]. These measures have had limited success because gait and joint kinematics do not reflect underlying pathophysiologic changes until damage to the joints has sufficiently progressed [10], [13].These sensor-based approaches provide a step toward longitudinal monitoring; however, they do not precisely capture the subtle changes in physiology or structure that would be most useful for quantifying MSK rehabilitation. There is a need for additional, and more sensitive, markers of joint health for augmenting the inertial sensor-based approaches to long term monitoring.

AEs could potentially provide information regarding the structural integrity of joints and health of their internal surfaces before their kinematic outcomes are apparent. These sounds produced during joint articulation have long been of interest to physicians, though little is known about their nature. Most existing research in joint AEs has focused on developing diagnostic techniques to differentiate “healthy” vs. “unhealthy” knee joints [14], [15]. In one study, osteoarthritic knees were found to produce more frequent, louder, and longer duration AEs when compared against healthy knees [16]. This work was limited by the technology available at the time of publication. Until recently, longitudinal assessment using AEs in healthcare was not feasible due to a lack of technologies for recording AEs outside of a laboratory or clinical setting. In 2016, for the first time, our lab used miniature microphones for wearable AE measurements outside of those settings with high reliability and repeatability [17].

With the feasibility of measuring AEs using miniature microphones proven, research to understand what type of physiologic information AEs contain is necessary. Before clinical adoption of AEs is possible, we must better understand the nature of AEs. This understanding will involve studying the origin and confounding factors of AE production and propagation. It will also need to be demonstrated that there is a direct and preserved relationship between internal structural disruptions of the knee (e.g. ligament tear, cartilage breakdown) and changes in AE patterns. With this understanding, AEs may one day be able to serve as a much-needed quantitative biomarker of joint health.

1.2 Major Contributions of this Work

AE analysis has broad applications in the realms of MSK diagnosis, screening, and monitoring. There has been scientific and clinical interest in using AEs to diagnose and track joint health status since at least 1902 [18]. Unfortunately, the technology has not made much progress in terms of clinical adoption. For the past three years, I have been recording the AEs of patients in clinic and working closely with their treating physicians and medical team. During this period, I have perceived three major impediments to adoption of AE screening technology in clinic. The first is the lack of scientific research explaining the nature of AEs which makes interpreting them difficult, and thus limits healthcare providers' (HCP) trust in the findings. The second is the perceived difficulty in recording and analyzing these signals. The third is a lack of understanding of the value that AEs will provide to a clinical workup outside of the current gold-standards of diagnosis and treatment. My thesis work seeks to address these impediments. As such, below I present the three major contributions of this work.

The major contributions of this work are:

1. Quantified, for the first time, the specific characteristics of AE features that change following acute knee injury in a cadaver model.
2. Demonstrated that AE characteristics are significantly different in patients with juvenile idiopathic arthritis compared to healthy matched controls, and that these differences diminish following successful treatment.
3. Developed and validated a head-worn sensor package for enabling high quality acoustic emission recordings from the TMJ in clinical settings.

1.3 Thesis Organization

CHAPTER 2 describes the specific aims of my research. Before detailing the findings of my research, In CHAPTER 3 I provide at a cursory introduction into the anatomy, physiology, and pathologies that will be heavily referenced throughout this work. The technologies, sensors and analytical frameworks I employed in recording and understanding AEs are detailed in CHAPTER 4.

The rest of this work is intentionally organized to create a cohesive narrative rather than strictly a chronology of my work. As such, I first present in CHAPTER 5, insights gained from our cadaveric model of knee AEs. This model informed much of our later analytical frameworks from the two clinical studies: JIA diagnosis and tracking (CHAPTER 6) and TMJ diagnosis (CHAPTER 7). Finally, I conclude and discuss potential future directions for those that begin this work where I finished in CHAPTER 8.

CHAPTER 2. SPECIFIC AIMS

The goal of this project is to understand what role AEs can served in describing joint health. Our lab has designed a recording setup using uniaxial accelerometers for recording joint AEs that is discussed in 4.1. In our previously published work, this setup was used to quantify knee recovery following acute injury in a small cohort of athletes [3]. These findings led to our hypothesis that AEs contain clinically relevant information pertaining to the underlying health of MSK joints. This central hypothesis is tested with the following aims:

1. Improve the fundamental understanding of joint sound origins by using a cadaver model

Researchers have used a wide range of instruments and analysis techniques to detect and interpret the sounds produced during movement of the knee [19]–[21]. However, the ability to interpret these AEs has had limited success due to a lack of mechanistic understanding of how these sounds are produced. We hypothesize that the exploration of AEs using human cadaver knees will reveal crucial information about the source, propagation and confounders of these sounds. A cadaver model will allow for a high level of control and manipulation of AEs. The effects of joint damage on the recorded AEs will be investigated using the lab's custom AE recording setup, inertial measurement units (IMUs), biplanar motion capture x-ray imaging, and computed tomography (CT) scans. Recordings will be performed at baseline, after a sham surgery, and after injury to the joint. These AEs will then be compared against AEs recorded from patients with similar injuries to understand if the changes noted in a cadaver model could translate to the clinic. This aim will help deduce the mechanism of joint sound production, validate a cadaver model

of AEs, and support the claim that changes to the internal joint architecture are measurable at the surface of the skin.

2. To determine the potential of AEs for diagnosis and longitudinal monitoring of joint health status.

Clinical assessment and treatment of joint injury and disease relies heavily on qualitative physical exams, patient reported symptoms, and hospital-based imaging studies. An initial step necessary for the integration of joint AE assessment into an MSK work-up is to determine their diagnostic potential. This will be studied in a group of children with JIA compared against matched control populations. If we find that the AEs from this group contain consistent features that are unique between the pathologic and healthy groups, it supports their inclusion in the respective diagnostic workups. In addition to diagnosis, longitudinal monitoring is an essential part of these patients' clinical management due to the chronic nature of the disease and the long recovery time of injury, but it is difficult to longitudinally monitor progress without repeat clinic visits [22], [23]. Our lab's preliminary research has shown consistency in subjects' AEs overtime [24]. Therefore, we hypothesize that changes in knee AEs overtime may be used to longitudinally monitor knee health status. To test this hypothesis, I will record the AEs of patients with JIA at critical timepoints in their treatment (e.g. at diagnosis and 3-6 months post-intervention). This aim will quantify the ability of AEs to classify the severity of arthritis, discover if AEs can track therapy efficacy, and provide support for using AEs to longitudinally monitor MSK conditions.

3. To expand the scope of AE analysis beyond the knee by exploring its use in the TMJ

Another commonly affected joint in JIA with severe outcomes if not monitored and treated is the TMJ. The TMJ is difficult to examine, and diagnosis relies mainly on imaging [25], [26]. We hypothesize that in JIA, AEs from the TMJ may indicate joint involvement and serve as a measurable biomarker for inclusion in a clinical workup. A custom headset with embedded AE sensors will be built and used to record patients as well as age- and sex-matched controls. This aim will provide proof of concept of AEs applicability in other joints and elucidate the diagnostic/screening potential of this sensing modality for TMJ involvement in JIA.

CHAPTER 3. BACKGROUND

3.1 Contextual Anatomy and Physiology

3.1.1 The Knee

The knee is a hinge joint that contains the distal femur, the proximal tibia and fibula, the patella, and the joints between those bony structures. It is the largest joint in the body [27]. There are three articulating surfaces in the knee: two between the femur and the tibia and one between the femur and the patella. The two rounded condyles of the femur rest on the flat tibial plateau. The two tibiofemoral joints are formed by the curves of the medial and lateral condyles of the femur as they articulate with the condyles of the tibia. The third articular surface in the knee is the patellofemoral joint. The patella slides along the groove of the anterior aspect of the distal femur called the trochlear groove [27]. The knee joint relies on four ligaments for stability. There is little fat or muscle covering the knee which makes it highly vulnerable to injury [28].

The menisci and two pairs of ligaments (the collaterals and the cruciate) are crucial for the stability of the knee [29]. The medial and lateral menisci cushion the femur on the tibia. They are crescent-shaped fibrocartilaginous discs that create a cup like surface for the articulating femoral condyles. Notably, the medial meniscus connects to the medial collateral ligament (MCL) – a broad, flat ligament connecting the medial femoral epicondyle to the medial condyle of the tibia. This connection makes the medial meniscus more prone to traumatic injury than the relatively free lateral meniscus [30]. The lateral collateral ligament (LCL) connects the lateral femoral epicondyle to the lateral condyle of

the tibia. Together, the LCL and MCL provide medial and lateral stability to the knee. The anterior cruciate ligament (ACL) crosses from the anterior medial tibia to the lateral femoral condyle. It prevents the tibia from sliding anteriorly on the femur. The posterior cruciate ligament (PCL) crosses with the ACL inside the knee joint. It connects the posterior tibia and lateral meniscus to the medial femoral condyle. It prevents the tibia from slipping posteriorly in relation to the femur. The ACL and PCL are crucial to the anteroposterior stability of the knee. These anatomical structures can be seen in Figure 1.

The knee is a synovial joint. The synovial cavity of the knee occupies the two concavities at each side of the patella (i.e. the "negative infrapatellar spaces") and the space above it - the suprapatellar pouch. This joint cavity covers the anterior, medial, lateral, and parts of the posterior aspect of the knee [29]. The synovial membrane is a type of specialized connective tissue that lines the inner surface of the joint capsules of synovial joints. This membrane creates a tight barrier that keeps synovial fluid inside the joint. This non-Newtonian fluid minimizes friction during movement. In addition to reducing friction, the synovium and synovial fluid prove a plane of separation, so that movement can occur. The synovium is not normally detectable but can become swollen and tender when the knee is inflamed or injured (as is seen in JIA) [31]. A diagram of the synovium is in Figure 2.

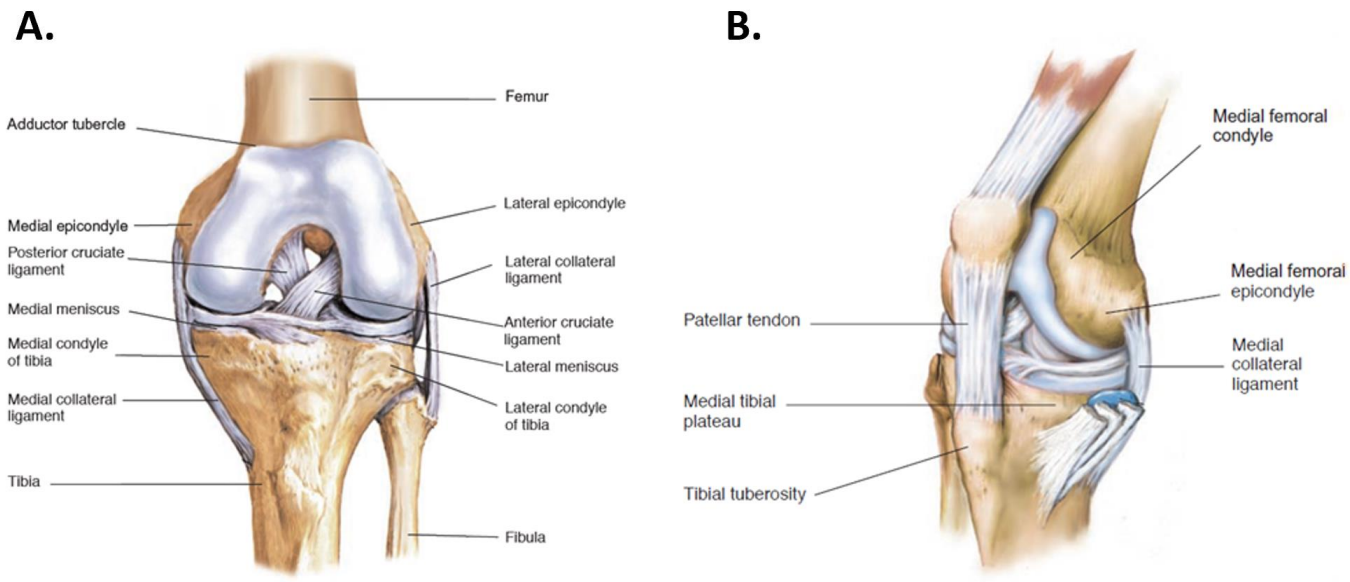

Figure 1. Knee Anatomy. Note the tibial tuberosity. That is the landmark used for determining placement of the contact accelerometers. Image courtesy of [29]**.**

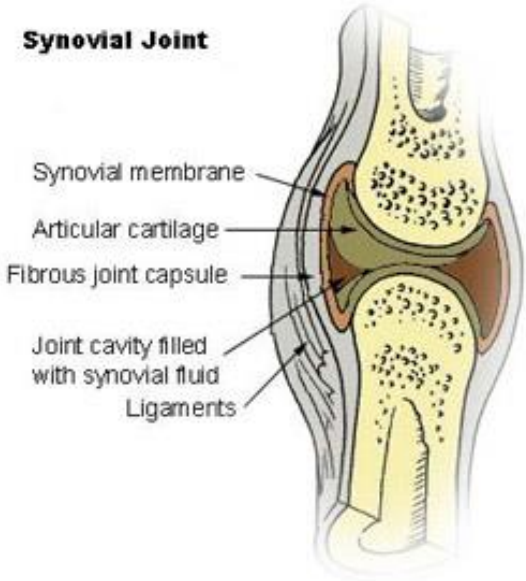

Figure 2. Synovial Joint. The synovial membrane is not normally detectable on physical exam but serves an important role in maintaining the synovial fluid balance which reduces friction in the joint. Image in public domain.

3.1.2 The Temporomandibular Joint (TMJ)

The TMJ is the most active joint in the body, opening and closing up to 2,000 times a day [5]. It connects the skull to the jaw. Particularly, the TMJ is formed by the fossa and articular tubercle of the temporal bone and the condyle of the mandible [27]. Like the knee, it is a condylar, synovial joint. There is also a meniscus (a fibrocartilaginous disc) in the TMJ. The TMJ's meniscus serves a similar function to the meniscus in the knee. It cushions the action of the mandibular condyle against the synovial membrane and capsule of the temporal bone. This meniscus divides the joint cavity into two small synovial cavities: the superior and inferior compartments. Retrusion and protrusion of the mandible occur in the superior compartment. Hinge movements of the mandible occur in the inferior compartment. The joint is reinforced by multiple ligaments connecting the mandible to the sphenoid and temporal bones and supported by the muscles of mastication. The size, depth, and coverage of the TMJ by surrounding tissue makes the TMJ particularly difficult to examine [32].

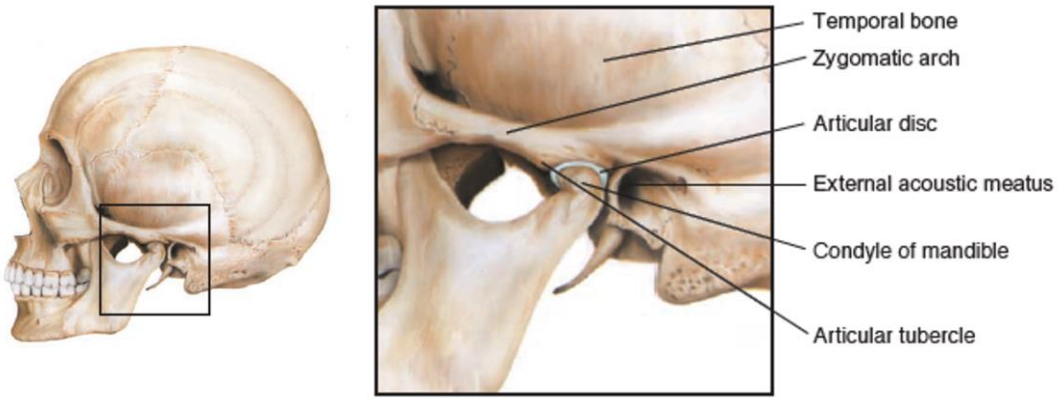

Figure 3. Temporomandibular Joint Anatomy. Image courtesy of [29]

3.2 Contextual Pathology and Injury

3.2.1 Juvenile Idiopathic Arthritis

Arthritis describes swelling within a joint, or limitation in the range of joint movement with joint pain or tenderness, which persists for at least six weeks, is observed by a physician, and is not due to primarily mechanical disorders or to other identifiable causes [33]. JIA encompasses all forms of arthritis that begin before a patient is 16 years old and are of an unknown origin. It is an autoimmune disorder characterized by persistent joint swelling caused by an accumulation of synovial fluid and thickening of the synovial lining (Figure 4) [34]. It is a leading cause of disability and the most common childhood chronic rheumatic disease with a prevalence of 150 cases per 100,000 [35].

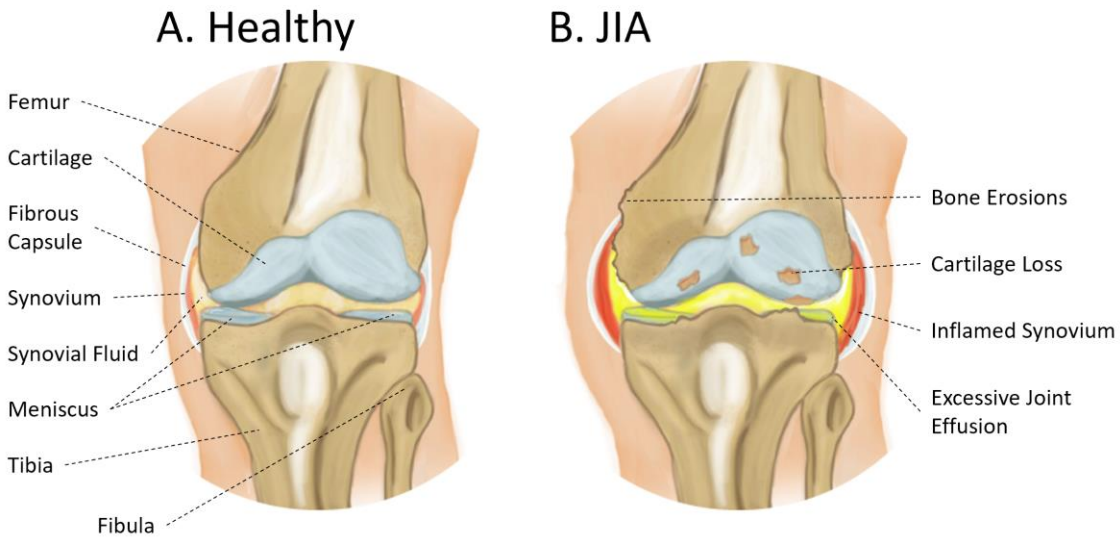

Figure 4. Depiction of healthy knee vs a knee with JIA. (A) A healthy knee articulates smoothly due to is smooth cartilage and appropriate amount/constituency of synovial fluid. In JIA, you may see cartilage loss, bone erosions, and a thickened/inflamed synovium with excessive joint effusions.

3.2.1.1 JIA Subtypes

JIA is highly variable with its presentation, symptomatology and course thought to be influenced by both genetic and environmental factors [34], [36]. The different forms of arthritis encompassed within JIA have been grouped based on clinical and laboratory features. These groups are systemic arthritis, oligoarthritis, polyarthritis, psoriatic arthritis, enthesitis related arthritis, and undifferentiated arthritis [33]. Systemic JIA is characterized by systemic features, such as fever, rash and serositis with arthritis in one or more joints[37]. Oligoarthritis affects one to 4 joints during the first 6 months of the disease. Polyarthritis affects 5 or more joints during the first 6 months of disease and is subdivided based on positive or negative rheumatoid factor (RF) tests. Psoriatic arthritis has features

of arthritis and psoriasis (dactylitis, nail pitting, onycholysis, or psoriasis in a first-degree relative). Enthesitis-related arthritis may have sacroiliac joint tenderness, inflammatory pain in the lumbosacral region, acute anterior uveitis, or a history of ankylosing spondylitis, sacroiliitis with inflammatory bowel disease, or Reiter's syndrome. Finally, undifferentiated arthritis contains all cases of all juvenile arthritis not covered by one of the other categories [33].

3.2.1.2 Diagnosis

No conclusive laboratory tests are available for the diagnosis of JIA [38]. Diagnosis is ultimately a diagnosis of exclusion, decided through a thorough history and physical exam. Taking a history involves asking questions about systemic manifestations (e.g. fever and rash), joint stiffness in the mornings, joint pain or swelling, and any history of auto-immune conditions in family members. The physical exam allows for assessment of pain, tenderness, swelling, limited movement, decreased strength, muscle atrophy and bony deformities. The physical exam also involves looking for lymphadenopathy, organ enlargement, rashes, nail abnormalities, or enthesitis [38]. The clinical features all depend on the subtype of JIA and differ based on the age of onset, number and location of joints involved, disease course and presence of antinuclear antibodies (ANA), RF, and uveitis [39], [40].

Finally, to fully subtype suspected JIA laboratory tests are needed. These include a full blood exam, inflammatory markers (erythrocyte sedimentation rate, C reactive protein), and autoimmune markers (RF, HLA B27, ANA). Imaging studies are also commonly used. Radiography can show narrowed joint-spaces or erosions as well as

growth abnormalities. Magnetic resonance imaging (MRI) can show the inflamed synovium and increased joint fluid [26], [39], [40].

3.2.1.3 Management and Treatment of JIA

Inactive disease followed by clinical remission with the goal of allowing the child to resume normal childhood activities and normal growth and development are the treatment goals for treating children with JIA [38]. Managing JIA requires a combination of pharmacological interventions, physical and occupational therapy, and psychosocial support. The three main goals of treatment are to: 1) lessen the pain and swelling during symptom "flares", 2) reduce the number of symptom flare-ups and put the disease into remission, and 3) prevent long-term joint damage [41].

Initial pharmacological treatment will depend on the severity of the presentation, but typically involves non-steroidal anti-inflammatory drugs (NSAIDS) and corticosteroid injections into the inflamed joint [42]. If those first line treatments are inadequate, treatment will move on to disease-modifying anti-rheumatic drugs (DMARDs). DMARDs can be divided into biologics and non-biologics. The most common non-biologic used to treat JIA is methotrexate. Low-dose methotrexate in combination with NSAIDs and corticosteroids is rapidly becoming the standard of care. Biologic DMARDs are antibodies that prevent the activation of various aspects of the inflammatory cascade or immune response – which results in less severe and less frequent flares. Anti-TNF drugs (e.g. abatacept) and anti-interleukin 1 and 6 antibodies are commonly used when symptoms are not controlled with more conservative measures [26], [42], [43].

3.2.1.4 Future Work in JIA

Clearly, the diagnosis and management of JIA is no small task. A foreseeable next step in the field will be to devise a method for combining the ongoing genetic/immunological mechanistic studies of the disease with observed clinical outcomes. In order to better subgroup patients, identify risk profiles, and predict an individual's response to treatment new technologies and modalities of objectively diagnosing and profiling JIA must be developed. One promising objective measure of disease activity is through joint AE assessment. With further research and development, AE sensing can hopefully lead to improved, personalized care, and help objectively quantify the results of the next generation of JIA clinical trials.

3.2.2 Acute Knee Injuries

3.2.2.1 Knee Injuries Addressed in This Work

The studies presented in this work focus most on acute ligament and meniscal tears. In CHAPTER 5, these types of injuries are analysed in an *ex-vivo* human cadaver model. An understanding of the current clinical standards of diagnosis and management is helpful to better understand the benefit that AE analysis may have in this clinical realm.

3.2.2.2 Diagnosis

3.2.2.2.1 Cruciate Tears

As discussed in 3.1.1, the ACL and PCL are the two cruciate ligaments that cross within the knee joint. The ACL provides 85% of the stability to the anterior translation of the tibia relative to the femur and restrains tibial rotation [3]. ACL tears are most commonly caused by non-contact pivoting injuries [44]. 54% of cases are associated with

acute lateral meniscus tears [44]. If left untreated ACL tears may lead to chondral injuries, unrepairable meniscus tears, and possibly arthritis [23].

Diagnosis of ACL tears begins with patient reported symptoms. This type of tear is four and a half times more common in females [45]. A patient commonly feels a large "pop", has pain deep in the knee, and immediate swelling. On physical exam, the physician may notice effusion or a quadricep avoidance gait (i.e. reduced active extension of knee). The Lachman's test is the most sensitive exam test, but the pivot shift, and KT-1000 tests can also aid in diagnosis [3]. Finally, MRI imaging is commonly used as the final step in diagnosis.

The PCL is the primary restraint to posterior tibial translation [3]. PCL tears make up 5-20% of all knee ligamentous injuries and are generally caused by a direct blow to the proximal tibia while the knee is flexed [46]. Isolated and combined PCL injuries are frequently underdiagnosed [47]. Patients with this injury may report instability and posterior knee pain[48]. The posterior drawer test is the most accurate physical exam maneuver for diagnosing PCL injury[46]. A confirmatory MRI is often ordered to finalize the diagnosis.

3.2.2.2.2 Meniscus Tears

Meniscus injuries are the most common indication for knee surgery, and patients with ACL deficiency are much more likely to suffer a meniscus injury. Medial tears of the meniscus are more common than lateral tears. However, in older patients a degenerative tear of the posterior horn of the medial meniscus is most common. In patients with acute ACL tears, a lateral tear is much more common. There are seven common patterns of

meniscus tears: longitudinal, bucket handle, oblique, radial, horizontal, complex and root tears. In our studies on the AEs related to meniscus tears we focused on radial and oblique tears. These two patterns are most closely associated with mechanical locking symptoms [30].

Diagnosis of meniscus injuries begins with patient reported symptoms. These may include pain localized to the medial or lateral joint line of the knee, mechanical symptoms such as locking or clicking, and swelling. Joint line tenderness is the most sensitive physical exam finding for meniscus tears. The examining physician may also notice effusions, and can perform three provocative tests: Apley compression, the Thessaly test, and the McMurray test [49]. During these provocation tests a palpable pop or click is a sign of a positive test. This finding clearly ties into AE analysis of the knee. The final stage of diagnosis is MRI imaging which is the most sensitive diagnostic test, but also has a high false positive rate [50].

3.2.2.3 Treatment and Management

Treatment of acute knee injuries can be either non-operative or operative. Non-operative approaches rely on rest, NSAIDs, and rehabilitation. Non-operative treatments are the first line of treatment for degenerative tears. Physical therapy and lifestyle modifications are commonly the first step in low-demand patients.

3.2.2.3.1 ACL

Nearly 400,000 ACL reconstructions are performed each year in the U.S.A [3]. To correct an ACL tear, either an auto- or allograft is threaded through the knee to replace the

torn ligament. An autograft, using the patient's own tissue, is considered the "gold standard" [51], [52]. Following surgery, patients are advised to use aggressive cryotherapy, immediate weight bearing, and full passive extension. Rehabilitation during the first 6 weeks after surgery focuses on exercises that do not place excess stress on the graft. The timing of when patients can return to play following surgery is largely dependent on psychological, demographic, and functional outcomes [53], [54].

3.2.2.3.2 PCL

The decision to operate on a PCL depends on the grade of the tear. If it's a grade 1 (partial) or 2 (isolated), protected weight-bearing and rehabilitation may be sufficient [55]. If the tear is more severe, operations may be performed with a grafted PCL in tandem with bony avulsion fracture correction or a high tibial osteotomy [56]–[59]. After the operation, the leg is immobilized in extension and protected from gravity. Rehabilitation is focused on strengthening the quadriceps [55].

3.2.2.3.3 Meniscus

In the meniscus, there are three main operative approaches: meniscal repair, partial meniscectomy, and meniscal transplantation [60]. A meniscal repair has 70-95% success, and is most successful when the tear is in the peripheral zone, has a vertical or longitudinal pattern, is 1-4mm in length, or is a root tear [61]. During a repair, vertical mattress sutures are used to seal the injury. A partial meniscectomy removes the torn portion of the meniscus and is performed when the pattern of tear is not amenable to repair (e.g. complex, degenerative or radial tears) [62]. Meniscal transplantations are performed in young

patients with near-total meniscectomies. The size of the allografted meniscus is essential for proper healing and is usually based on radiographs of the knee [63].

3.2.2.4 Opportunities for Advancement

The prevention, diagnosis, treatment, and rehabilitation of acute knee injuries are all areas of active research. The diagnostic capabilities of AEs could be included as part of the initial work-up prior to MRI imaging. In planning surgical approaches/techniques having a full visualization of the injury is still an essential piece of the treatment plan. Therefore, AE assessment is unlikely to completely subvert the use of MRI. AEs could however be used as a screening step prior to imaging which would greatly help reduce the time and resources spent unnecessarily imaging non-surgical cases. AEs may also be of particular use is in the prevention and rehabilitation realms of MSK injury. The small form-factor, and affordable hardware, makes AE assessment feasible outside of the clinic. This type of assessment could easily be performed by a physical therapist, sports trainer, or a patient at home or in the field. After injury, the AE feedback could help personalize the rehabilitative efforts to optimize recovery time. It may even one day be possible to predict if an injury is likely to occur – perhaps from overuse – if AEs are regularly monitored during intense physical activity. With further research and development, these technologies can hopefully lead to improved, personalized care that helps reduce the tremendous burden of MSK injury on patients and the healthcare system [64].

CHAPTER 4. SENSORS AND ANALYTICAL FRAMEWORKS

4.1 AE Sensing Technology

In 2016, our lab published an extensive comparison of the different techniques for recording MSK AEs in the knees [17]. In this work, Teague et al. compared two classes of airborne microphones (electret and MEMS) and a piezo film contact microphones. They found that joint sound measurements from air microphones were repeatable in subjects and across the two joints. Most of their contact microphone recordings were corrupted due to noise from the interface technique, but several benefits of contact microphones were still discovered in this preliminary work. One main benefit of using contact microphones was that they did not record background noise. The main drawback was that the sensor-skin interface could be lost in a wearable setting. However, for my work, adapting the recording setup for a wearable setup was a secondary concern. The data presented in this thesis was recorded in a variety of clinics and laboratories with different levels and types of background noise. The contact microphones presented a clear advantage given the many recording locations. Additionally for TMJ AEs, it had previously been shown that contact accelerometers provide the highest mean amplitude in the time domain waveform [65]. Many of the features that appeared to differentiate clinical groups based on our proof-of-concept AE recordings were related to the amplitude of the spikes in the time domain waveform (e.g. RMS power, energy, b-value). Ultimately, either approach to recording AEs would have been suitable depending on the application. In the work presented in this thesis, I exclusively used surface mounted accelerometers to record AEs because they were

easy to place on the joints, able to capture high amplitude spikes in the time domain, and were resistant to background noise.

The AEs in all studies were recorded using Dytran uniaxial, miniature accelerometers (Model 3225F7, Dytran Inc, California, USA 91311) with a diameter of 6.35 mm. They are highly sensitive to changes in acceleration (sensitivity is 10.2 $mV/m/s^2$) and the frequency response curve is flat from 2 Hz to 10 kHz. The accelerometers were connected to a data acquisition device that enables the simultaneous and synchronous capture of up to four accelerometers at a rate of 100 kHz. The sensors and the data acquisition device (DAQ) are plugged into a laptop that powers the devices and is running a custom data-capture program written in MATLAB (Figure 5A). This program controls the length of the recording and converts the voltage readouts from the sensors to units of acceleration (using the manufacturer-provided calibrated sensitivities of the specific microphones). The program also performs preliminary steps to ensure that the data are successfully recorded including bandpass filtering (between 250 Hz – 20 kHz) and plotting the recordings.

With the recording hardware built and the software acquisition program written, the two main design decisions remaining were how to surface mount the accelerometers, and where to place them around the joint. A strong skin-to-sensor interface is essential for capturing AEs using contact accelerometers. In the cadaver model, we used sutures and superglue to adhere the accelerometers to the skin. With this "gold-standard" adhesion technique, we could better appreciate the joint sounds without adding interface noise. Upon listening to the AEs, we found that double sided sticky pads were a suitable means of attachment during our human subject studies. Particularly, Rycote brand sticky pads introduced very little "rubbing" noises into the recordings. They also provided a strong

interface with the skin, were quick to place, and did not irritate our pediatric patients. The location of the microphones in this work was selected based on the underlying anatomy with the guidance of the clinical mentors on the team. We opted to place one accelerometer approximately 3 cm medial and lateral to the distal patellar tendon (Figure 5B). This location had minimal anatomical structures impeding the path from the articulating surface of the knee joint (where we hypothesized was the source of the sounds) to the surface of the skin.

With the recording apparatus built, and the method for placing and recording AEs in place, it was then time to begin recording AEs in our populations of interest. The next section details the techniques we developed for analyzing these signals.

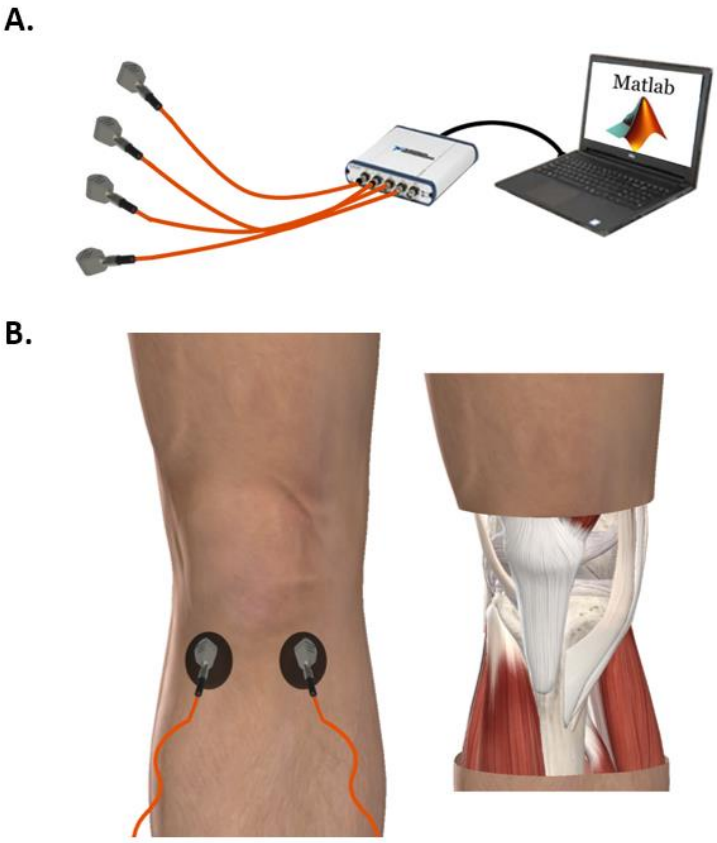

Figure 5. Recording Setup and Microphone Location. (A) Up to 4 accelerometers could be simultaneously recorded using the DAQ and custom Matlab recording script. (B) The ideal recording location in the knee was selected based on the minimum impedance between the surface location and joint line (B). Image courtesy of Complete Anatomy.

4.2 Analytical Frameworks

All analysis is done using Matlab (MathWorks, Natick, MA).

4.2.1 Signal Pre-Processing

The raw recorded signals in each experiment are first processed using a digital finite impulse response (FIR) band-pass filter with 100Hz - 10kHz bandwidth to maintain emissions in the audible range while removing motion artifacts. This frequency range was previously shown to contain the majority of knee AEs information [17]. Once filtered, the signals are compared against the simultaneously recorded IMU motion data and the beginning and ends of the AE signal are trimmed to remove the excess periods of noise before and after the flexion/extension movements began. This trimmed noise is used as a basis for a noise suppression algorithm using spectral subtraction from the AE recordings in several of our studies [66].

4.2.2 b-Value analysis of Acoustical Data

The first approach we pioneered for interpreting AEs for the work presented here was by calculating and comparing the b-value. The b-value metric is computed for AEs to differentiate the sounds based on their amplitude distributions. The b-value was first proposed by Gutenberg and Richter in earthquake seismology to quantify a logarithmic relationship between the magnitude and frequency recorded in a seismic trace, using the empirical formula expressed in Equation 1 [67].

$$\log(N) = a - bML \quad (1)$$

Equation 1. Logarithmic Relationship between Magnitude and Frequency

Where ML is the Richter magnitude of events, N is the number of events with magnitudes greater than ML, and a and b are the constants. Based on this relationship, the b-value is the negative gradient of the log-linear AEs frequency/magnitude plot and thus represents the slope of the amplitude distributions. Our previously published work successfully used the computed b-value of joint sounds to differentiate knee injury status in athletes [68]. An example calculation of the b-value is shown in Figure 6.

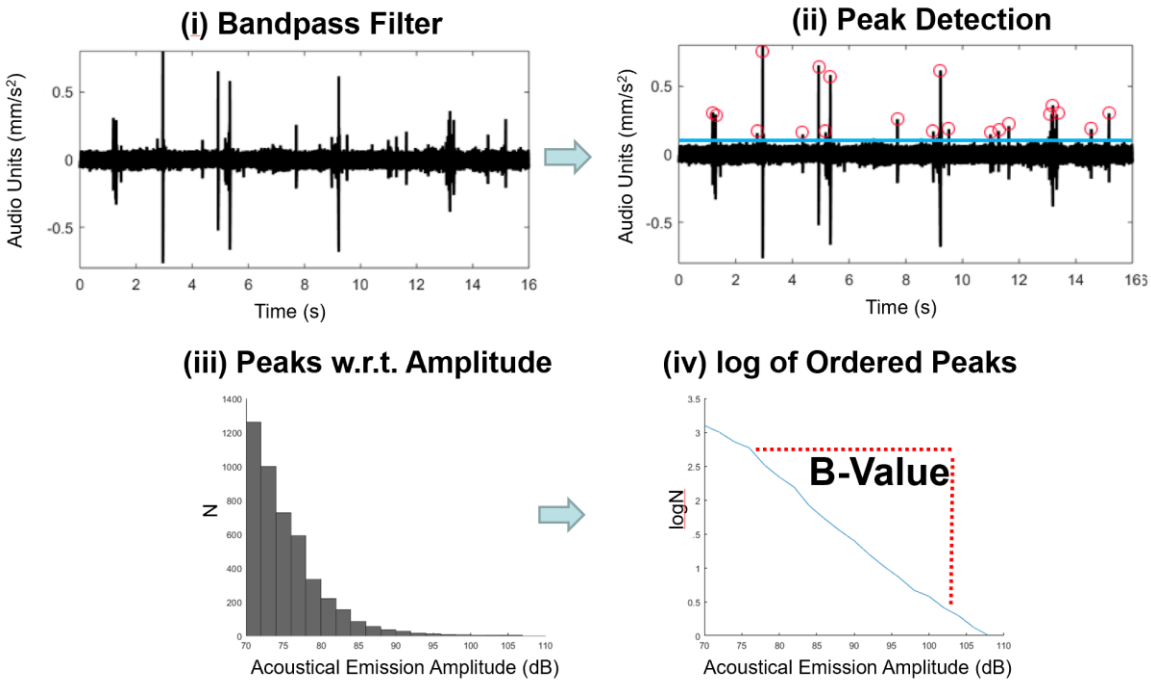

Figure 6. b-Value Calculation. (i) Each microphone's recorded acoustic signal is bandpass filtered. (ii) The peaks are detected above a threshold (RMS Power + noise

maximum). (iii) The peaks are ordered based on amplitude, and (iv) the log of the ordered peaks is taken. The slope of the line in (iv) is the b-Value.

4.2.3 Logistic Regression

The b-value analysis was useful for several of the studies in this work; however, when the changes in AEs were more subtle, b-value analysis was found to not be robust enough to fully characterize the AE characteristics. To make the analysis more robust, we calculated up to 49 features for each cycle of joint articulation for each microphone, and organized them into a feature matrix. A machine learning approach known as logistic regression was used to classify the input signals based on that feature matrix. Logistic regression is a common machine learning technique borrowed from statistics for binary classification problems. At the core of this algorithm is the logistic function, which was originally developed by ecologists to describe population growth – it is a sigmoidal curve that rises quickly and levels off at a given environment's carrying capacity [69], [70]. The algorithm uses this function to map any real number input to a value between 0 and 1.

$$1/(1 + e^{-value}) \tag{2}$$

Equation 2. Logistic Function

In logistic regression, the input values (x) are combined linearly to predict an output value (y) using weighted coefficients (b_n) that are calculated during training. These coefficients describe the n-dimensional hyperplane that best separates the two classes [71]. Unlike linear regression, in logistic regression the output values being predicted are binary

(0 or 1, or in our case healthy or diseased/injured). The logistic regression equation thus takes on the following format:

$$\mathbf{y} = \mathbf{e}^{\mathbf{b_0+b_1x_1+\cdots+b_nx_n}} / \mathbf{1} + \mathbf{e}^{(\mathbf{b_0+b_1x_1+\cdots+b_nx_n})} \quad (3)$$

Equation 3. Logistic Regression Mapping Function

Where y is the predicted output, b_0 is the intercept, b_1-b_n are the coefficients for the input feature values (x_1-x_n). x corresponds to one feature from one cycle of movement within the larger feature matrix. The individual features used are described in 4.2.5. Each of the features in the input feature matrix are given a coefficient learned through training (b_1-b_n). The vector of b_1-b_n are stored in the coefficient vector (β). β is found using a maximum-likelihood estimation, specifically the quasi-Newton method, that minimizes the error of the predicted probabilities [71]–[73].

The predicted output (y) is the probability that a given feature input belongs to the selected default class. A classification label for each feature is assigned using a probability threshold. In this work, the threshold is described in Equation 4. This threshold was chosen heuristically, but theoretically could be adjusted to increase the sensitivity or specificity depending on the application of the algorithm in the future. Once each of the feature classifications has been determined, the cycle that those features describe is labeled based on the majority of individual feature labels. Similarly, a subject-level classification label is determined by the predicted class of the majority of that subject's component cycles.

$$mean\big(p(x)\big) \leq 0.5, y = \text{Healthy} \quad (4)$$

$$\text{mean}(\text{p}(\text{x})) > 0.5, y = \text{Injured or Diseased}$$

Equation 4. Threshold for AE Classification. p(x) is the output probability of feature 'x' belonging to the selected class. *y* is the label assigned to that feature.

4.2.4 Decision Trees and Bagged Trees

At times when both the b-value analysis and logistic regression provided sub-optimal separation of the groups being studied, we moved to another common machine learning algorithm: the decision tree. The decision tree algorithm like logistic regression is a type of supervised learning, but it can classify more than two classes in a single model. In this algorithm, a tree representation is generated to solve a classification problem. Each leaf (node) of the tree corresponds to a class label, and attributes of the splits are represented on the internal nodes of the tree. See the diagram in Figure 7.

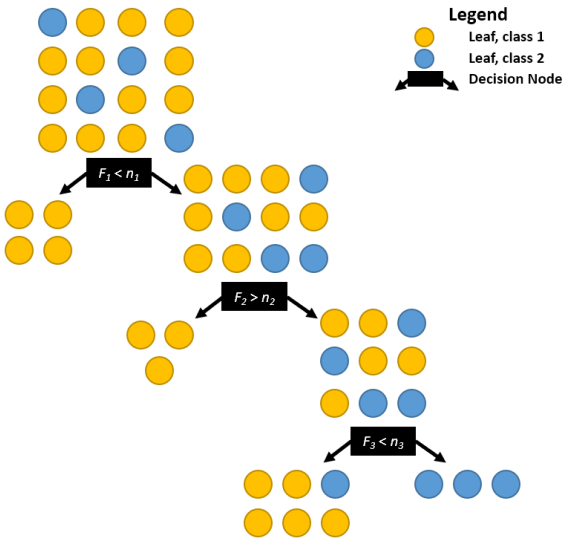

Figure 7. Classification Tree Diagram

There are several benefits of using a decision tree in classifying data. They are very easy to interpret. They handle missing data and outliers well which requires less data pre-

processing. They can handle nominal and ordinal variables, and they can model non-linear relationships. However, as the complexity of the branching increases, the model becomes increasingly difficult to interpret and may overfit the training data – resulting in poor generalizability [74].

In this work, rather than relying on a single decision tree, we performed bootstrap aggregation or bagging, to make the classifier more robust [75]. Bagging improves performance by combining multiple trained decision trees and taking a consensus vote before determining a classification. Each component decision tree is trained on a training set sampled uniformly and with replacement from the original dataset. Bagging helps reduce overfitting, reduces variance, and improves stability and the generalizability of a model. Figure 8 demonstrates the classification process in an ensemble bagged tree decision model.

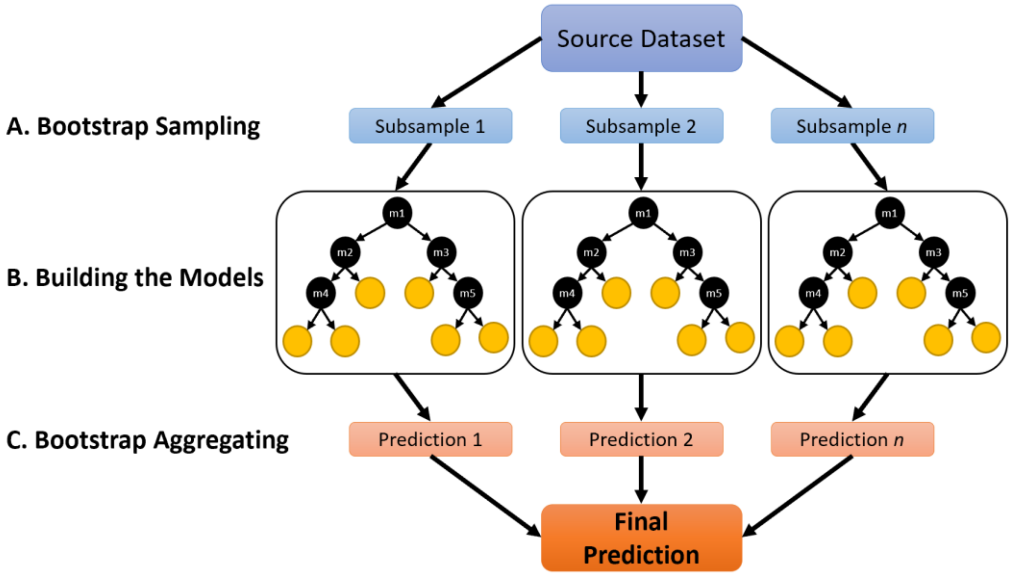

Figure 8. Bagged Tree Diagram. A) During bootstrap sampling a percentage of the source data is sampled with replacement and arranged into *n* subsamples. B) A decision tree is constructed on each of the subsamples based on a random set of *m* features. C) Results from the constructed trees are averaged to make a final prediction.

When the bagged tree model is used, information gain at each decision split is used to quantify the importance of that feature. Information gain is calculated as the decrease in entropy after a dataset is split on a feature. In an ensemble method, where many trees are simultaneously developed, the average entropy decrease across all trees is used. This average information gain describes the impact that each feature had had on the model's performance and is thus a useful proxy of feature importance [74].

4.2.5 Audio Feature Dictionary

The following features are calculated for each accelerometer during each cycle of flexion/extension for use in the classification algorithms. These features thoroughly describe both the time-domain and frequency characteristics of the recorded AEs.

Table 1. Features used in machine learning model. Each feature had there mean, median, and standard deviation used in the feature matrix as appropriate. Descriptions found at [76], [77].

#	Feature Name	Description
1	Mean Frequency	Center of the power distribution across all frequencies of a signal.
2	RMS Amplitude	Root mean square (RMS) of the signal amplitudes.
3	Zero Crossing Rate	The rate of sign-changes of the signal.
4	Energy	Square of the time-domain signal amplitudes, corresponds to loudness of signal.
5	Fundamental Frequency	Pitch of a signal - for a harmonic signal it's the frequency such that its integer multiple best explains the content of the signal spectrum.
6	Spectral Centroid	The barycenter, or "center of gravity", of the frequency spectrum
7	Spectral Spread	Standard deviation around the spectral centroid.
8	Spectral Crest	Ratio of the maximum of the spectrum to the arithmetic mean of the spectrum , indicates peakiness of spectrum.
9	Spectral Decrease	Represents the amount of decrease of the spectral amplitude – related to perceptual studies.
10	Spectral Entropy	Measures the disorder and peakiness of the spectrum
11	Spectral Flatness	Measures the noisiness to sinusoidality of a spectrum.
12	Spectral Flux	Measures the variability of the spectrum over time
13	Spectral Kurtosis	Measures the flatness, or non-Gaussianity, of the spectrum around its centroid.
14	Spectral Rolloff Point	Measures the frequency so that 95% of the signal energy is contained below that frequency
15	Spectral Skewness	Measures of the asymmetry of the frequencies around the centroid
16	Spectral Slope	Measures the amount of decrease of the spectral amplitude.
17	Harmonic Ratio	Measures the amount of energy in the tonal part of the signal compared to the total energy.

4.2.6 Leave-One-Subject-Out Cross Validation

Cross validation is a model evaluation method that gives an indication of how well a trained model will do when it is asked to make new predictions for data it has not already seen. This is achieved by not using the entire data set when training a learner. Specifically, in leave-one-subject-out cross validation (LOSO-CV), all rows in the feature matrix corresponding to a specific subject are left out of the training phase. Then, after training is done, that subject's data is used to test the performance of the trained classifier [78]. This process is repeated for each subject in the dataset and the accuracy of the model is the average of the predicted-label accuracies during each fold of testing.

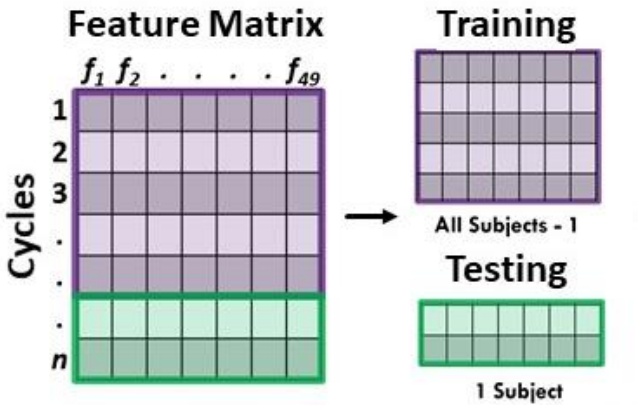

Figure 9. Leave-one-subject-out Cross Validation. In LOSO-CV, the feature matrix is split into a training and testing set. This splitting is iteratively performed to assess the accuracy of the model with each subject left out.

CHAPTER 5. AE ANALYSIS IN A CADAVER KNEE MODEL

5.1 Overview

The assessment of joint health is a long-standing issue in the management of musculoskeletal injuries. AEs could serve as a biomarker for joint health assessment, but their use has been limited by a lack of understanding of the mechanism of their production and propagation. In this work, we investigate AEs using an injury model in human lower limb cadavers and relate AEs to joint kinematics. This relationship helps us better understand the nature of these AEs. Using our custom joint sound recording system (described in 4.1), first, we recorded the AEs from 9 cadaver legs in four stages: at baseline, after a sham surgery, after a medial meniscus tear, and post-meniscectomy. We compare the resulting AEs using their b-values as described in 4.2.2. We also explore the relationship of swelling on AE production. Finally, we compare joint anatomy and kinematics to the AEs using the X-ray Reconstruction of Moving Morphology (XROMM) technique. XROMM analysis showed a close correlation between the minimal inter-joint distances and a large increase in the AEs. In our next cadaver study, we analyze the change in AEs from a lateral meniscus tear and determine if AEs can determine the severity of a meniscal tear. This work provides key insight into the nature of joint AEs and details a novel technique and analysis for recording and interpreting these biosignals.

5.2 Introduction

The knee is one of the most frequently injured body parts, with 18 million knee related patient visits occurring per year in the United States [2]. However, the number of acute knee injuries pales in comparison to the number of people suffering from chronic joint diseases such as osteo- and rheumatoid arthritis. It is predicted that by 2040 26% of the overall population will be diagnosed with some form of arthritis in the United States [79]. This prevalence coupled with the severe reduction in quality of life presents a significant burden on patients and healthcare systems [1]. Currently, clinical assessment and treatment relies on qualitative mobility assessments, patient reported symptoms, and imaging studies. A suitable marker of knee joint health that is quantitative, and measured with affordable hardware, could reduce this burden on healthcare systems and greatly improve patient outcomes and quality of life.

One such possible marker of knee joint health is the AE signature produced by knees during movement. These joint sounds have been explored as a means of assessing the joint's structural health since at least 1902 when Blodgett reported on auscultating the knee [18]. Since then, researchers have employed a wide range of instruments and analysis techniques to detect and interpret the sounds produced during movement of the knee. These findings have often led to attempts to diagnose joint conditions [20], [80]. However, the ability to interpret AEs from the knee for clinical decisions has had limited success. One of the main reasons for this is a lack of understanding of how these joint sounds are produced and what factors influence them.

In this work we investigate joint AEs using a human lower-limb cadaver model to address this knowledge gap in the field. This model allows for highly controlled analysis of the joint sounds from the knee in a reproducible and anatomically relevant manner. To record

the AEs, the limb is passively flexed/extended through its range of motion with contact microphones attached (as described in 4.1). The AEs produced during this movement are recorded.

To better understand the source of these AEs, we explore the relationship between internal contact of the articulating structures within the knee and AE production. This is done using a system of IMUs, biplanar motion capture x-ray imaging, and computed tomography (CT) scans. The output from that system is synced with the AE data. We created a medial meniscus injury to understand how alterations of the underlying anatomy can correlate with the AEs recorded on the surface of the knee. Combining literature on internal joint pressure, our findings of minimum articulating surface distances, and joint sounds at each stage of injury led to our proposed model of joint sound production (Figure 10). To provide more physiologic context to the model, we next emulated the biomechanical alterations associated with swelling following an acute injury by serially injecting saline into the joint capsule. The b-value of the AEs was calculated at each stage of testing as described in 4.2.2 [68]. During our next phase of cadaver testing, we arthroscopically performed serial cuts to the lateral meniscus. The impact of surgical approach and ability to quantify severity of cut is explored.

This chapter presents the first time that an analysis of knee AEs has been performed on a controlled, cadaver model with associated incorporation of anatomical complexity, confounding physiological factors that occur in an injured state (i.e. swelling), and specific structural changes in the knee. Our findings allowed us to propose a model of knee AE that better localized the source of these sounds while remaining consistent with the prior literature's findings that these sounds can be useful in classifying the health status of a knee

[68], [81]. If characteristic alterations of these AEs can be linked with knee health status, joint sounds may offer a biomarker for early detection and assessment of musculoskeletal injury.

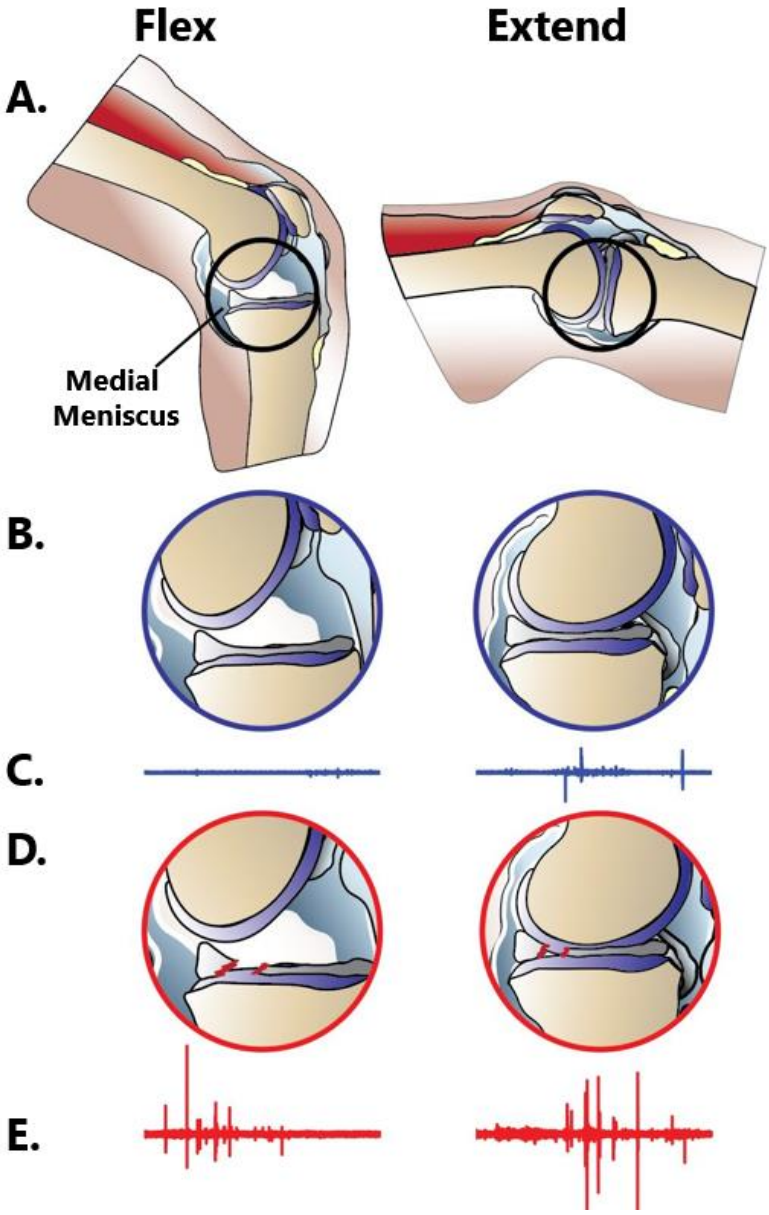

Figure 10. Concept model of knee acoustic wave creation before and after a meniscus tear with representative acoustic wave forms. A. Diagram of the knee during flexion and extension. B. Medial femoral condyle compressing the medial meniscus from flexion to extension. C. Representative acoustic waveform produced by the knee's movement. D. Compression of the radially torn, medial meniscus from flexion to extension. E. Representative acoustic waveform produced by the knee.

5.3 Effect of Medial Meniscus Tear, Swelling, and Joint Distance on AEs

5.3.1 Materials and Methods

5.3.1.1 Cadaver Specimen Procurement and Preparation

Experiments were conducted on 9 fresh, frozen human cadaver lower-limbs. The specimens were procured from MedCure, Inc (Orlando, FL, with permission for use in a research experiment), had an average age of 63.6 ± 9.5 years of age, stored at -20°C, and thawed to room temperature in a water bath for 8 h prior to testing. The age of these cadaver specimens may not be fully representative of the overall population, but the exclusion criteria helped limit the impact of confounding comorbidities. The joints were selected from donors with no known arthritis, injuries or past surgeries of the knee, and that were mobile at time of death. Prior to use, the legs were clamped to the laboratory benchtop and preconditioned with manual flexion/extension movements for five minutes.

5.3.1.2 Knee AE Setup and Acquisition

Two uniaxial analog accelerometers (3225F7, Dytran Instruments Inc. Chatsworth, CA) were sutured (4-0 Nylon Kit, Your Design Medical, Brooklyn, NY) 2 cm medial and lateral to the patellar tendon. These accelerometers and anatomical location are described in 4.1.

To record the knee AEs, the cadaver legs were suspended on the side of a lab bench and passively flexed and extended through their full range of motion (~90° to 180°). This suspension ensured the cadaver limb did not contact the surface of the lab bench at any stage of the motion. To pre-condition the leg prior to AE recording, it was flexed/extended through its full range of motion for 5 minutes at a rate of 1 cycle every 4 seconds. After

pre-conditioning the AE recording began. The legs were extended for two seconds, and then flexed through the same range for two seconds. The recordings contained a total of ten flexion/extension cycles with 5 seconds of background, environment noise recorded before and after the exercise for a total recording time of 50 s. An IMU (MPU6050, TDK InvenSense, San Jose, CA) was attached 5 cm proximal to the ankle and used to validate the joint angle and rotational velocity during these exercises. The signals from the accelerometer were sampled at 100 kHz and recorded using the DAQ described in 4.1. The recording setup is depicted in Figure 11.

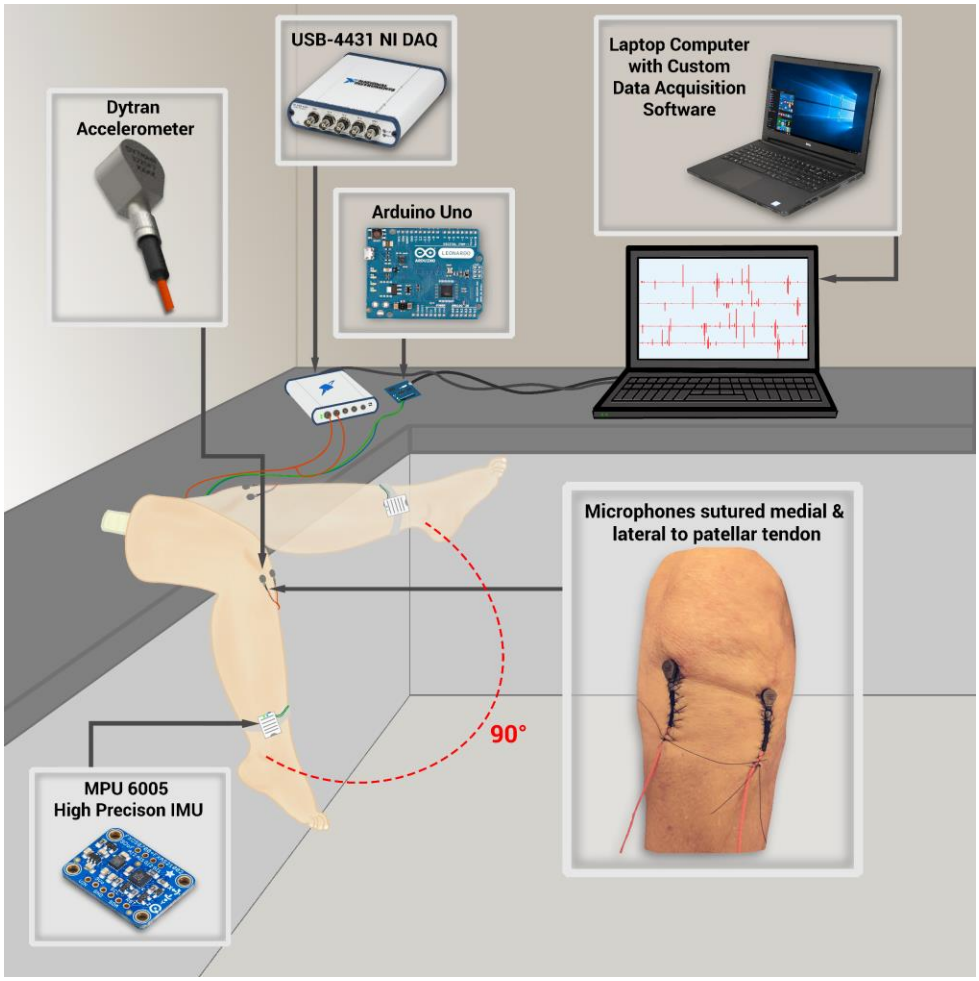

Figure 11. Testing setup for the generation, acquisition, and analysis of knee AEs on a cadaver model. The cadaver knee is outfitted with two accelerometers and a high-precision IMU. The accelerometers are sutured medial and lateral to the patellar tendon and record the surface vibrations (AEs) created by the manual flexion/extension of the leg. The IMU captures and syncs the 3D motion data to the joint sounds providing anatomical relevance to the recorded signals. A DAQ captures the audio waveform data and a microcontroller captures the IMU data. All data is transmitted to a laptop computer with custom acquisition and analysis software written in MATLAB.

5.3.1.3 Tear Protocol

Each of the knees (n=9) were serially, surgically altered in four stages to isolate the effects that a medial meniscus tear has on the joint's AEs. The four stages of testing were baseline, sham surgery, meniscus tear, and the meniscectomy. After thawing and pre-conditioning, the joint sounds were first recorded at their baseline status. Next, a sham surgery was performed on the leg. The sham surgery was performed with the knee at 90° of flexion with a 5-cm oblique incision made just posterior to the superficial MCL at the level of the vastus medialis curving over the medial epicondyle onto the anteromedial aspect of the tibia. This cut exposed the interval between the posteromedial joint capsule, semimembranosus, and medial head of the gastrocnemius [82]. Next, the posteromedial joint capsule was cut 2 cm to expose the medial meniscus. Without damaging the meniscus, the incisions were closed with simple continuous, running sutures [83]. The sounds were recorded at this sham surgery status. Next, the meniscus tear was introduced. The sutures were cut to re-expose the meniscus and a 10mm transverse (radial) incision on the posterior (zone A) portion of the meniscus was performed. The surgical entry path was again sealed with a simple continuous running suture and the sounds were recorded. Finally, a meniscectomy was performed on the injured meniscus. The sutures were cut to re-expose the meniscus and a 5mm margin anterior and posterior to the transverse/radial meniscus cut was surgically removed. The incisions were resealed and sounds re-recorded.

5.3.1.4 Saline Injection Protocol

To emulate the altered mechanical environment within the knee resulting from swelling following acute injury [48], [84], varying levels of saline were injected into the knees prior to meniscus surgery (n=5). A superolateral approach into the suprapatellar pouch was used due to its reliability as a route of entry into the knee joint and the obstruction of the attached microphones impeding other approaches [85]. The leg was fully extended and a 1.5 inch 25-gauge needle was inserted underneath the superolateral surface of the patella and directed posteriorly and inferomedially into the knee joint. 5 mL aliquots of saline were serially injected from 0 to 50 mL. After each injection, the joint sounds were recorded using the above AE acquisition protocol.

5.3.1.5 Acoustic Data Pre-Processing

Noise was trimmed from the beginning and ends of the recordings and a bandpass filter was applied as described in 4.2.1.

5.3.1.6 b-value Analysis of Acoustic Data

The b-value metric was computed for the AEs to differentiate the sounds based on their amplitude distribution of the AEs. This calculation is described in 4.2.2.

5.3.1.7 Acoustic Data Statistical Methods

The mean and standard deviation of the b-values were calculated for each dataset. The data were assessed for normality using a Lilliefors test. It was found that the groups were non-normal, so the Scheirer-Ray-Hare extension of the Kruskal Wallis test was performed. This test is often used as a non-parametric equivalent to the two-way analysis of variance (ANOVA) test. Finally, multiple Wilcoxon signed rank tests were performed to compare

between the data groups. A Bonferroni correction was applied to correct for the multiple comparisons. The same series of tests were performed on the saline injection data.

5.3.1.8 Joint Distance Imaging

The geometry of the tibial plateau is complex and asymmetric. In order to calculate the distance between the femur and tibia during articulation we used a two-part imaging protocol relying on a high-speed, biplanar x-ray video (100 fps, 79° between beam angle) and a computed tomography (CT) scan of the patella, tibia, and femur. Three 1mm diameter, radiopaque, tantalum markers (X-Medics, Frederiksberg, Denmark) were implanted into each bone on the posterior-medial, posterior-lateral, and anterior aspects of the bone. These three markers appeared in both the CT scan and x-ray videos. The CT scan allowed for a 3D model of the bones to be constructed. These markers were tracked in the x-ray videos using XMALab (Brown University, Providence, RI) and aligned with the markers on the CT-derived 3D model using Maya (Autodesk, San Rafael, CA) in order to track the complex interaction between the bones. This motion tracking technique is known as XROMM [86], [87]. With the articulation of the bones fully visualized, the distances between each of the 7076 vertices of the triangles making up the 3D mesh of the tibial plateau and its nearest femoral counterpart were calculated using custom Python scripts.

5.3.2 Results

In this study, the effects of a meniscus tear on the AEs produced by manually articulating a human cadaver knee are explored. The results from these tests are presented below. The b-value is the principle metric for comparison. It is a unitless metric that describes the slope of the amplitude distribution of an acoustic signal.

5.3.2.1 Sham surgery on a cadaver model of joint AEs does not significantly alter the AEs from a baseline state.

A sham surgery was performed to expose the medial meniscus of the cadaver. All the successive layers from the skin to the joint capsule were surgically resected *(*Figure 12A/E to Figure 12B/F*)*, and the AEs were recorded. Qualitatively, the time domain sound signature at this stage of testing appears very similar to the baseline state *(*Figure 12I, J*)*. The b-value statistic of the joint sounds at baseline was 1.99±0.54. After the sham surgery, the b-value dropped to 1.87±0.40. This shift was not statistically significant ($p=0.25$). This lack of statistical significance indicated that the sham surgery, with its alteration to the tissue external to the joint cavity and exposure of the joint capsule to the air and laboratory atmosphere, had minimal influence on the AEs of the knee.

5.3.2.2 Introducing a meniscus tear significantly alters the AEs from the sham state.

A full width, radial tear was performed on the posterior, medial meniscus *(*Figure 12 B/F to Figure 12 C/G*)*. After closing the resection, the AEs were again recorded and analyzed. At this stage, the AEs appear much more chaotic, with several large spikes in the amplitude of the sounds. This increase in amplitude was reflected in the b-value after the meniscus tear (b-value = 1.33±0.15). This drop in the b-value was significant when compared to the baseline and sham stages ($p=0.0039$), indicating that the meniscus tear was responsible for the change seen in the AEs. It also indicates that knee AEs can describe the internal environment of the knee.

5.3.2.3 Further removal of the meniscus via meniscectomy does not significantly alter the AEs.

After the meniscus cut was completed, the cadaver was reopened and a larger portion (with clean margins) of the torn meniscus was removed resembling a meniscectomy (Figure 12 C/G to Figure 12 D/H). Qualitatively, the acoustic signal appeared to diminish at this stage from the meniscus tear state (Figure 12 L). When analyzed, there was a marginal increase in the b-value (1.34±0.29) toward the baseline/sham values. However, this increase was statistically insignificant when compared to the meniscus tear group ($p=0.91$). This indicates that the size of the meniscal defect or border tear patterns may not significantly alter the AEs of the knee.

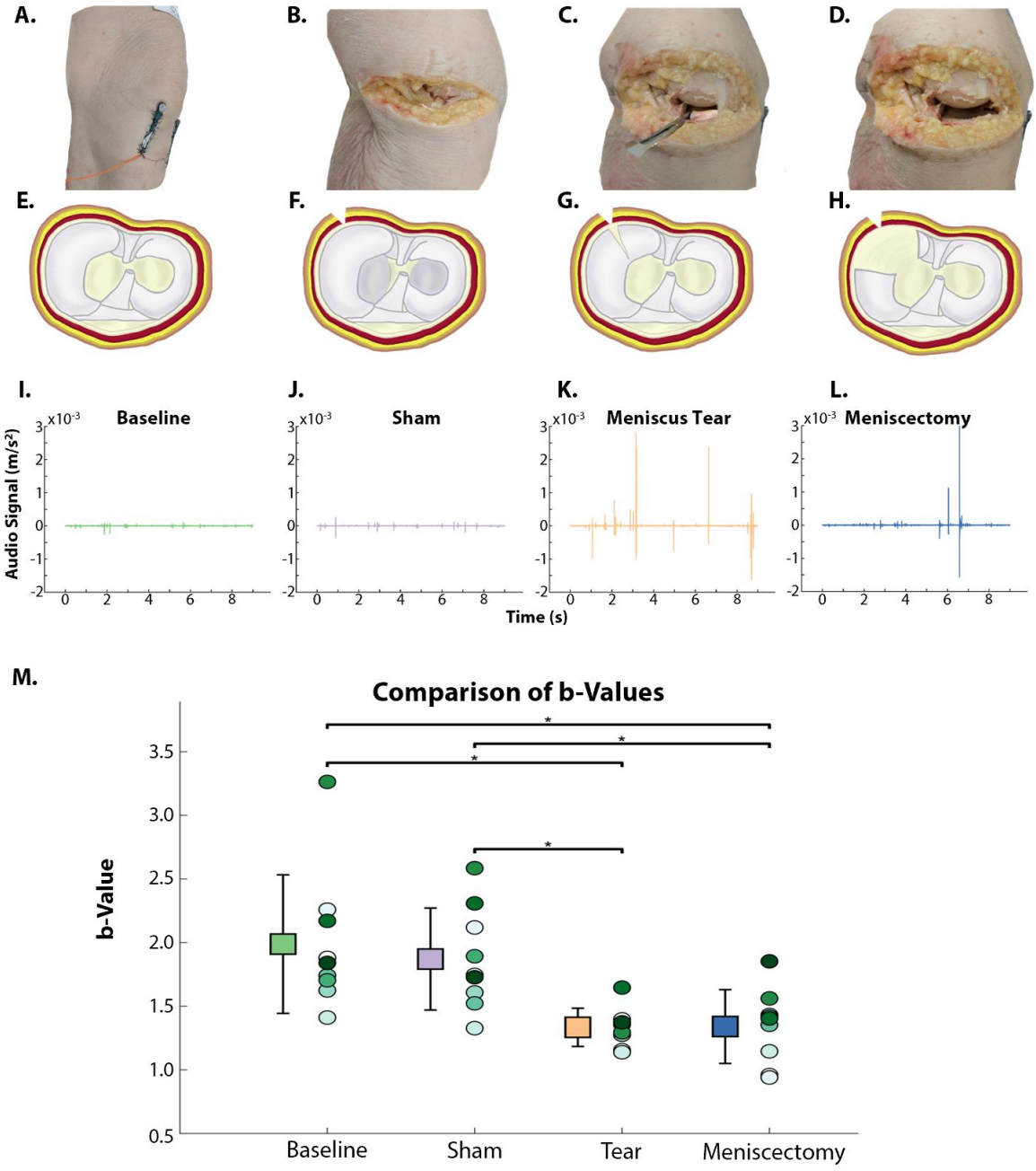

Figure 12. Acoustic data and b-values from four stages of meniscus intervention: baseline, sham, meniscus tear, and meniscectomy. Surgical stages are presented as photos (A-D) and transverse plane view of tibial plateau diagrams (E-H). Each leg's AEs were recorded at baseline (A,E), after a sham surgery (B,F), after a posteromedial radial cut (C,G), and post-meniscectomy (D,H). Representative time-domain sound data from one flexion/extension cycle at each stage are presented in I-L. Note the increase in spikes and amplitude from baseline to meniscus tear (I-J) and slight decrease from tear to

meniscectomy (K-L). Statistically significant changes in the b-value are indicated with *, n=9 and $p<0.05$). (error bars= 1 standard deviation from mean)

5.3.2.4 Saline injected within the knee capsule as a surrogate for effusions does not significantly impact AEs.

After a meniscus tear occurs *in vivo*, a series of physiologic events begin in response to the injury. Principal among these regarding the effect on mechanical articulation is localized swelling. To better understand the extent to which this swelling affects joint AEs we serially injected 5 mL aliquots of 0.9% saline solution into the knee capsule. (Figure 13A). After each injection, the AEs were recorded and b-value calculated. The b-values ranged from a minimum of 1.6±0.3 to 2.1±.6. The data were highly variable with no clear trends or statistical significance ($p>.05$ for n=5) (Figure 13 B). Therefore, the injection of saline into the knee capsule does not directly influence the production or propagation of AEs.

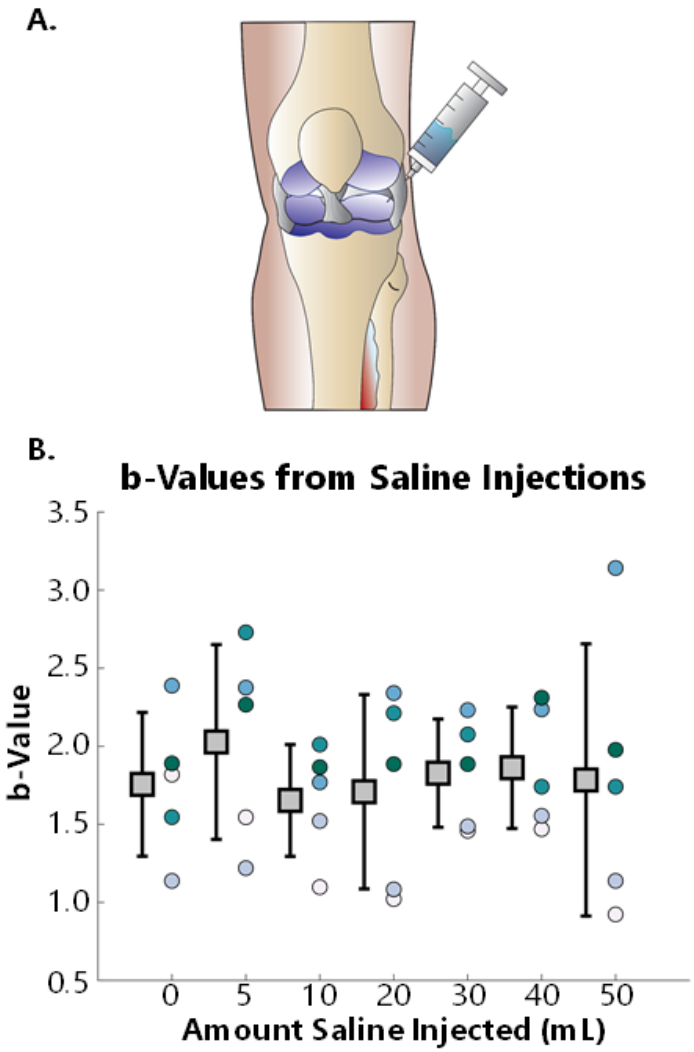

Figure 13. Acoustic Data and b-values from serial saline injections. Saline was serially injected from 0 to 50 mL into the joint cavity. (A) Demonstration of the superolateral approach used for injection of the saline. The corresponding b-values at each amount of injection are presented in (B). There were no significant differences from 0-50 mL of injected saline indicating that there was not a statistically significant change in the AEs of the knee from this intervention. (n=5, error bars= 1 standard deviation from mean of the b-value from the 5 legs tested.)

5.3.2.5 The distance between the femoral condyles and tibial plateau is minimized when the knee is between 140° and 150°:

The tibio-femoral distance was measured on one of the cadaver legs using the XROMM imaging analysis technique on one of the cadaver legs [86], [87]. Distances from the tibial plateau to the nearest point on the femur were computed for 7076 vertices that made up the tibial plateau of the CT-derived 3D model. We found that the distance between the two articulating structures was minimized between 140° and 150° during both flexion and extension. The point distances at three demonstrative angles during extension (120°, 150°, and 180°) are presented as a heat map (Figure 14 A). Of note, the minimum distances (darker red) trend to the anterior as the leg articulates. This agrees with reported tibiofemoral distances in the literature [88]–[91].For reference, the posteromedial meniscus tear was located on the bottom-left portion of the tibial plateau as presented in Figure 14A. Of note, the minimum dimensions did not significantly change between any stages of the experimental protocol (Figure 14 D,E). This indicates that the interventions did not cause significant changes in the biomechanics and articulation pattern of the cadaver leg.

The RMS power of the acoustic waveform and the rate of change of the minimum tibiofemoral distances differed during the extension and flexion phases of movement. The minimum tibiofemoral distance during extension is 0.251±0.082 cm and during flexion is 0.265±.003 cm both occurring at 145° (Figure 14 B,C). The tibiofemoral distance sharply increased from the minimum at 145° to 180° (full extension) during both the flexion and extension phases of movement. During extension, there is a large increase in the RMS power from when 145° to full extension. This increase in RMS power mirrors the increase seen in the extension-phase tibiofemoral distance plot (Figure 14 B,D). The relationship

between RMS power of the acoustic signal and distance is slightly different during the flexion phase of movement. During flexion, the peak in RMS power of 0.34±0.02 occurs at 150° flexion nearly coinciding with the minimum tibiofemoral distance at 145° flexion. During flexion we again note an increase in the rate of growth of the RMS power from 160° to 150° - closely resembling the rate of change in the tibiofemoral distance plot. The difference in the relationship between RMS power and tibiofemoral distance during the flexion and extension phases along with the slight delay in the RMS power of the signal following maximum compression indicates that there is a more complex interaction occurring to create these joint AEs than rigid-body bone-on-bone compression alone.

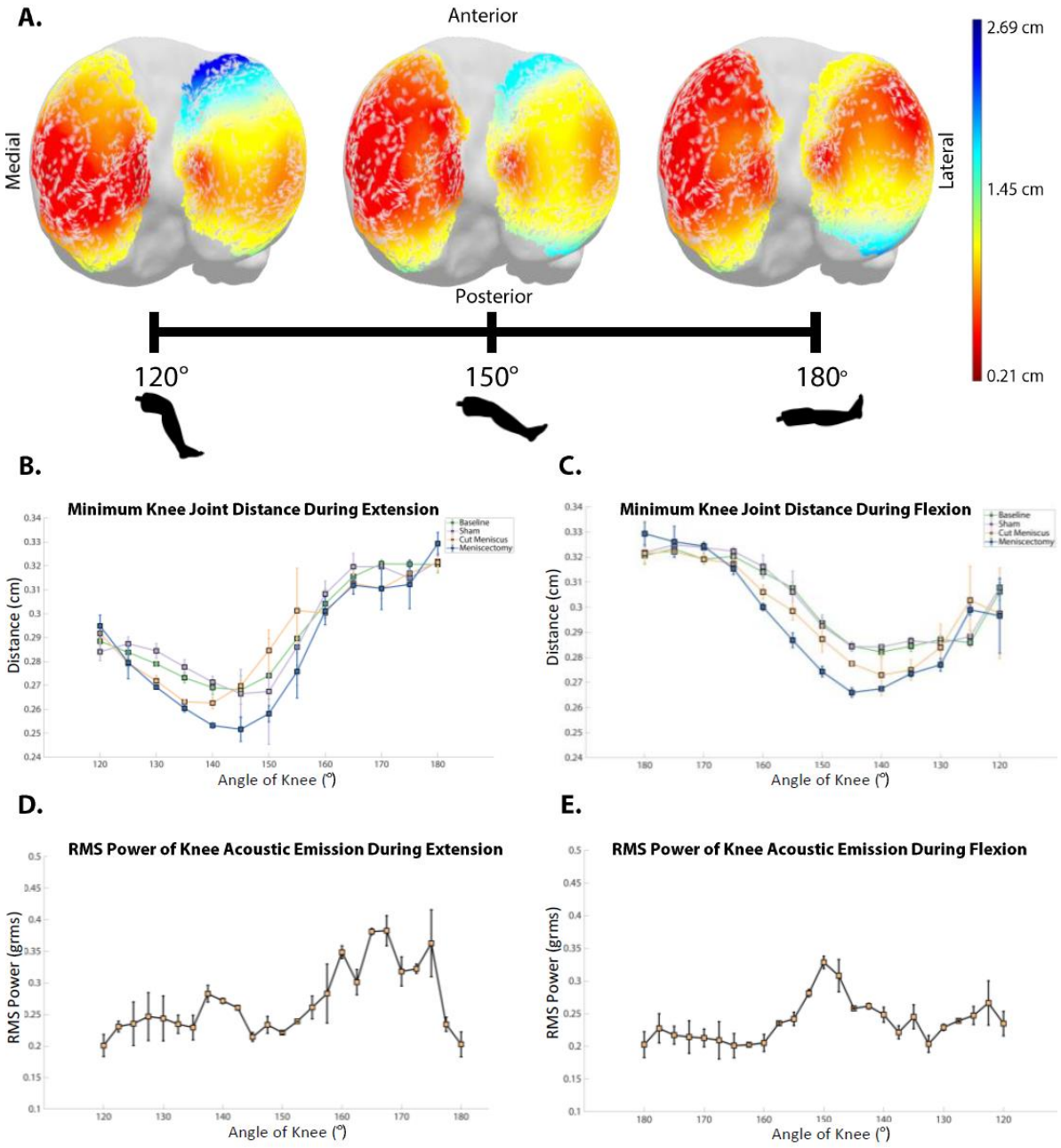

Figure 14. Comparison of tibiofemoral distances to sound recordings during a flexion-extension cycle. (A) Heatmap of distances from femoral condyles to tibial plateau at select distances (i.e. 120°, 150°, 180°). These heatmaps appeared nearly identical during flexion and extension. Minimum tibiofemoral distances at each degree of movement during (B) extension and (C) flexion (Error bars indicate one standard deviation from the mean of three trials at each data point). In B and C, the 1000 nearest vertices of the 7076 total vertices creating the 3D mesh are averaged with their standard deviations displayed. RMS Power of the joint AEs at each degree of movement during (D) extension and (E) flexion (Error bars indicate one standard deviation from the mean of the AEs of all n=9 cadaver legs tested).

5.3.3 Discussion

Meniscal tears are the most common knee injury, and partial meniscectomies are one of the most common orthopedic surgical procedures. These injuries are seen in all age groups and have a variety of causes as described in 3.2.2 [92]. The primary functions of the menisci are load distribution and stability during ambulation. Upon injury, the menisci have an impaired ability to distribute loads and resist tibial translation – destabilizing the joint. During extension, the medial and lateral menisci transmit 50% and 70%, respectively, of their compartmental loads. During flexion those increase to 85% and 90% respectively [30]. It has previously been shown that after medial meniscectomy, contact stresses can increase by 100% [18]. The type of tear has been shown to have a significant effect on contact pressures within the knee. In particular, complete radial tears significantly increase mean contact pressure and decrease contact area compared with the intact state [92]. Significant research has been performed on tear morphologies [93], [94], compartment pressures [84], [95], and outcomes of different corrective surgical approaches [92].

Diagnosis of meniscal injury is described in 3.2.2.2. This diagnostic workup is extensive, but not without its shortcomings. The patient's history and their pain ratings are highly subjective. The physical exam is dependent on the practitioner's expertise. The MRI is by far the most powerful and objective tool for evaluating the meniscus; however, it is time-consuming, costly, and often uncomfortable for the patient. Post-surgical monitoring and rehabilitation efforts rely on the same assessment techniques, typically with even fewer imaging studies. We believe that the AEs produced by the knee during flexion/extension could serve as a suitable marker of knee joint health that is quantitative, affordable, and easily incorporated into a clinical assessment.

The AEs explored in this section were produced by the articulation of the tibiofemoral joint. This joint is a hinge synovial joint as described in 3.1.1. The friction during this articulation creates a complex series of vibrations [96]. These vibrations travel to the skin where they encounter a large impedance mismatch between the tissue and air and manifest as vibration signals on the skin [24]. In this work, we proposed a cadaver injury model to better understand the impact that factors such as injury and swelling have on these acoustic vibrations. In addition, the tibiofemoral distances during movement were measured and correlated to the AEs. This was the first time that human knee AEs have been studied in such a controlled setting.

Our exploration of knee AEs began with a four-stage surgical intervention. We had previously shown that injuries to the knee resulted in significant alterations to the AEs as measurable by the b-value metric [68]. On 9 fresh-frozen cadaver legs, the sounds were first recorded at baseline after the legs were pre-conditioned and thawed in a water bath. A sham surgery was performed to expose but not damage the medial meniscus. The lack of significant changes between the acoustic signals from baseline to sham indicated that the exposure of the meniscus with the cutting of the various skin and fascial layers was not responsible for the bulk of the change in AEs. Next, we re-entered and performed a 10mm radial incision on the meniscus. With the meniscus torn, the AEs significantly increased, and this meniscus-torn state was classified using the b-value metric. In the final stage, we removed a 5 mm margin around the meniscus tear. This removal resembles a surgical meniscectomy – a commonly performed reparative surgery for this type of injury. After meniscectomy, the b-value of the AEs returned toward baseline but was still not significantly different from the meniscus cut stage. This lack of significant change in the

b-value following meniscectomy indicates two possible outcomes: 1) The cadaver model was not a suitable substitute for a reparative treatment given the lack of blood flow/synovial fluid, or 2) this sensing modality may not be suitable for monitoring post-surgical repairs. In earlier work, the knee AEs were recorded from athletes at the start of their season and after suffering injuries such as torn anterior cruciate ligaments, torn menisci, and sprained medial collateral ligaments. In that study we found that the b-value and this sensing modality was able to track their recovery post-surgical intervention [68], [81]. Thus, the lack of return toward baseline is most likely due to differences in the physiology/anatomy of the cadaver model and young, collegiate athletes.

To examine this discrepancy in findings between athletes and cadaver model, we explored a possible confounding factor – swelling. Intra-articular knee joint effusions accompany nearly all knee injuries [97]. The serial injections of 5 mL aliquots of 0.9% saline solution did not significantly alter the knee AEs. This was counter to our expectations. We had expected an increase in intra-joint fluid volume to lead to an increase in the tibiofemoral distance and less interactions between the articulating components in the knee thereby decreasing AEs. This lack of significance was promising for the sensing technology to be used clinically since the level of swelling will not need to be controlled for when interpreting joint AEs; however, it did present data counter to our prevailing notion of how these sounds were produced. This finding led to our interest in exploring the relationship between tibiofemoral distance and AEs.

To explore the relationship between anatomical distances and joint AEs a series of biplanar, video x-rays and CT scans were performed on a cadaver specimen. Following the XROMM protocol, tibial and femoral movements were tracked and segmented 3D models were

animated using the x-ray videos as reference. The distance between the tibia and femur were calculated continuously through the leg's range of movement [86], [87]. The RMS power of the joint AEs was calculated along each legs range of movement for comparison. The RMS power metric of AEs was calculated rather than the b-value since the b-value relies on a longer signal with many peaks occurring. In Figure 5, the RMS power of the signal was calculated for every 2.5° of movement which amounted to a signal time of 55.4±0.2 ms. Often, there were no large amplitude spikes in that small-time interval and the b-value would tend toward infinity by virtue of its derivation. There was a slight delay in the increase in the RMS power after the minimum joint distance was reached. We believe that this delay in sound production may indicate that the sounds are a result of the viscoelastic expansion of the menisci. The mechanical properties of the meniscus have been extensively characterized [97]. Principle among these analyses are the characterization of the complex viscoelasticity and anisotropy of the meniscus. We propose that knee AEs are heavily influenced by the compression of the menisci and the consequent release of compression during movement. The viscoelasticity of the meniscus may be responsible for the slight delay between the AE RMS spiking and the minimum tibiofemoral distance occurring.

In the future, the possibility of the viscoelastic properties of the meniscus contributing to the AEs of the knee should be further explored. If this theory is correct, its result may be far reaching regarding diagnosing meniscal health from joint sounds. There may also be merit in correlating the AEs not only with experimentally measured tibiofemoral distances but also simultaneous, joint pressure mappings. In this work, we relied on previously published research to classify the pressure profile within the knee.

Using experimental pressure data with our setup may provide a better correlation between the anatomical orientations and sounds produced. For this novel sensing modality to become clinically valuable, we need to discover the full extent of its capabilities both for longitudinal monitoring as well as diagnostics. The next steps in that development will involve expanding the scope of joint AE studies to different injuries and pathophysiological conditions (e.g., anterior cruciate ligament tears, arthritis, etc.). Future work should also focus on developing more sophisticated signal processing techniques for reducing noise, optimizing signal quality, and potentially isolating the sources of these sounds within the knee, so that the nature of these sounds can be better understood. Aside from clinical merit, there is also substantial intellectual merit to be gained by further researching the influence of confounding variables such as the ideal protocol for generating AEs, physiological variables such as pain, effusions, and degree of tear, and the ability to localize injuries based on an array of microphones. In the future, we intend to explore more fundamental AE analysis techniques to accurately and effectively characterize the differences between these responses.

This work presents the first time that knee AEs have been characterized in a controlled setting with a cadaver model of knee injury. The insights gained on the application of AEs for identifying meniscus tears are promising and warrant future work in the field. Additionally, the correlation of tibiofemoral distance to AE patterns provided the first of its kind attempt to correlate AEs with anatomical positions. The relation between joint anatomy, the associated interactions upon articulation, and the resulting AEs should be further explored to help understand the full utility of this novel sensing modality. With

more research, joint AEs could soon serve as a readily measurable, non-invasive biomarker of joint health.

5.4 Lateral Meniscus Tear, Grading Severity, and Surgical Approach

As a follow-up to our earlier work in analyzing the AEs recorded in the cadaver model, we next recorded the AEs of a lateral meniscus tear in a new cohort of cadaver limbs. In this study, the b-value metric was not able to statistically separate the different experimental groups. Ultimately, a bagged tree algorithm was applied, as described in 4.2.4. This approach enables a more thorough and methodical analysis of these sounds to determine if characteristic changes in AEs could be used to screen for and grade lateral meniscus tears.

5.4.1 Materials and Methods

Ten fresh-frozen cadaver knee specimens with no history of arthritis or significant injury were obtained for this study (5 male, 5 female, average age: 64 ± 3.5 years, average body mass index: 22.3 ±2.5 kg/m^2). Two accelerometers were sutured 1 cm medial and lateral to the distal patellar tendon. AEs were recorded using the same custom developed hardware and software used in our previous cadaver work [98]. This setup captures the skin-surface vibrations up to 10 kHz simultaneously from both accelerometers simultaneously. Joint position was synchronously recorded using an IMU attached on the distal shank of the leg. With the sensors in place, each leg underwent serial, arthroscopic surgeries in three stages: a sham surgery (scopes were placed to visualize the untouched menisci), a partial tear (half of the posterior lateral meniscus was radially torn), and a full tear (the previous tear was worsened to a full thickness tear) (Figure 15 A-C). After each

stage of surgical intervention, the AEs were recorded while the knee was moved through its full range of motion 10 times at a rate of 1 cycle every 4 seconds in triplicate. The IMU data were used to window the data into individual cycles. To quantify the recorded AEs, first, the b-value was calculated for each stage as described in 4.2.2. The b-value results were inadequate, so a more thorough analysis was performed using a bagged tree classification algorithm. First 49 audio features were calculated for each cycle of movement for both microphones during each of the three AE recording trials for all ten legs (features are described in Table 1). This amounted to 300 cycles of movement captured during each of the 4 stages of intervention for each microphone.

These features were stored in a matrix, $\boldsymbol{F_{stage}}$, with stages being baseline, sham, partial tear, and full tear. In order to remove outliers from this feature set, the interquartile ranges of each feature for each surgical stage was calculated. An interquartile range (IQR) is a measure of statistical dispersion, and is particularly useful when data is not normally distributed [99]. Any cycle that had more than half of its features outside of 1.5 interquartile ranges was labelled an outlier and removed. This resulted in 11% of the cycles being removed. With outliers removed, since we sought to compare four surgical stages, $\boldsymbol{F}$ was fit to a bagged tree classification algorithm. This type of ensemble algorithm is described in 4.2.4. This model used LOSO-CV, with 30 learners, and a learning rate of 0.1.

The algorithm classified each cycle as baseline (1), sham (2), partial tear (3), or complete tear (4). The average of all the cycle labels for a given leg was calculated. This average classification score per leg is proposed as a joint health score and presented in Figure 17. The joint health scores were not normally distributed, so a Wilcoxon rank sum

test with a Bonferroni correction for multiple comparison was used to test for statistical significance.

Finally, the relative importance of each of the input features was calculated as described in 4.2.4. Importance was determined by calculating the information gained from each stage of the tree splitting as quantified by the decrease in entropy of the dataset after the split. These features are presented in Figure 18.

5.4.2 Results

5.4.2.1 Qualitative Differences in the AEs at Different Stages

The recorded AEs had noticeable, qualitative differences in their time domain signals at each stage of surgical intervention (Figure 15 D-G). As the tear was made more severe there were more frequent and larger "clicks" observed in the signal.

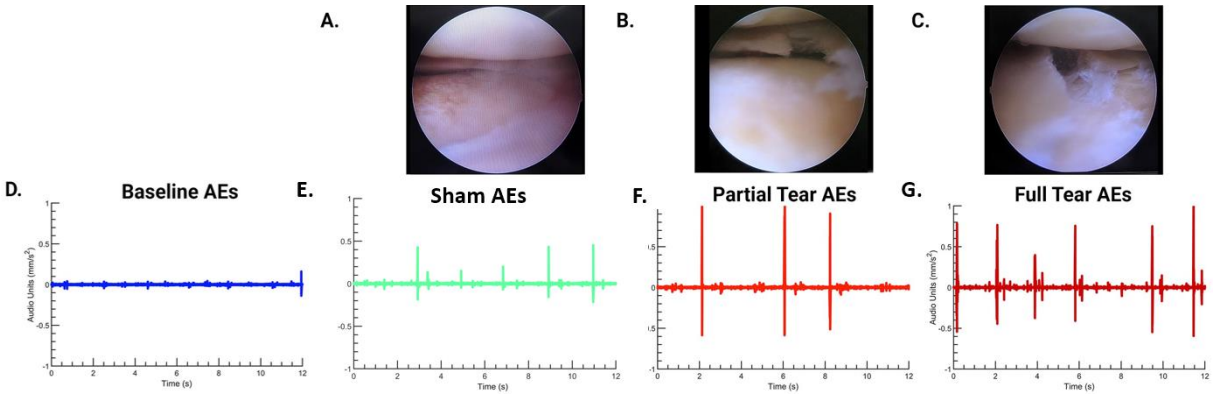

Figure 15. Stages of Lateral Meniscus Cadaver Injury with Associated AEs Profile. (A-C) Arthroscopic images of the lateral meniscus. (D-G) Time Domain waveforms of the AEs at each stage.

5.4.2.2 b-Value based AE analysis was unable to stratify the surgical stages

The b-value was the primary quantitative metric we had previously applied to our work in the cadaver model of AEs. However, as seen in Figure 16, the b-value was not able to statistically separate the four stages of interventions. A different surgical approach was used in this work (arthroscopy), but similar to our earlier work the b-value did not find any differences in the AEs from baseline to sham. However, since the b-value found no statistical differences in any of the stages a more thorough analysis was warranted.

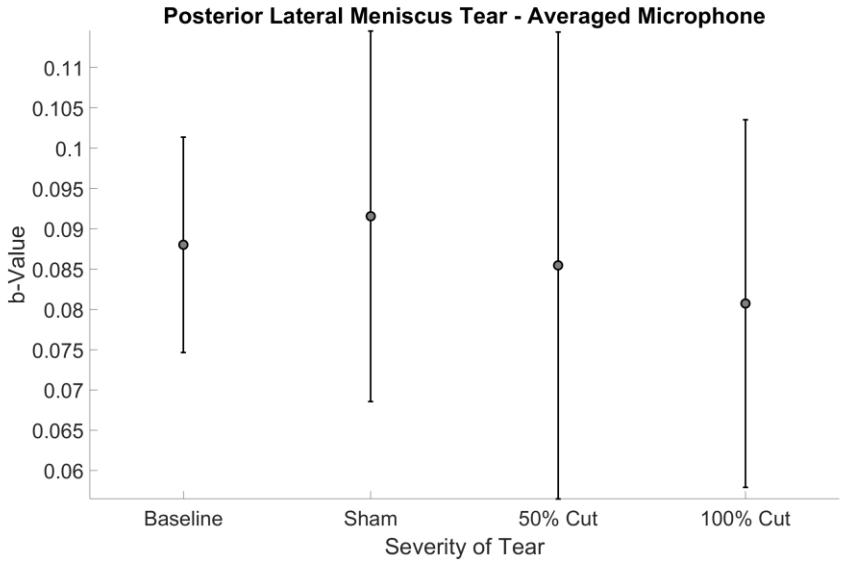

Figure 16. b-Value Classification of Lateral Meniscus Tears

5.4.2.3 Surgical Approach Impact – Baseline to Sham

The surgical approach used in this work differed from our earlier work. In this project, an orthopaedic surgeon performed the interventions to induce injury arthroscopically. The approach resembled what is normally performed to repair a meniscal tear, as described in 3.2.2.3. This approach involves filling the knee with saline fluid to

open the joint space. The joint health score was calculated for each surgical stage using the bagged tree algorithm. In Figure 17, there is a small increase in the joint health score from 1.89±0.79 to 2.20±0.55, but this increase was not statistically significant.

5.4.2.4 Introduce a meniscal tear alters the AE profile

The joint health score increased from 2.20±0.55 to 2.72±0.50 between the sham stage to the partial tear stage. This 23.6% increase in the joint health score was statistically significant ($p<0.0001$). After re-entering the knee and increasing the severity of the tear to a full thickness tear, the average joint health score continued to increase to 3.08±0.64 – a 13.2% increase. Again, this was a significant change from the sham stage ($p<0.0001$).

5.4.2.5 AEs can be used to grade the severity of a lateral meniscus tear

The potential for AEs to grade a meniscal tear has never before been studied. In our model, the knee goes from a relatively minor (~50%) posterior tear to the lateral meniscus to a complete thickness tear. With the increased severity of injury, there was a similar increase in the joint health score from 2.72±0.50 to 3.08±0.64. This increase was also statistically significant, but with a higher p-value than in our other comparisons ($p<0.005$).

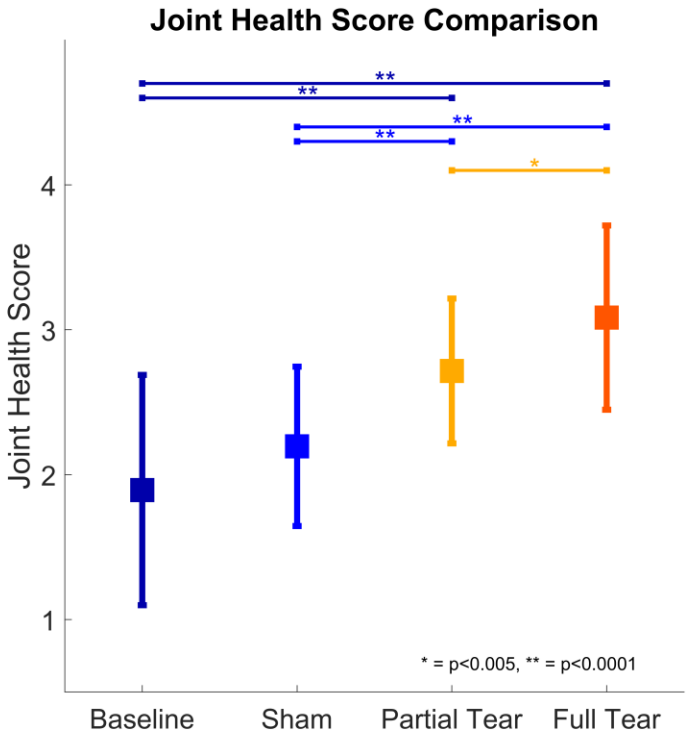

Figure 17. Average Joint Health Score of Each Lateral Meniscus Injury Stage. (n=10)

5.4.2.6 Feature Importance Ranking

49 features were used as predictors in the bagged tree model. These features are described in Table 1. Figure 18 presents the top 13 features used in classifying between the four stages of surgical intervention. The importance of each feature was divided by the leading feature – the median of the spectral centroid. With this normalization, the relative importance of the features ranged from the lowest-ranked feature (median energy) at 0.18 to the median spectral centroid at 1.0.

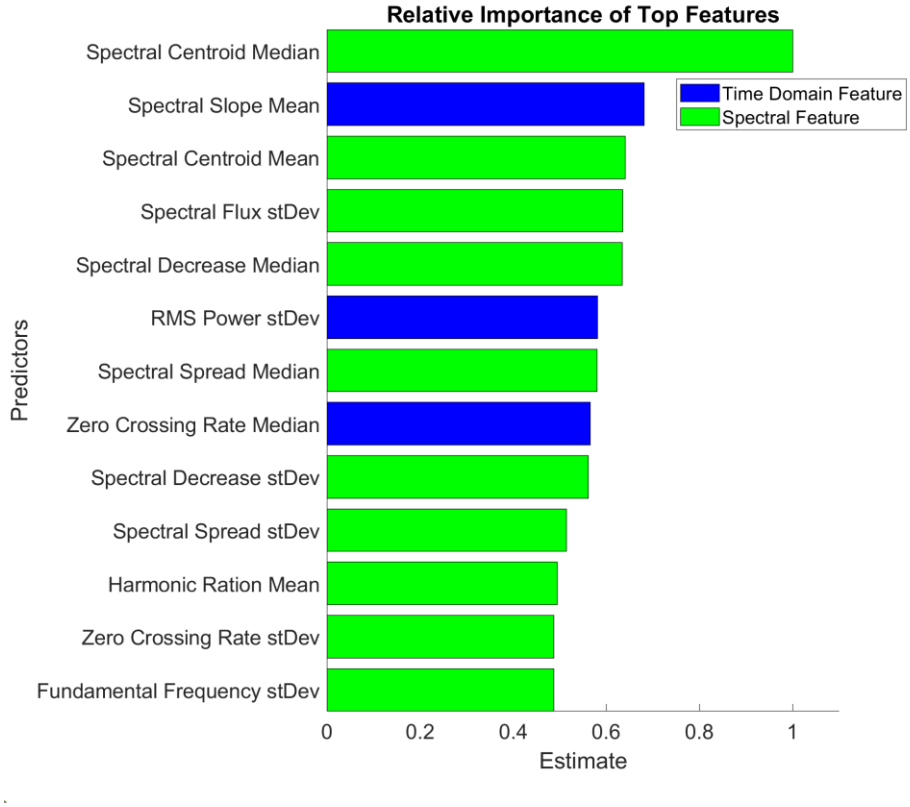

Figure 18. Relative Feature Importance of Top 12 Features for Classifying Four Stages of Lateral Meniscus Injury.

5.4.3 Discussion

The bagged tree classification algorithm successfully labeled the knee status at a healthy baseline, and after partial and full thickness lateral meniscus tears using only the AEs from the controlled cadaver injury model. This study shows for the first time that AE analysis can diagnose and grade lateral meniscus tears in a human cadaver model. The knee is one of the most frequently injured body parts, and this technique could one day serve a screening tool for triaging possible knee injuries prior to imaging studies. This non-invasive sensing modality holds promise for future clinical applications, but the study was

not without limitations. Principal among these is that the cadaver knees acquired were of an advanced age. They did not have any reported arthritis, but age-related degeneration may still alter the AE profile. Further research is needed to determine the effects of physiologic degeneration and aging on AE production.

In the studies presented in this chapter, the surgical approach to reveal the meniscus did not significantly alter the AE profile. The significant changes in the AEs occurred following a tear to the meniscus. These controlled cadaver studies support our central hypothesis that alterations to the internal environment of the knee are reflected in the surface AE recordings, and that AEs contain clinically relevant information.

CHAPTER 6. DIAGNOSIS AND MONITORING OF JIA

6.1 Overview

In this work, we quantify the AEs from the knees of children with JIA and support their use as a novel biomarker of the disease. JIA is the most common rheumatic disease of childhood; it has a highly variable presentation and few reliable biomarkers, which makes diagnosis and personalization of care difficult. The knee is the most commonly affected joint with hallmark synovitis and inflammation that can extend to damage the underlying cartilage and bone [100], [101]. We hypothesize that AEs from the knee contain clinically relevant information about the joint, and that this information could be used to aid in the diagnosis, personalization of treatment and longitudinal monitoring of JIA. In this study, we compare the AEs from 25 patients with JIA -- 10 of whom were recorded a second time 2-3 months later -- alongside 18 healthy age- and sex-matched controls. We compute 49 features from each flexion/extension cycle of each subject. Those features are used to train a logistic regression model, which can classify individual cycles of flexion/extension as having JIA or being healthy with 84.4% accuracy using leave-one-subject-out cross-validation (LOSO-CV). When analyzing the complete AE recording of a subject, which contained at least 8 cycles of flexion/extension, a majority vote of the cycle labels accurately classified the subjects as having JIA or being healthy 100% of the time. To better understand the longitudinal monitoring capabilities of AEs, we use the output probability of a JIA classification as a basis for a joint health score. We compute this score for the patients with JIA at their first and follow-up visits. In all 10 of our follow-up

recordings, the trend in joint health score accurately tracked with successful treatment of the condition. Our proposed AE-based classification model of JIA presents a compelling case for incorporating this novel joint health assessment technique into the clinical work-up and monitoring of JIA.

6.2 Introduction

JIA describes a heterogenous group of arthritides that present in children. It is a leading cause of disability and the most common chronic rheumatic disease of childhood with a prevalence of 150 cases per 100,000 [35]. It is an autoimmune disorder with a complex etiology thought to be related to a combination of pre-disposing genetic factors and environmental influence [34], [36]. JIA is discussed in depth in 3.2.1.

6.2.1 Diagnosis and Subtyping

The heterogeneity of JIA's presentation makes diagnosing JIA difficult. This difficulty is exacerbated by the lack of conclusive, diagnostic laboratory tests. Diagnosis currently relies on taking a thorough history, physical exam, and several laboratory and imaging studies [38]. Once diagnosed, in order to select the most suitable treatment for JIA the disease should be classified into its subtype. JIA is divided into seven subtypes based on laboratory and clinically observed features [39], [40]. Determining the most appropriate subtype, and thus the most effective therapy, requires extensive workup that is both time and resource heavy. For a child with swollen and painful joints, receiving a proper diagnosis can be exceedingly difficult, but it is only the beginning of management.

6.2.2 Treatment

After diagnosis, the goal of treatment is to achieve clinical remission [38]. Current treatment protocols are discussed in 3.2.1.3. In summary, treating JIA is difficult due to the variability of the condition, large number of treatment options, and inability to predict patient response. The treatment protocol is largely reactive with decisions made based on subjective and qualitative measures of response to therapy.

6.2.3 AE Opportunity Space

Early diagnosis with effective treatment is necessary for preventing the long-term sequela of JIA [38]. Pediatric rheumatologists are the most well-suited physicians for diagnosing and treating JIA; however, there is currently a severe shortage of pediatric rheumatologists. As of 2019, there are fewer than 400 board-certified and practicing pediatric rheumatologists in the United States. This shortage contributes to only 1 in 4 children with JIA being able to regularly see a pediatric rheumatologist [102], [103]. To address the difficulty of diagnosis, subjectivity of treatment, and severe lack of access to pediatric rheumatologists, more research must be performed in to develop objective biomarkers of JIA. A suitable biomarker could help more effectively diagnose patients, identify risk profiles, and predict/track an individual's response to treatment. Additionally, the development of such a biomarker could allow for more effective translation of the many genetic and immunological mechanistic studies of the disease to further improve clinical outcomes. Ideally, this biomarker would also be readily measurable with affordable technologies, so that JIA could be easily diagnosed and monitored by non-specialist healthcare workers.

The use of AE analysis could provide a basis for developing such a biomarker [104]. In the case of a chronic condition – such as JIA – AEs could serve as a means of not only diagnosing but also longitudinally monitoring the conditioning. If AEs show a correlation with disease status in JIA, they could regularly be measured to help personalize the management of JIA. Until recently, longitudinal assessment using AEs in healthcare was not feasible due to a lack of technologies for recording AEs outside of a laboratory or clinical setting. However, the development and application of piezoelectric accelerometers to AE assessment has substantially advanced the field. This type of sensor is sensitive to physical vibrations (such as those seen on the skin during joint articulation), but does not substantially record external noises [105]. AE joint assessment technologies if properly applied to JIA, could lead to earlier diagnosis, improved, personalized care, and could serve as an objective measure in the next generation of clinical trials.

6.2.4 The Goal

In this chapter, we explore the potential of using AE analysis to diagnose and longitudinally track JIA. AEs are recorded from the knees - one of the most commonly affected joints in JIA [64], [106]. In CHAPTER 5, we found that by damaging the meniscus in a cadaver model of the knee, the resulting AEs were substantially altered (published in [98]). In JIA, affected joints are characterized by persistent joint swelling caused by an accumulation of synovial fluid and thickening of the synovial lining [34] (Figure 19A). We hypothesize that these pathologic changes in the knee will similarly alter the AE profile of the knee. If that hypothesis is supported, the AEs of the knee could then be correlated with disease status.

6.3 Knee AEs as a Biomarker of JIA Status

6.3.1 Experimental Design Overview

In order to test this hypothesis, we use our custom hardware and software setup to record AEs during 10 cycles of flexion/extension (Figure 19B-D). We place two piezoelectric accelerometers medial and lateral to the distal patellar tendon, and an IMU around the ankle. The AEs from the knees of two groups of children are recorded: one group had active JIA and the other was an age- and sex-matched healthy control group. To assess the effectiveness of AEs for tracking therapeutic efficacy and changes in disease status, we also recorded the AEs from the children with JIA 2-3 months after successful treatment. Our proposed algorithm, powered by logistic regression, analyses 49 signal features of each individual cycle of flexion/extension and outputs the probability that a cycle of movement came from a patient with JIA. This output probability forms the basis for our proposed JIA digital biomarker. Finally, we assess the importance of each signal feature in the algorithm as well as the accuracy and generalizability of the model using leave-one-subject-out cross-validation (LOSO-CV). An overview of this experimental design and analysis is presented in Figure 19.

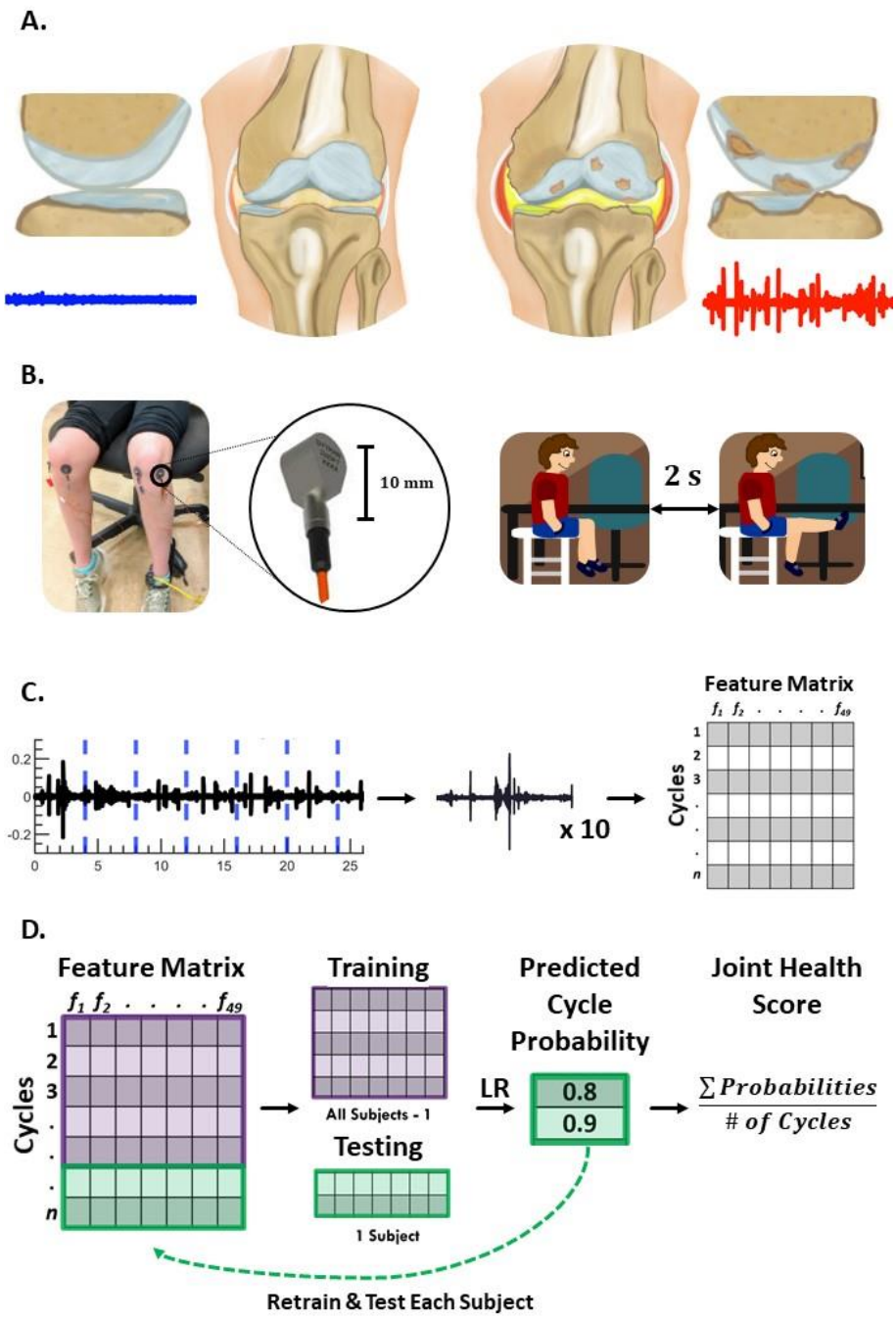

Figure 19. Joint AE Overview. (A) A healthy knee articulates smoothly due to its smooth cartilage and appropriate amount/constituency of synovial fluid. This smooth articulation creates a noise-like AE (blue). In JIA, you may see cartilage loss, bone erosions, and a thickened/inflamed synovium with excessive joint effusions. These changes are hypothesized to create an AE with several large spikes (red). (B) To record the knee AEs, two contact accelerometers were placed on each child's knees. They viewed and replicated the movements in an instructional cartoon during AE recording

such that their movement speed and range of motion was controlled. (C) The resulting AEs were split into their approximately ten component cycles. 49 features were calculated to describe these cycles. (D) Using logistic regression and LOSO-CV, the probability of each cycle belonging to JIA were calculated. The average of those cycle probabilities is used as a "joint health score" to indicate the severity of JIA. If the majority of cycles for a given subject had a probability of JIA greater than or equal to 0.5, that subject was classified as having JIA.

6.3.2 Materials and Methods

6.3.2.1 Human Subject Protocol and Subject Demographics

The study was conducted under a protocol approved by the Georgia Institute of Technology and Emory University Institutional Review Boards. 43 subjects participated in this study. 25 of the subjects were diagnosed with JIA by a pediatric rheumatologist and 18 of the subjects were healthy controls with no history of JIA or acute knee injuries. The group with JIA consisted of 20 females and five males (12.2 ± 3.1 years old, BMI 20.1 ± 4.1 kg/m^2). The healthy control group consisted of 15 females and three males (12.9 ± 2.7 years old, BMI 22.3 ± 2.8 kg/m^2) with no history of joint disease, surgery or significant joint injury. In order to capture longitudinal changes in the knee AEs during the course of treatment, data were acquired from ten of the subjects (1 male, 9 female, 12.5 ± 3.3 years old, BMI 20.8 ± 3.5 kg/m^2) with JIA a second time, 3-6 months after initial measurements (follow-up group). Note, that JIA is more prevalent in females with estimates ranging from 65-78% of all cases occurring in females [107], [108].

The data acquisition setup for each subject is shown in Figure 19B. To record the sounds produced by the joints, two uniaxial analog accelerometers (3225F7, Dytran Instruments Inc. Chatsworth, CA) were attached to each knee using double-sided adhesive

pads (Rycote Microphone Windshields Ltd, Stroud, Gloucestershire, GL5 1RN, United Kingdom). These professional-grade pads tightly coupled the accelerometer to the subject's knee. This accelerometer placement location was the same used in CHAPTER 5 that allowed for the capture of high-fidelity signals capable of differentiating meniscus injuries in a cadaver model.

To record the knee AEs, each subject performed ten unloaded knee flexion/extension exercises, while seated on a height-adjustable stool to prevent foot contact with the ground. The subjects repeated the movement as seen on an instructional cartoon that encouraged a cycle to be completed every four seconds through their full range of motion (RoM). The signals from the accelerometer were sampled at 100 kHz and recorded using a DAQ (USB-4432, National Instruments Corporation, Austin, TX). An IMU attached around the ankle of the subject recorded synchronous positional data during AE recording at 50 Hz to allow for analysis on a cycle-by-cycle basis, as well as to ensure the subject maintained an appropriate speed and RoM. The ideal speed and angles to move through have previously been explored using a cadaver model of joint AEs [98]. The exercise and recording protocol were repeated for both knees for all subjects. The recorded signals were analyzed using Matlab (MathWorks, Natick, MA).

6.3.2.2 Knee Movement Tracking

The goal of this project was to compare the AEs of the knees between subjects with JIA and healthy controls. We hypothesized that these sounds were originating from the articulation of the joint. In order to properly compare the AEs, we needed to control the movement as feasibly as possible. However, knowing that there would be a certain amount

of inter- and intrasubject variability in the performance of the flexion/extension cycles, we also chose to quantify key characteristics describing this movement. We calculated the range of motion (RoM), angular speed, cycle duration and angular speed variability of each movement cycle for every subject in the three groups. These features were calculated from the synchronously recorded IMU data as described in 4.1. The range of motion was taken to be the difference between the maximum and minimum angles of each cycle. The cycle duration was calculated as the time spent from the minimum of flexion to the next minimum of flexion. The angular speed is the angular distance (number of degrees the sensor moved through in one cycle) divided by the cycle duration. Finally, the angular speed variability is the rate of change of the angular speed during one cycle. This parameter was used as a means of describing the fatigue of a subject's movement. If the subject got fatigued during the recording it would be expected that the speed would decrease, and the angular speed variability would reflect that change.

The movement features were found to be normal, so an unmatched, two-tailed student's t-test was applied to the JIA data against the healthy control data, and the healthy data against the follow-up data. A matched, two-tailed t-test was applied between the JIA data and the follow-up data. Significance was set to an alpha = 0.05, and with a Bonferroni correction for multiple comparisons, that meant a significant result would need a p-value less than 0.167.

6.3.2.3 Signal processing and feature extraction

The joint AEs were analyzed in the time and frequency domains. Figure 20 shows a representative plot of the time domain signal after bandpass filtering from one subject with

JIA, that subject's AEs at 6-weeks follow-up, and a healthy, matched control's AE recording. The AEs from these subjects have high bandwidth frequency content as expected from earlier work [81], [109], [110]. Figure 19C graphically depicts the signal analysis workflow for knee acoustical emissions. The signals are pre-processed using a digital finite impulse response (FIR) band-pass filter with 250Hz - 10kHz bandwidth. In order to segment the AE data into individual flexion / extension cycles, a FIR low-pass filter (5 Hz) is applied to the raw AE signals to visualize the movement of the knee through its RoM. This motion data is compared against the synchronized IMU data and the proper indices for the beginning and end of each flexion / extension cycle are selected. These individual cycles are separated and subdivided into 400ms long frames. 49 signal features are extracted from each frame for each microphone. The ten frames corresponding to one cycle are averaged to give 49 descriptors of each cycle of flexion / extension. This process was repeated for all four microphones – two on each knee. These feature sets were stored in the matrix, $\boldsymbol{X}$. The rows of $\boldsymbol{X}$ each represent a single cycle of movement as recorded from each microphone, and the columns represent each of the 49 features extracted. The matrix $\boldsymbol{X}$ was standardized to zero mean and unit variance by subtracting the mean of each column and dividing by its standard deviation (see Feature Matrix in Figure 19*C*).

The features extracted can be categorized into two groups: either time domain or spectral features. The time domain features include the zero-crossing rate (ZCR), energy, root-mean-square (RMS) power, and entropy. The frequency characteristics of the joint sounds are described by the spectral features including the spectral centroid, spectral flux, spectral density, spectral roll-off, spectral spread, and spectral entropy (A full list of the features is available in Table 1). The mean, standard deviation, and median are calculated

for each feature for the set of 400ms windows on each cycle to better classify these features. This approach using these particular features to classify joint AEs is based on the appearance and sound of the signals. Their selection was supported by previous pilot work on this topic as well our cadaver model work, and is described in detail in 4.2 [76], [104].

6.3.2.4 Knee audio score classification using logistic regression

With the data appropriately organized, we trained a logistic regression classification model. Logistic regression is a common machine learning technique borrowed from statistics for binary classification problems (e.g. healthy vs JIA). Logistic regression is discussed in 4.2.3. It takes the input (*x)* which correspond to each of the forty-nine features in one row of the feature matrix $\boldsymbol{X}$, and outputs the classification label *y* for that input. This classification is decided based on a threshold of the probability that a given feature in *x* belongs to the JIA class. If the majority of the forty-nine features have a classification probability of JIA greater than the threshold, that row (or cycle) is classified as JIA. This threshold is described in Equation 5. A subject is classified as healthy or having JIA based on which class the majority of their component cycles are predicted to belong to using that same threshold.

$$\begin{aligned} mean\big(p(x)\big) \leq\ & 0.5, y\ =\ \text{Healthy} \\ \text{mean(p(x))}\ >\ & 0.5, y\ =\ \text{Injured or Diseased} \end{aligned} \tag{5}$$

Equation 5. Threshold for Healthy Control vs JIA Classification

In order to assess the accuracy of this model, LOSO-CV is applied to the logistic regression model as described in 4.2.6. All rows for one subject can be removed from $\boldsymbol{X}$ to leave behind $\boldsymbol{X'}$ and $\mathbf{X}_{\mathbf{subject}}$. Each row in these matrices corresponded to one accelerometer's output from one cycle of flexion/extension of the subject's leg. Each subject had two accelerometers on each leg and was asked to perform ten cycles of movement. However, clipping often occurred at the beginning or ending of a recording, so the average number of rows in these submatrices was 36 ± 3 rows, with the average number of rows per accelerometer being 8±1 rows. The mode of the labels in any given $\boldsymbol{X}_{\mathbf{subject}}$ is taken to be the subject classification. For example, If the majority of the rows are predicted to be 1's, the subject is labeled as having JIA. Similarly, if the majority of rows are predicted as 0's, the subject is labeled as healthy. To quantify the accuracy of the logistic regression classification, each $\boldsymbol{X}_{\mathbf{subject}}$ was serially tested using LOSO-CV.

In each fold of LOSO-CV, the logistic regression classifier is trained using the data in $\boldsymbol{X'}$ with one subject omitted - $\boldsymbol{X}_{\mathbf{subject}}$. The trained model predicts the classification of the AE signal of the excluded subject's knee AEs. During LOSO-CV, the matrix $\boldsymbol{X'}$ was standardized after the removal of $\boldsymbol{X}_{\mathbf{subject}}$. The mean and standard deviation of $\boldsymbol{X'}$ were then subtracted and divided, respectively, from the columns in $\boldsymbol{X}_{\mathbf{subject}}$. By doing this, the calculated features for $\boldsymbol{X}_{\mathbf{subject}}$ were not prematurely included in the standardization of $\boldsymbol{X}$. The model estimates the probability of JIA for each row (cycle) in $\boldsymbol{X}_{\mathbf{subject}}$. These probabilities are stored in the vector, $\boldsymbol{p}_{\boldsymbol{predicted}}$. The overall subject's audio scores are calculated by averaging the contents of $\boldsymbol{p}_{\boldsymbol{predicted}}$ (Figure 19*D*).The 0.5 threshold is applied to this average probability to assign the predicted label of healthy (0) or JIA (1). The cross-validation is completed by calculating knee audio scores for all forty-three subjects,

excluding one subject per fold. The generalizability of the model is assessed by calculating the accuracy of our algorithm in labelling each cycle, as well as in labeling each subject and comparing those labels to the clinical diagnosis of the patient.

The average probabilities of each cycle are used not only for predicting labels, but also as an indicator of joint health status, or a proposed "knee health score". In this way, as the average probability of a subject trends toward 0, the signal more resembles a healthy knee. As those probabilities trend toward 1, the signal resembles a more actively inflamed knee having JIA. For subjects with JIA that have follow-up recordings, this process is repeated to calculate the change in the probability of JIA between the first recording and second. Importantly, the follow-up recordings are never used as part of the training set, since at the time of recording those subjects the ground-truth of their disease status is unknown.

6.3.2.5 Feature importance ranking

The relative weighting of each of the features in the model needs to be explored to understand which features most relate to differentiating JIA AEs from healthy AEs. In order to quantify the importance of each feature, the standardized data from every subject with JIA (excluding the follow-up data due to it lacking a ground truth classification) is used to train the classifier. The resulting model is used to generate relative feature importance scores. In this case, no testing set is required to quantify feature importance since we are not assessing the generalizability of the model. In the case of logistic regression, the model computes a coefficient for each input feature that describes the *k*-dimensional hyperplane that best separates the two input classes. When the input matrix $\boldsymbol{X}$

is standardized to zero mean and unit variance, the absolute value of each of the coefficients output from the model can be directly compared to assess relative importance to the model. In this way, a coefficient with a large absolute value has a larger effect on the model than one with a smaller value. All forty-nine features are ranked in order from most to least important, and the top 20 features as provided in **Figure 23*A***.

6.3.2.6 Effect of number of features and cycles of movement on model performance

After ranking the forty-nine features, we further assessed the impact on the accuracy of the model's predictive capabilities by training the model on one to forty-nine features in order of their relative importance. We first trained a model on only the most important feature, and assessed the accuracy of the model as detailed above using LOSO-CV. Next, we iteratively added each new feature in order of descending relative importance to observe how that accuracy improved with the addition of each new feature. We simultaneously assessed the importance of the number of flexion/extension cycles by testing each iteration of the model on a subset of all the cycles. For example, we first trained the model on the most important feature, and tested the model using one cycle from the subject left out, next two cycles, then three cycles, all the way up to the full number of recorded cycles. In doing so, we calculated how the model responded for each feature input and for each additional cycle of movement input. Of note, when choosing the subsets of cycles to test we iteratively tested up to 1,000 unique permutations on any given sized subset of cycles and the average of those cycles was reported. A heatmap of these results was generated and can be seen in **Figure 23*B*.**

6.3.3 Results

6.3.3.1 Qualitative Comparison of Knee AEs.

The AEs were recorded from the knees of two groups of children. One group had actively inflamed knees with either newly diagnosed or poorly controlled JIA, the other group was composed of age- and sex-matched healthy controls with no JIA or significant injuries to the knee. There are several notable differences in the time-domain patterns of the AEs between these groups. The 18 healthy controls had no noticeable peaks in their audio signals and upon listening the recorded AEs resembled white noise (Figure 20A). The 25 subjects with JIA consistently show periodic, high-energy clicks in each flexion-extension cycle. These "clicks" have a spike-like appearance in the time domain (Figure 20B). Ten of these JIA patients had a second recording after 6 weeks to 3 months of treatment as prescribed by their treating pediatric rheumatologist. The AEs of this follow-up group showed a large reduction in the amplitude and frequency of the clicks noted during their actively inflamed stage (Figure 20C). The post-treatment AEs more closely resembled the healthy controls both in the plot of the time domain of the AEs and in listening to the recordings. Representative subject AE recordings from each of these groups is presented in Figure 20.

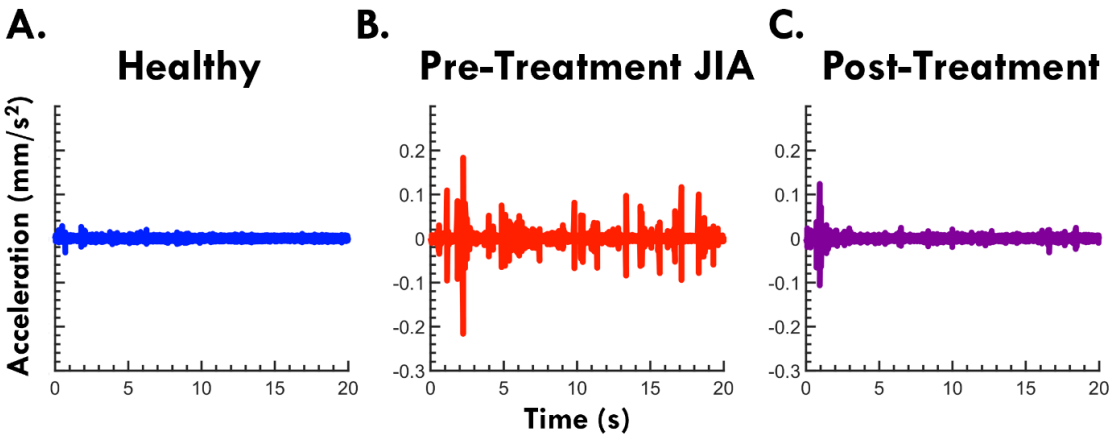

Figure 20. Representative time-domain plots of AEs from a healthy control (A), a subject with active JIA (B), and that same subject after 6-weeks of successful treatment (C). Each recording is twenty seconds and presents the AEs from five flexion/extension cycles.

6.3.3.2 Knee Movement Classification

The movement patterns of patients in the three groups had several significant differences, but these differences do not align with the change seen in the AEs. Only the cycle duration was statistically different between the JIA and healthy control groups. The RoM, angular speed and cycle duration were statistically different for the healthy and follow-up groups. The RoM and angular speed were significantly different between the JIA subjects and their matched follow-up recordings. The results of the knee movement features are shown in Figure 21.

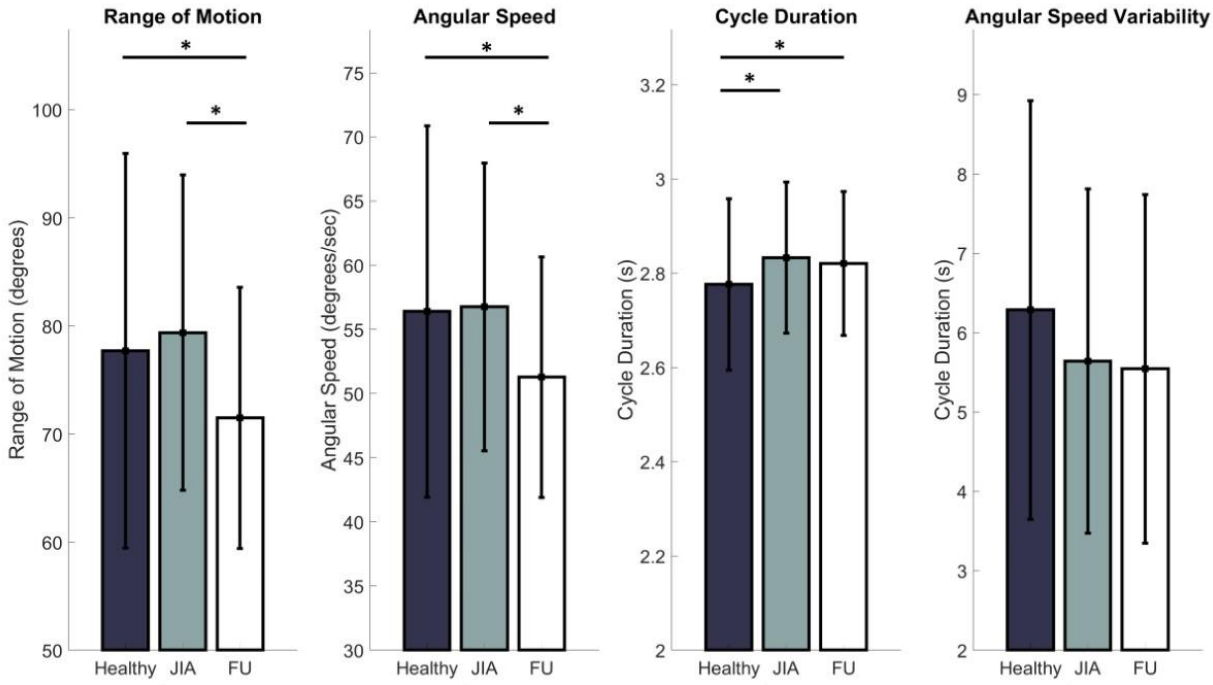

Figure 21. Knee Movement Classification. * indicates statistical significance <0.05 or 0.0167 (with correction for multiple comparisons). FU=Follow-up

6.3.3.3 Knee Audio Score Classification

The knee audio score for each subject was defined as the probability of a cycle belonging to a subject with JIA. In this manner, a knee score of 0 indicates 0 probability of having JIA, and a score of 1 indicates an actively inflamed joint with JIA. A threshold was set at a score of 0.5 to delineate the classification of the two groups. A threshold cutoff of 0.5 was chosen heuristically but could theoretically be changed to place an emphasis on sensitivity vs specificity as desired. A subject's joint health score was calculated by averaging all of the computed cycle probabilities of that individual subject's flexion/extension cycles. The subject-level joint scores are presented as a histogram in Figure 22A. Notice the heavy overlap between the healthy (blue) and post-treatment, follow-up subjects (purple). This was expected based on the success of the treatment as

reported by the treating pediatric rheumatologist. The JIA distribution is centered around a score of 0.82 with clean separation from the other two distributions. The overall cycle-based logistic regression analysis had an accuracy of 82.7% for classifying individual cycles. The ROC curve and confusion matrix are presented in Figure 22 B,C. The ROC curve had an AUC of 0.899. The cycle classification had a specificity of 80.4%, a sensitivity of 84.5%, an error rate of 20.1%, a positive predictive value (PPV) of 84.7%, and a negative predictive value (NPV) of 90.2%.

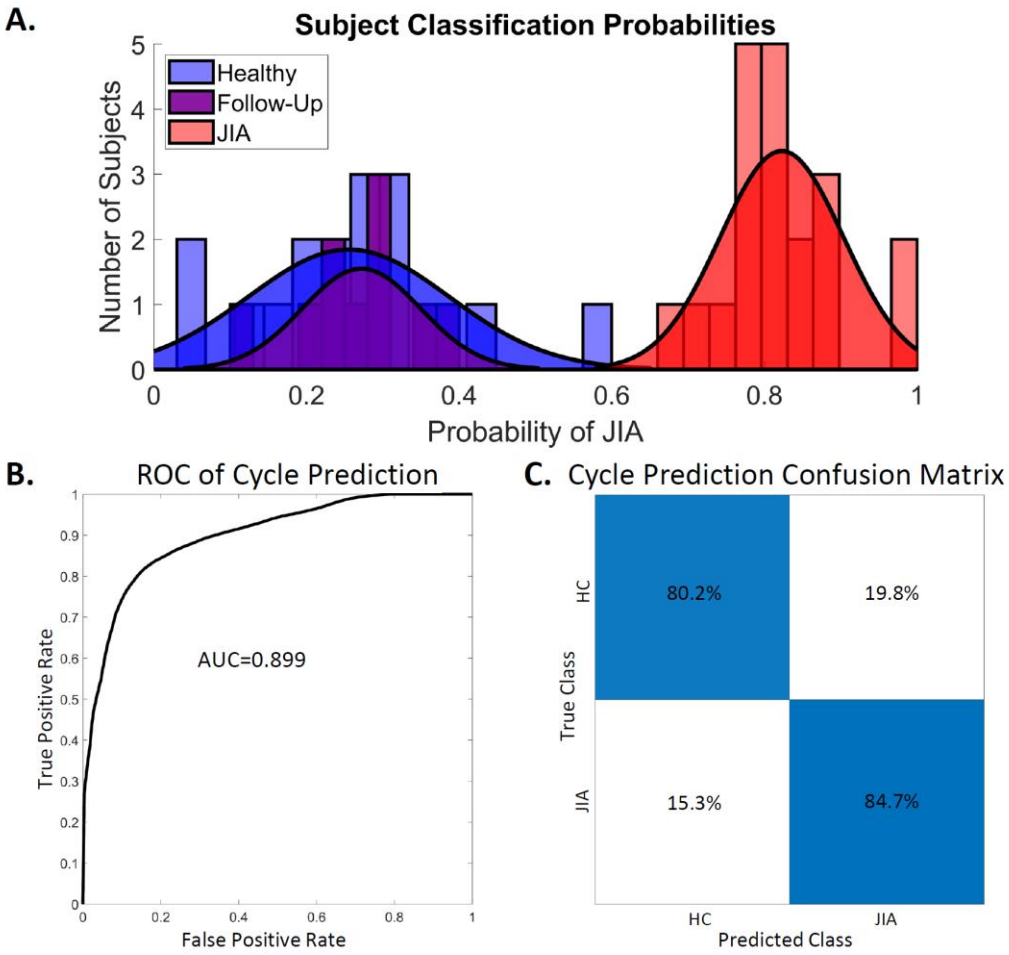

Figure 22. Assessing the performance of the logistic regression classifier on subjects (A) and cycles (B-C). (A) There was little overlap in the computed joint health score of the healthy control group and the group with JIA. A sub-group from the JIA group after effective treatment had JIA scores heavily overlapping with the healthy control group at follow-up. (B-C) The logistic regression model overall classified the individual cycles accurately 82.7% of the time. The model achieved adequately high sensitivity (84.5%) and specificity (80.4%).

6.3.3.4 Feature importance ranking and model performance.

Logistic regression is a binary classification algorithm that finds the best hyperplane in the feature space which separates the 2 classes: in this case, healthy and JIA[71]. The absolute values of the individual feature weights describing that hyperplane are used to quantify the impact that each feature has on the model and thus its importance. Figure 23A shows the relative importance of the top 21 features used in computing the knee health score. The majority of these features for classifying the two classes are in the spectral domain which agrees with the results from our earlier pilot work on the topic [104].

To assess model performance, the number of features and cycles were varied in training the model to quantify the change in accuracy that the inclusion of each consecutively less important feature and each recorded cycle had on the classification accuracy of each subject. The output of this testing is visualized as an accuracy heatmap in Figure 23B where the color represents the average accuracy from testing on each subject in the dataset using LOSO-CV using the depicted number of features and cycles of movement. At the bottom left of this plot is the accuracy of the model when only trained on the most important feature – the mean spectral spread - and tested on just one randomly selected cycle of flexion/extension from the subject. All permutations of possible cycle selection were performed and averaged to yield the accuracy under these conditions. In the case of just one cycle and one feature, the average cycle classification accuracy was only 11.1%. Ascending along the y-axis, one feature is consecutively added based on its relative importance, such that at the top left corner of the heatmap the model

has been trained on the top 20 most important features. Still, when tested with only one cycle from a subject, the accuracy remains low at 25.0%. From left to right, the algorithm is tested on an increasing number of cycles recorded from a subject. The model has an accuracy of 42.8% in the bottom right corner, where it was trained on just the mean spectral spread and tested using all recorded cycles of a subject from all 4 microphones. The algorithm had the highest accuracy of 80.6% when trained on the top 20 most important features and tested using all recorded cycles. This is slightly less than the 82.7% observed in Figure 22. This discrepancy is because the model in Figure 22 had the added benefit to the classification of all forty-nine features, not only the top twenty most important (Figure 23).

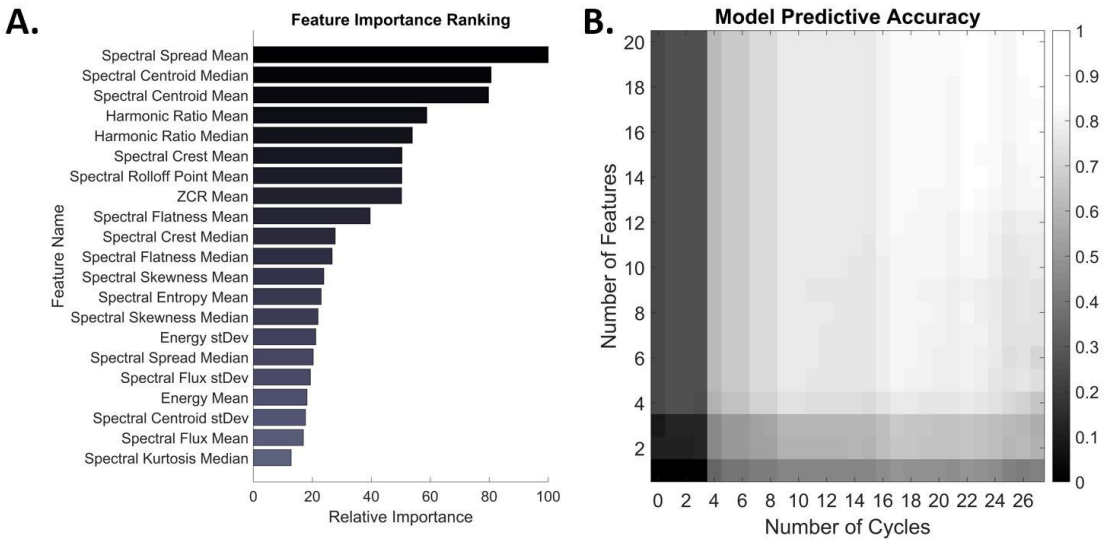

Figure 23. Feature Importance and Model Performance Based on Number of Features and Cycles. (A) Features are ranked based on their weighted coefficients as output by the trained logistic regression model. The most important feature was the mean spectral spread. (B) The model was trained on a feature set containing just one and up to twenty of the top features and the accuracy was assessed based on including those features and number of cycles recorded from a subject. The colors represent the average accuracy across all subjects for all permutations of cycle selection for a given set of testing parameters. The maximum accuracy of 80.6% is seen in the top right corner when trained on the twenty most important features and tested on all cycles of a given subject.

6.3.3.5 Knee audio score's longitudinal health tracking capability.

The knee audio scores were calculated for ten of the subjects with JIA before and after 6 weeks of treatment. At first visit, these subjects were either newly diagnosed with JIA, or having a resurgent flare of the condition. Their treatments were prescribed according to the current clinical standards by their treating pediatric rheumatologist and were recorded but not controlled for in this study. Every subject at follow-up reported a reduction in symptoms and the treating physician reported an overall improvement in their symptoms and disease status. In Figure 24, the calculated joint health scores are shown before and after treatment for this cohort. The average joint health score at initial visit was $0.84 \pm .08$. At follow-up, the scores dropped to an average of $0.19 \pm .09$. This drop in joint health scores is statistically significant with a p-value <0.001, when tested with a one-tailed t-test. The individual subjects scores are represented with dashed lines in Figure 24 and in all cases mirror the clinical assessment of their improvement.

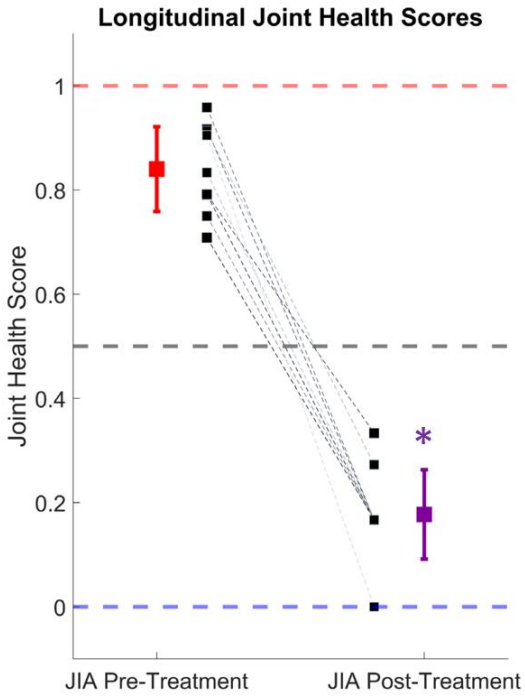

Figure 24. Longitudinal Joint Health Score Tracking. The average joint health score, which describes the probability of having JIA, dropped from 0.84 ± .08 to 0.19 ± .09 after successful treatment of the condition in ten subjects. The individual subject scores are denoted by the black squares and dashed lines. The mean and standard deviation of the actively inflamed subjects with JIA is shown in red, and the purple marker indicates the mean and standard deviation at follow-up. This drop in joint health score was statistically significant ($p<0.001$)

6.3.4 Discussion

There is a compelling need for the development of a non-invasively measurable biomarker that can both diagnose and track the status of affected joints in JIA. Assuming a child is properly diagnosed, determining which treatment regimen will work best for them is largely reactive. This means that a certain course of treatment is prescribed and adjusted based on patient-reported feedback and infrequent clinical assessments. In this work, we explore the impact that joint AE monitoring could have on the diagnosis and treatment of JIA. If AEs were found to contain clinically relevant information, they could potentially be used as an initial screening tool by primary care medical professionals -- reducing the burden on the healthcare system of unnecessary referrals to specialists. Furthermore, this could help diagnose patients earlier, which may prevent the long-term sequelae of JIA [106]. After diagnosis, if joint sounds were found to closely track with treatment efficacy and joint health longitudinally, they could be used as an objective biomarker to decide or even predict the most effective course of treatment. This would reduce the burden of frequent JIA flare-ups on patients and allow for a tightening of the treatment feedback loop leading to overall better management of the condition.

In this study, the effects of JIA on the AEs produced by articulation of the knee were explored. The study population was made up of 43 subjects, 25 of whom had JIA, and ten of whom had repeat recordings six weeks after the initial visit. The AEs from a pediatric population with JIA of this size have never before been compiled and analyzed. These AEs were first compared qualitatively to better visualize the differences in the recordings as

seen in 6.3.3.1. It was noted that there are characteristic high-frequency clicks in the AEs of subjects with JIA that fade away with successful treatment and are not present in matched healthy controls' AEs. More work is needed to determine the origin of these clicks, but we hypothesized that they occur due to increased internal friction in the joint caused by the characteristic inflammation of the synovial membrane, breakdown of cartilage, and reduced joint space in JIA[34], [111]. Of note, similar clicks are apparent in the case of acute injury as was recently discovered by our work in a cadaver model of knee injury[98] and a similar study in an injured athlete model[68]. Rather than relying strictly on one or even a few characteristics of these AEs as was done in previous work, in this study we attempt to more thoroughly quantify the differences between the recorded AEs. We do this by splitting the joint sound recordings from each subject into their component flexion/extension cycles. On each cycle, forty-nine features (from the spectral and time-domain) were calculated to describe the observed AEs. These features and cycles were organized into a feature matrix which was used to train a machine learning, classification model using logistic regression. This technique should provide a more exhaustive analysis of the features of the AEs, and overall be more generalizable than past efforts to interpret AEs. The results of this model are described below.

6.3.4.1 Knee audio score classification.

Logistic Regression is a binary classification algorithm that attempts to find the best hyperplane in k-dimensional space for separating the two classes (e.g. healthy and JIA) while minimizing logistic loss [71]. Logistic regression outputs the probability that a given test cycle belongs to the healthy or JIA class. In this case, if the probability of JIA is ≥ 0.5, that input cycle was labeled as JIA, if <0.5 it was labeled healthy. This threshold

could be adjusted in the future to favor an increase in the sensitivity or specificity depending on the use case. This concession may lead to more false positives in the classification, but since the outcome of a false positive test result are relatively minor compared to the potential damage from a missed diagnosis, this decision was appropriate in this case. We have also proposed that the estimated JIA probability could be used as a basis for quantifying knee joint health. In this paradigm, a probability of 0 indicates a healthy knee with no signs of JIA, whereas a score of 1 indicates a knee clearly inflicted with JIA. The classification accuracy of the model is presented in 6.3.3.3. First, the subject-level classification histogram showed clear separation of the joint health scores when the 0.5 classification threshold was applied to the output probabilities (Figure 22.A). This finding helps support the idea that knee AEs could be used as part of the screening and diagnosis of JIA. The accuracy of labeling each cycle is then quantified to better understand the performance of the logistic regression model (Figure 22**.**B,C). The overall accuracy of the cycle labeling was 82.7%, which corresponds to a sensitivity of 84.5% and a specificity of 80.4%. As discussed, JIA is difficult to diagnose not only due to the highly variable nature of the condition and presentation, but also because of the shortage of pediatric rheumatologists who are specially trained to identify the disease. One potential use of AE-based assessment in JIA is to allow for better screening of the condition by healthcare providers that are less trained to identify it. AE based assessment is entirely non-invasive and achievable with affordable hardware. The high sensitivity of this technique means that few false negative test results will occur. In the future, AEs could at least be used as a preliminary screening tool that gates whether a patient should pursue a specialist consult for further diagnostic workup.

6.3.4.2 Feature importance ranking and model performance

In order to understand the effects of feature selection and length of recording on JIA AE assessment, we presented our findings on which signal features are most important for the algorithm, and how it performs with less cycles to classify using a subset of features. In our model, there were 49 features describing each cycle of movement from each subject. A feature weights vector of length 49 was output from the model describing the hyperplane that best separates the JIA from healthy labeled cycles. The absolute values of the individual feature weights were used to quantify the importance of a given feature for the model. The relative importance of the top twenty features in the algorithm are presented Figure 23.**A**. These results were highly-consistent with our team's earlier pilot study on a much smaller sample of patients with JIA [104]. Each subject had two microphones on each of their knees recording the AEs during ten cycles of flexion/extension at a rate of 1 cycle every four seconds. These four audio files were subdivided into the individual cycles of movement based on the simultaneously recorded motion data captured by the inertial measurement unit (IMU) attached to the subjects' ankles. The resulting data structure thus had approximately 40 segments of data describing one subject's movement. Figure 23b graphically depicts the results of varying the number of those segments included in the testing dataset. Each square in Figure 23b describes the average accuracy when each subject was tested with the described parameters as a part of LOSO-CV on the trained model. Along the y-axis, Features were sequentially added in order of descending importance, such that at the bottom of the plot, only the most important feature – the mean spectral spread – was used to classify the cycles. As the y-axis is ascended, each of the 20 features as described in Figure 23A are consecutively included in training the logistic

regression model. This figure thus depicts the impact that feature selection has on the accuracy of the classification. There is a clear benefit on the accuracy of the model by including more features, and this should help with the generalization of the model to novel data. In the past, attempts have been made to describe knee AEs using only one or a few different signal features [15], [110], [112]. These attempts generally have success on a small data set, but when applied to a data set of this size were suboptimal when compared to the accuracy of the model proposed in this work.

The impact of the length of the AE recording is also demonstrated in Figure 23B from left to right. Each step to the right includes an additional, and randomly selected, flexion-extension cycle, and the color of the square indicates the accuracy of classifying a subject with that many cycles. On the left, we test the model with only one cycle recorded from one microphone on each subject. On the far right, every cycle recorded for every microphone is used to test any given subject. The impact is similar to increasing the number of features in the trained model – as the number of cycles increases the classification accuracy similarly increases. It should be noted that there is some possible redundancy in having two microphones recording the AEs from each knee. Overall, this analysis demonstrates the impact that the feature selection and length of AE recording has on the accuracy of the model. In our case, the accuracy was at its lowest with one feature and one cycle at 11.1% and achieved a high of 80.6% with the top twenty most important features and every recorded cycle from a subject. It can also be seen that the performance plateaus after the addition of around half of the cycles. This lack of change in the accuracy with additional recordings indicates that the model is relatively robust to varying lengths of AE recordings as long as a minimum number of cycles are met. This analysis also demonstrates

why past approaches have had only limited success in generalizing their findings. If only a subset of these features were used to describe AEs, the accuracy would significantly diminish. Many features are needed to fully describe the nature of these sounds and separate the differences between populations. Later work comparing a different clinical scenario, or a larger dataset may find that a different feature is more important for delineating two study groups, but the approach applied in this paper should hopefully provide guiding influence on future assessments of AEs.

6.3.4.3 Longitudinal joint health tracking

In order to discover if knee AEs had the potential for quantifying joint health longitudinally, ten subjects with JIA had their AEs recorded during an active flare-up of the condition and 6 weeks later at their follow-up visit. In this particular cohort, every subject showed clinical improvement and reported a lessening of symptoms. To calculate these subjects' knee scores, the logistic regression model was trained on all subjects not in this cohort. The recordings before and after treatment were tested on the trained model and the knee audio scores computed as described in 6.3.2.4. The hypothesis was that as a child's knees healed from effective treatment, their knee scores would decrease from the JIA range (0.5-1.0) toward the healthy range (0.0-0.5). In all subjects, this hypothesis was shown to be valid. There was a statistically significant drop in the average scores of 0.65, or a 77.4% improvement in the joint health score. This closely tracked with the reported clinical workup of the subjects indicating that joint health scores based on AEs may be clinically applicable for not only diagnosing JIA (as discussed in *6.3.4.1*), but also monitoring the condition over time.

In this study, these ten patients represent a subset of the overall JIA population and before claiming how consistently joint sounds track with knee health status in an individual the sample size of those studied should be further increased. However, these findings represent the first time that a population large enough to adequately power a study of children with JIA has been studied longitudinally. The close correlation between the change in joint sounds and the observed clinical status will hopefully support further research into this relationship. Overall, this study represents an early, but important step toward understanding the nature of joint AEs. The strong separation of the classes alongside the close tracking of disease activity make it clear that joint AEs contain clinically relevant information. This information if properly leveraged could one day enable better more personalized treatment of JIA.

6.3.4.4 Motion Data Comparison

In the comparison of JIA to healthy controls, there was a statistical significant reduction in the joint health score. The same occurred from the JIA case to those patients at follow-up. It appears then that AEs can be used to both longitudinally track and diagnose JIA involvement in the knee. However, it could be possible that a change in the technique of flexing and extending the leg could be confounding the results. In order to test this, we look at four features that describe the movement of each subject: the RoM, angular speed, cycle duration, and angular speed variability.

Each child is asked to watch and copy the movements of an instructional, cartoon video showing a flexion/extension cycle at a rate of 1 cycle every 4 seconds. This is a relatively slow pace, but one found heuristically in the cadaver experiments to produce the

best AEs. The only significant difference in the movements between the JIA group and the healthy control group was in the cycle duration. The angular speed, range of motion, and angular speed variability were not significantly different. The cycle duration however can not completely explain the gap in the joint health scores of the healthy subjects and those with JIA because the cycle duration was not significantly different between the JIA group and their follow-ups (where a similar drop in joint health score was observed).

Of the groups assessed, the children with JIA are expected to have the smallest range of motion – given the pain and symptoms in the knees recorded. With clinical improvement, their RoM would be expected to increase at follow-up. We however measured a drop in the RoM at follow-up. Interestingly, the follow-up subjects' angular speed also significantly dropped. These significant changes in their movement could be the cause of the change in joint health index calculated. However, we also observed a similar change in the mean RoM and angular speed between the healthy controls and the follow-up group, and there were not significant differences in the joint health score between these two groups.

Thus, rather than the changes in movement features being related to JIA status (our variable of interest), it appears they are more closely related to a learned effect. The follow-up group had already watched the video that showed them how to control their flexion/extensions. At follow-up, the patients were not as nervous and more confidently reproduced the slow movements seen on the instructional video. Since they are moving slower (lower angular speed), but there cycle duration has not significantly changed, the follow-up subjects are going to move through a smaller range of motion. The ideal angular speed is 45° per second, which equal one cycle of 90° every 4 seconds as is shown in the

video. The FU group's average angular speed was closest to that at 51°±9°. Together, this evidence indicates that the changes seen in the IMU data were more related to a learned effect rather than a limitation due to the JIA status. Additionally, there were no shared statistically significant motion features that could explain the shared drop in the joint health index based on AEs seen in the follow-up and healthy cases. In the future, to fully test this hypothesis, a cohort of healthy patients will have their joint sounds recorded a second time and their motion data re-recorded. If these healthy follow-ups have a similar change in the motion as the JIA follow-ups, we can conclude that it is a learned effect rather than an unknown confounding factor in AE analysis.

6.3.4.5 Limitations of results and steps to clinical adoption.

JIA is a chronic condition that affects multiple joints in the body. The knee is one of the most commonly affected joints and made for a viable target for this attempt at analyzing AEs. To better understand the clinical utility of this sensing modality, AEs should be studied in other commonly affected joints in JIA. Additionally, the sensitivity of this method should be compared against the performance of the current clinical standard procedure for diagnosing and staging the condition. Treatment of JIA seeks to reduce the frequency of acute, symptomatic flare-ups, and to ultimately achieve clinical remission. In this study, the treatments our subjects underwent were not controlled for due to the small sample size. In the future, the effectiveness of therapy should be quantified using a prospective study design. Additionally, in this cohort all subjects improved with treatment and we saw a corresponding drop in the joint health score. Since no patients got worse at follow-up, we were unable to discover if AE assessment could track worsening of the condition. The sensitivity of joint sounds for detecting not only different severities of the

condition but also the course of the condition should also be assessed. AEs would present significant clinical significance if they were able to determine the difference between an acutely inflamed joint and a more chronic, undiagnosed state. Determining that duration of disease activity would help with selecting the ideal treatment for a patient. Classifying subjects into the different subtypes of JIA and delineating joint sounds caused by JIA versus all other causes would also offer clinical merit. Finally, though this study was performed on the largest sample size of subjects to date, increasing the enrollment would better support the generalizability of the discussed results. Joint AEs are a novel technique for analyzing the health of a joint. The findings in this paper present significant clinical merit to this type of analysis, but there is still much to be discovered.

CHAPTER 7. SCREENING IN THE TMJ

7.1 Introduction

7.1.1 AE Opportunity Space

Assessment of the temporomandibular joint (TMJ) can be difficult; clinical signs and symptoms are non-specific[25], examination is challenging and imaging is often necessary [26]. TMJ disease in children can cause pain and growth disturbances leading to malocclusion and/or skeletal deformities [6],[64]. The presentation, difficulty in diagnosis, and severity of sequelae of untreated disease present a compelling need for the development of a biomarker for TMJ health [6]. Ideally, this biomarker would be objective, noninvasive, and readily measurable with affordable hardware. AEs (AEs) from the TMJ could serve as such a biomarker. They contain information related to the structural integrity of the joint and the health of internal articulating surfaces [24],[98]. Changes to AEs could serve as an objective diagnostic indicator of TMJ pathology.

7.1.2 Previous Work on TMJ AE Analysis

AEs from joints were first reported in 1902 by Blodgett [18]. In the 1930s, Steindler correlated joint malfunctions and sounds using several types of sound detecting equipment [113]. In 1961, Brackin filed the first patent detailing an apparatus for recording and analyzing joint disorders with unique acoustic patterns recorded from different pathologies [114]. These attempts to facilitate diagnostic procedures by microphonic detection of

emissions did not gain widespread use because of discrepancy in the nature of the sounds and the recording technique. In 1984, Molan found that the use of a piezoelectric accelerometer detector in direct contact with the skin gave a robust signal and allowed for detection in the subsonic frequency range [115]. Five years later, Gay filed a patent for a diagnostic procedure and apparatus that quantitatively correlated joint-induced sound patterns relative to the joint position in time, and noted that it could be particularly useful in diagnosing TMJ disorders [116]. Gay's technique was the first to move away from qualitative descriptors of the joint sounds to quantitatively comparing the sound profiles.

Prior to the 1990's, joint AE analysis was limited by the computational power and by the physical size of the sensors, so research focused on larger, more accessible joints (e.g. the knee). As a result of those limitations, comparisons were often qualitative and inconsistent between researchers. The advent of miniaturized sensors and the increasing computational power of the 1990's presented the opportunity for more powerful (and quantitative) AE analysis of smaller joints (e.g. the TMJ). Since then, two main approaches for recording TMJ AEs have gained prevalence in the field: binaural miniature microphones placed at the intra-auditory meatus and contact accelerometers placed on bony prominences around the joint [117]. Microphones at the intra-auditory meatus provide a broad signal-to-noise bandwidth, while contact accelerometers provide the highest mean amplitude in the time domain waveform [65]. Either approach is suitable depending on the application and nature of the underlying signal being recorded. In our project, we used surface mounted accelerometers because they are easy to place on TMJs and are able to capture high amplitude spikes in the AEs, as discussed in 4.1.

Significant steps have been made in the quantitative classification of these audio signals. Prinz showed that the time domain is where most of the characteristic differences of the various TMJ AEs are found, and that the frequency domain was much less distinct than the time domain [96]. To study key signal features in the time domain, several computationally rigorous approaches have been applied such as the reduced interference-distribution (RID) of the time-frequency energy distributions and neural networks. The RID technique was shown to have a more detailed classification of TMJ AEs than by auscultation [118]–[120]. Neural networks were used to classify TMJ sounds based on their narrow-band, wide-band, and time-varying frequency components [121]. Previous research on TMJ AEs resulted in several patents for devices which capture TMJ AEs. However, this type of analysis has not gained widespread clinical usage perhaps because a standard protocol for best capturing and analyzing AEs does not exist.

7.1.3 The Goal

AEs of TMJs must be better understood, characterized and a standardized technique for recording and interpretation needs to be developed. The purposes of this project were to: (1) present our custom, wearable headset with embedded contact accelerometers and (2) assess the repeatability and reliability of our headset in children. We hypothesize that this headset will allow for the convenient recording of AEs, which will ultimately facilitate their inclusion as a biomarker in a clinical workup of the TMJ. The work presented here is an early, but crucial step toward the design of a system for augmenting the current diagnosis and monitoring of TMJ disease in children.

7.2 Assessing the Feasibility and Repeatability in Children

7.2.1 Subject Recruitment

IRB approval was obtained (#00081670), and all subjects were recruited in accordance with the Helsinki Declaration guidelines. Inclusion criteria consisted of children age 6-18 years of age who presented to Oral and Maxillofacial Surgery (OMS) clinic for a non-TMJ related reason. The presence or absence of TMJ sounds was verified via a clinical examination by a board-certified OMS. Exclusion criteria were: (1) systemic disease which has potential to affect the TMJ (e.g. juvenile idiopathic arthritis, [JIA]), (2) history of craniofacial syndromes with potential for TMJ involvement (e.g. hemifacial microsomia), (3) history of TMJ/facial trauma, (4) ongoing orthodontics, and/or (5) complaints of temporomandibular joint dysfunction (TMD).

7.2.2 Device/Headset Setup

When a subject opens and closes his / her jaw, TMJ articulation creates vibrations that are detectable on the surface of the skin. We built a headset adjustable to fit 95% of users younger than 18 years old based on anthropometric head circumference data[122]. The headset is positioned on the subjects' heads with skin contact accelerometers against the articular eminences of TMJs (Figure 25) [123], [124]. This location and skin contact previously demonstrated detection of TMJ sounds with the highest quality time domain waveforms [65]. This method provided sufficient contact force without hindering portability of the device or causing discomfort [117].

The AEs are recorded using the setup described in 4.1 fit inside the headset as described above. After each recording, the software preliminarily filters (between 250 Hz – 10 kHz) and plots the recordings, so that the researcher can ensure everything was

functioning properly. This filtering range isolates the frequencies containing the majority of TMJ AE signals, and removes artifacts associated with large-scale movement of the jaw, low frequency muscle sounds, and environmental noise [96], [104], [117]. With the setup in place and the software running, the subjects perform 10 repetitions of opening/closing their mouth at a rate of 1 repetition every 4 seconds (Figure 25).The raw and filtered data were recorded and locally stored for further processing.

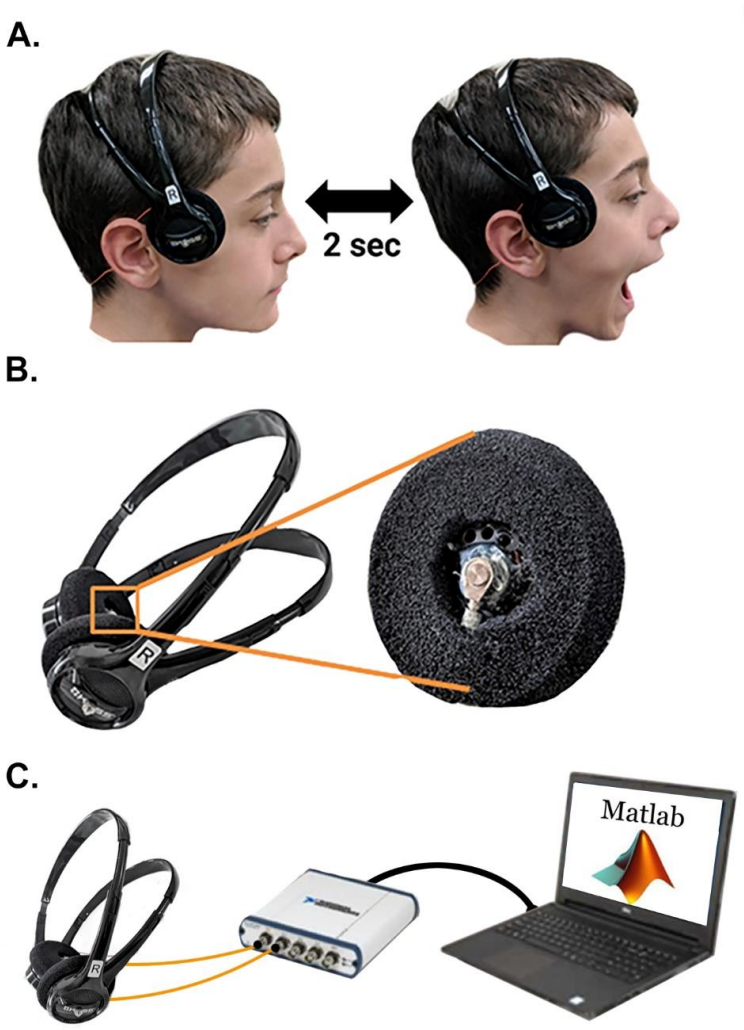

Figure 25. Recording Setup Used for Capturing TMJ AEs. (A) Each subject wore the headset and performed 10 repetitions of opening and closing their mouths, at a rate of 1 cycle per 4 seconds while watching an animation to help maintain consistent speed and movement. (B) AEs were recorded from both TMJs simultaneously while performing the exercises using uniaxial accelerometers embedded into a headset form-factor for convenient placement superficial to the TMJ. (C) The signals captured by the accelerometers in the headset were acquired using an data acquisition unit, and analyzed via a connected laptop computer running Matlab.

7.2.3 Feasibility and Repeatability Testing Protocol

To assess the feasibility of using our TMJ AE recording headset, AEs from one healthy control (i.e. no TMJ sounds) and one patient with clinically noticeable TMJ sounds were recorded. These recordings were qualitatively compared to ensure that there were differences in the sounds and that the headset was recording AEs properly. To assess the repeatability of the recording device, 9 subjects performed three trials of open/close movements while their AEs were recorded. Between each trial, the headset was removed and repositioned on the subject's head to test for repeatability of the placement of the device.

7.2.4 Analysis Technique

To analyze repeatability of measurements from the AEs of TMJs, we calculated three features that describe the signals in the time domain: the root mean square (RMS) power, the signal energy, and the zero-crossing rate (ZCR). The RMS power is a measure of the absolute value of the magnitude of the signal, so signals with larger spikes would be expected to have a larger RMS power. The energy feature is computed as the integral of the squared signal magnitude. This feature describes how "loud" the audio signal is. The ZCR describes how often the signal crosses zero, which estimates how quickly its values change. We are using the ZCR to quantify how often the signal is moving from negative to positive and back indicating a change in direction as the skin vibrates. If the skin was vibrating back-and-forth faster, then the ZCR would increase. All together, these three features comprehensively describe the qualitative differences that we observed.

Repeatability of measurements on each subject was calculated using the intra-class correlation coefficient (ICC). The ICC indicates how strongly the different sessions of TMJ recordings resemble each other. The ICC varies from 0 to 1 (1 indicates completely the same, 0 indicates no overlap) with values above 0.9 typically representing excellent repeatability [125]. We calculated the ICC for each of the features we selected to describe the signal (i.e. the RMS power, energy and ZCR) for all trials. Each TMJ (left and right) of a patient was a separate group. We did this because we were not trying to compare the features of different joints, but rather ensure that the device was recording a repeatable signal from each specific TMJ. There is inherent inter-subject and intra-subject variability in the AEs of each TMJ since each individual TMJ has unique anatomy and kinematics (Figure 3).

7.2.5 Results of Feasibility and Repeatability Testing

To test the headset's recording capabilities, recordings were obtained to ensure the device was working properly. We recorded sounds from TMJs of a healthy subject with TMJ sounds and sounds from TMJs of a healthy subject without TMJ sounds. There are several qualitative differences between the two subjects' recordings (Figure 26). The patient with sounds had large spikes (with amplitudes of ~0.1 mm/s^2) that occurred approximately every four seconds (Figure 26*B*), These spikes sounded like loud clicks or pops when listening to the recordings. These sounds were occurring at the same point in the articulation of the jaw during each cycle of opening and closing. In addition, the TMJ sounds were more heterogenous and variable than the ones from the child without TMJ sounds. The child without TMJ sounds had numerous smaller spikes in the sound (with

magnitudes of ~0.5 mm/s^2). When listening to these smaller spikes, they resembled a grinding sound.

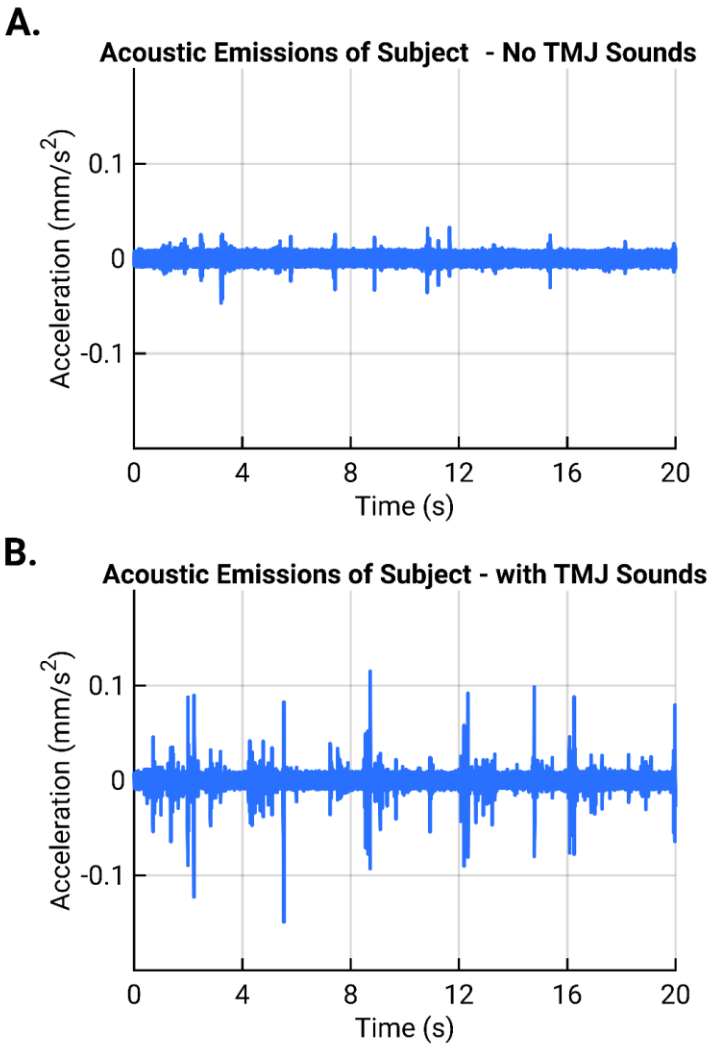

Figure 26. TMJ AE Feasibility Recording. (A) Time domain recording from a subject without TMJ sounds. (B) Time domain recording from a subject with TMJ sounds. Each spike in the acoustic signal represents a large click or pop.

Next, nine healthy children (6 females, 3 males) with mean age of 10.8 ± 3.2 years (range, 7 to 16 years) had their AEs recorded in order to assess the repeatability of TMJ

AE recordings. The three signal features discussed above (RMS power, energy, and ZCR) were calculated for each of the three recordings from each TMJ on all the subjects. The goal of this analysis was to quantify how similar the signals from each recording sessions were for each subject. A representative example of the three recording trials for one subject can be seen in Figure 27*A.* The distribution of feature values across all the recording sessions and subjects can be seen in Figure 27*B-D.* Of note, though the individual feature values vary from subject to subject, the three sessions' features were tightly clustered for each individual TMJ for all subjects. This tight clustering of feature values indicated that the signals were repeatable. To further quantify this repeatability, the ICC values are presented in Figure 27*E.* The ICC values were 0.96 for the RMS feature, 0.91 for the energy feature, and 0.995 for the ZCR feature. As discussed above, an ICC score >0.9 is considered to represent excellent similarity of the signals being assessed. Here, it indicated that the AE recordings are highly consistent across multiple recording sessions and placements of the headset.

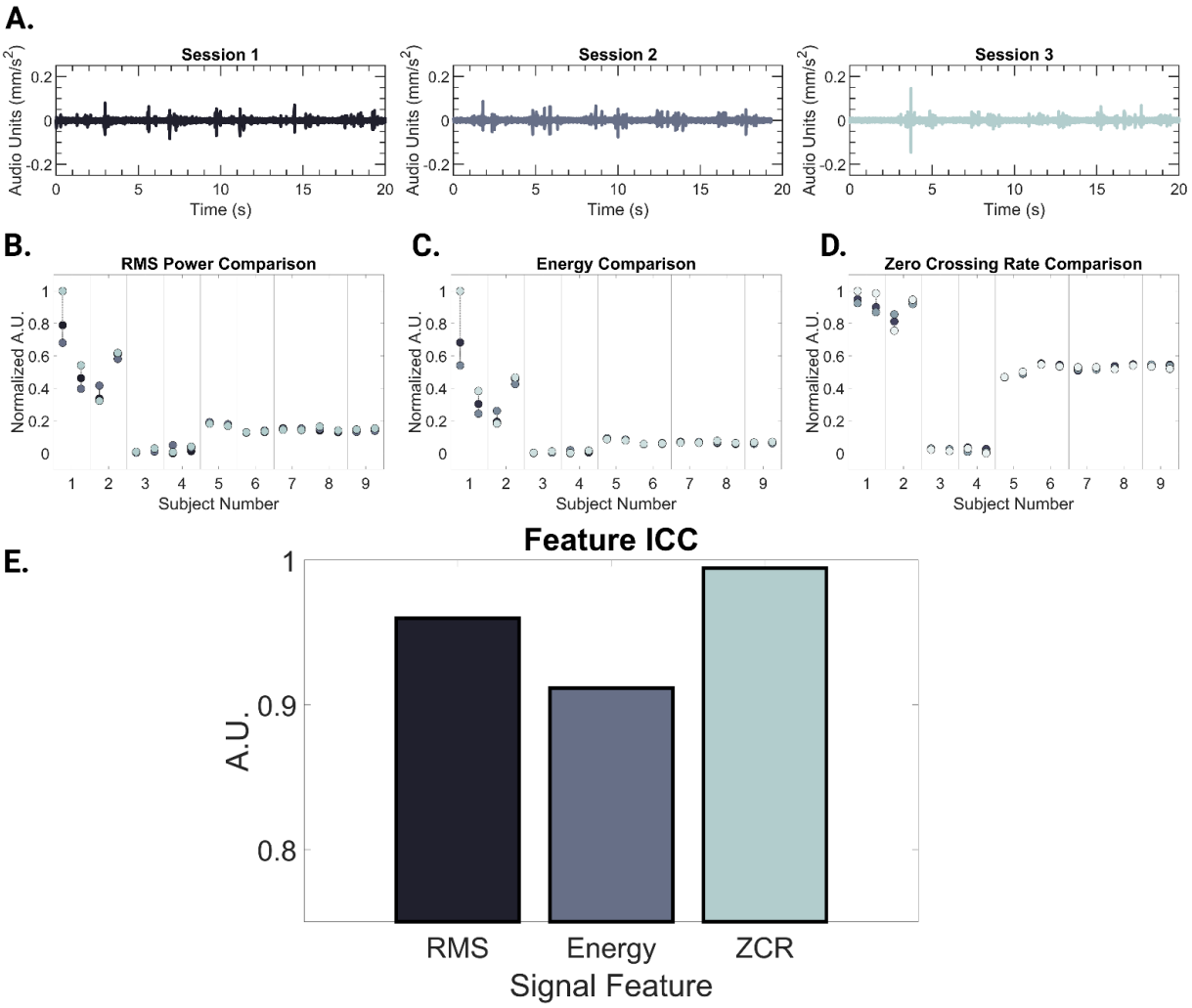

Figure 27. Repeatability of TMJ AEs. (A) Example time domain recordings from the three sessions of one subject. (B-D) The RMS power, energy, and ZCR for the three recording sessions of each subject show that there was very little change from one recording to the next. The recordings from the left TMJ are on the left in each subject number division, and likewise the right TMJ data are on the right. (E) The ICC values of each feature presented in B-D; each ICC value is >0.9, so signals have excellent repeatability between recording sessions.

7.2.6 Discussion of Feasibility and Repeatability Results

TMJ health is evaluated by a combination of physical exams and imaging studies. Physical exams rely on health care worker expertise. Imaging is not always feasible due to its high cost, need for occasional sedation in children, length of time, need for specialized equipment (e.g. magnets), and potential contraindications[5], [6]. A TMJ AE headset has the potential to serve as a screening tool prior to obtaining imaging. The purposes of this manuscript were to (1) present our custom, wearable headset used to record AEs of TMJs, and (2) assess the repeatability and reliability of this headset in children.

The technique for measuring TMJ AEs has evolved since it was first proposed in 1902 [18]. The field has progressed from manual auscultation, digital stethoscopes, condenser microphones, electret microphones, and now favors miniaturized contact accelerometers [14], [15], [65], [114]–[117]. Our headset is based on findings of earlier work in selecting an ideal accelerometer with high sensitivity, and a bandpass filter to remove confounding low frequency muscle sounds and environmental noise[126]. It was designed to obtain the highest amplitude signal in the time domain – which contains the majority of the characteristic differences in TMJ AEs[96]. Our device places the accelerometers superficial to the TMJ[117] and was designed specifically for children who are likely to be uncomfortable with an intra-aural device. This sensor location and comfortable form-factor minimized the time required to place the sensors accurately and firmly on the TMJs. The acquisition software was written to minimize computational time. Together, the form-factor, hardware, and recording scripts allowed for reproducible recordings of TMJ AEs with minimal time required for setup and acquisition (< 2 minutes).

Minimizing the time needed to assess the joint is of critical importance for a busy clinical setting.

Before exploring the diagnostic capabilities of our headset, we needed to confirm that its recordings were repeatable and consistent. In order to quantify this repeatability, we calculated three time-domain signal features: the RMS power, zero-crossing rate, and energy. It was previously shown that time-domain features contained nearly all of the characteristic differences of TMJ AEs [96]. In particular, the energy of the signal has been used extensively to describe characteristics of TMJ AEs [119], [121]. In our study, the ICC values were all >0.9, which indicated high consistency from one recording session to the next; thus, excellent repeatability (*Fig. 3*). These findings support the claim that this wearable headset can consistently record AEs from the TMJ of children.

When listening carefully to these sounds, we noticed that sounds occurred at the same point in the articulation of the jaw during each cycle of opening and closing. We hypothesized that the cyclical occurrence of these loud sounds may indicate that there is an anatomical variation producing them. The TMJ sounds produced by the patient without clinically-evident sounds may simply indicate friction of the TMJ during articulation.

This study has a few limitations. Although the headset was removed multiple times, AEs were recorded during the same visit. This study shows that TMJ AEs can be successfully and repeatedly captured by a wearable headset. However, it does not address the variability in sounds overtime as disease progresses. Additionally, all the subjects recorded in this study were healthy with no history of TMJ dysfunction. This resulted in relatively small AEs, since the TMJs of healthy children are not expected to produce much

sound. In the future, to better understand the feasibility of this technology for clinical diagnosis, it will be applied to children with systemic disease known to affect TMJ such as JIA and may be compared to MRI findings. This is the subject of an ongoing investigation in our center.

In conclusion, this project provides the foundation for the eventual clinical use of a TMJ AE device. In the next section, we use this technology on a cohort of patients with JIA and age/sex matched healthy controls to evaluate the effect of arthritis on the AEs of TMJs. In a chronic condition such as JIA, AE assessment may extend beyond just screening/diagnostics and instead be used as a longitudinal biomarker of disease activity within the joint. Overall, these exciting preliminary results should inspire further research into the acquisition, analysis, and classification of TMJ AEs.

7.3 TMJ AEs as a Screening Tool for JIA

With the feasibility of recording TMJ AEs from a pediatric population using our custom headset proven, we now move on to understand if AEs could detect inflammation caused by JIA in the joint. The TMJ is one of the most commonly affected joints in juvenile idiopathic arthritis (JIA) (up to 45% of cases) [25]. There is a discrepancy between clinical signs and presence of arthritis of the TMJ, which makes recognizing involvement and effective intervention difficult [32]. In order to diagnose JIA in the TMJ, combined imaging studies are necessary, but they are time consuming, expensive, and may not show involvement until the disease has sufficiently progressed [5], [6]. Crepitus is a common complaint in the TMJ, but it is still poorly understood [32]. In this work, we build off of our earlier understanding of AE analysis, and attempt to determine if AEs could be used as

a biomarker of JIA in the TMJ. This is done by recording the TMJ AEs from 4 groups of patients: two healthy control groups without JIA – one with and one without observable TMJ sounds, and two groups of children with JIA one with and one without observable TMJ sounds. The AEs from each of the groups will be compared. There are two main goals behind this comparison: 1) determine if AEs can separate healthy patients from JIA patients, and thus be used as a screening tool for the disease, and 2) determine if AEs can differentiate "healthy" TMJ sounds from pathologic TMJ sounds.

7.3.1 Subject Recruitment

This project was covered under our previously acquired IRB approval (#00081670). Inclusion criteria consisted of children age 6-18 years of age who presented to the OMS clinic. The presence or absence of TMJ sounds was verified via a clinical examination by a board-certified OMS. Exclusion criteria were: (1) history of craniofacial syndromes with potential for TMJ involvement (e.g. hemifacial microsomia), (2) history of TMJ/facial trauma, and (3) ongoing orthodontics. Unlike our previous work, in this study subjects with JIA were included as one of the groups. A summary of the current enrollment statistics is seen in Table 2.

Table 2. TMJ JIA Study Enrollment Statistics

	# Enrolled:	Age:	STD:	# Male	# Female
HC no Sounds:	20	11.89	3.18	5	15
HC w Sounds:	8	15.13	2.36	4	4
JIA w Sounds:	9	13.00	2.06	0	9
JIA no Sounds:	14	12.50	2.77	5	9
Total:	51	13.13	1.40	7	37

7.3.2 Setup and Testing Protocol

The AE recording setup was the same as described in 7.2.2 and seen in Figure 25. With the setup in place and the software running, the subjects perform 10 repetitions of opening/closing their mouth at a rate of 1 repetition every 4 seconds. In order to synchronize their movements, the subjects watched an instructional cartoon while performing the exercises as seen in Figure 28.

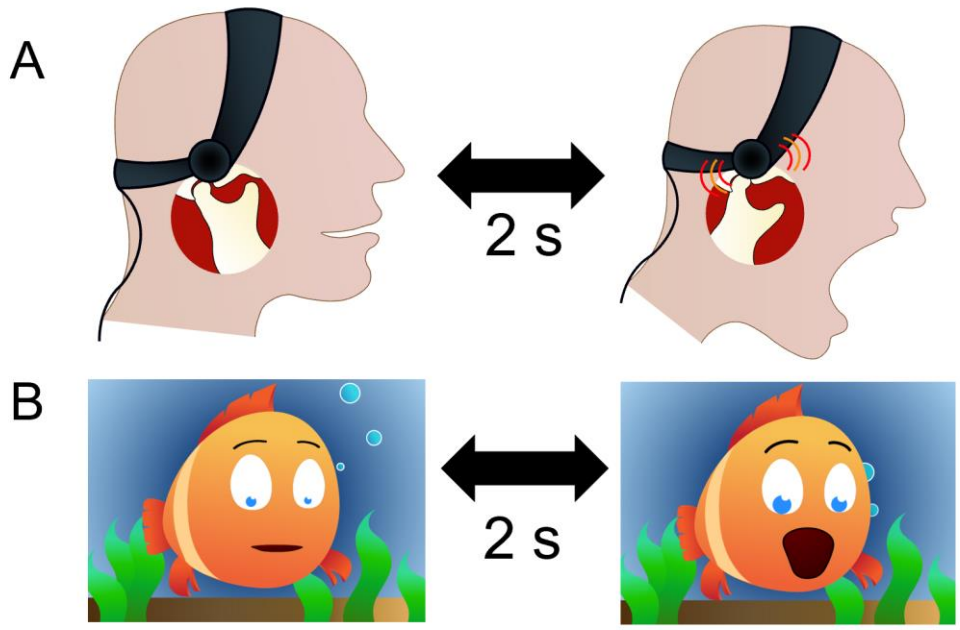

Figure 28 TMJ recording setup synchronized with instructional cartoon. A) Example graphic of a subject wearing the TMJ recording headset while performing open/close exercises. B) Animation video used to synchronize and standardize exercises for comparison across subjects. The animation performs the movement at the selected rate (1 cycle / 4 sec), so that children of all ages can easily reproduce the exercises.

7.3.3 Analysis Technique

Qualitatively, the differences in the AEs of the TMJ are more subtle than those seen in the knees. To analyze the AEs of TMJs, for each cycle of opening and closing of the jaw, the features described in Table 1 were calculated and stored in a feature matrix labeled $\boldsymbol{F_{TMJ}}$. Next, outlier cycles were removed from the data set. To determine if a cycle was an outlier, first $\boldsymbol{F_{TMJ}}$ was separated into 4 submatrices based on if that cycle belonged to a subject with or without JIA, and with or without clinically observed TMJ sounds. The median absolute deviation (MAD) was calculated for each feature in each of the 4 submatrices. The MAD is a robust measure of the variability of quantitative data and is calculated as described in Equation 6. Any cycle where the majority of the features were

more than three scaled MAD away from their respective feature median were labeled as outliers and removed from $\boldsymbol{F_{TMJ}}$.

$$MAD = median(|Feature_i - median(Feature), \quad for\ i = 1,2, \ldots number_{cycles} \tag{6}$$

Equation 6. Median Absolute Deviation.

With the outliers removed, the $\boldsymbol{F_{TMJ}}$ matrix was tested using logistic regression as described in 4.2.3. In this study, we were interested in two distinct binary classifications: 1) if the subject had JIA or not, and 2) if the subject had TMJ sounds or not. The classification accuracy for each of those labels was determined by testing $\boldsymbol{F_{TMJ}}$ twice using LOSO-CV. In the algorithm, JIA status was the first response variable. The presence of sounds was the second test's response variable. Logistic regression had sub-optimal separation of the classes, so three other common classification algorithms were tested on the feature set. The three other algorithms tested were: a cubic support vector machine, weighted *k*-nearest neighbors, and a bootstrap aggregated (bagged) decision tree. Each model's accuracy was calculated using LOSO-CV (as described in 4.2.6). The accuracy of these tests is seen in Table 3. Ultimately, the bagged decision tree model was found to have the best accuracy and used for the rest of the testing. Bagged decision trees combine the results of many decision trees, which reduces the effects of overfitting and improves generalization. In our algorithm, a random subset of predictors is selected to use at each decision split – similar to the random forest algorithm [128]. To better understand which features were most important for differentiating between the healthy AEs and the AEs from

a TMJ with JIA, the 49 features used in training the model were ranked in order of relative importance. The bagged tree algorithm and feature ranking technique is described in 4.2.4.

The goal of this project was to design a TMJ screening algorithm that would provide value to the clinical workup of these hard to diagnose patients. In order to do so, a dual decision algorithm is proposed as seen in Figure 29. In this algorithm, a subject's TMJ's are first tested to determine if they should be categorized as having or not having JIA. This step has the additional benefit of serving as a screening tool for JIA. Next, the AEs are assessed in the JIA group to determine if the subject has abnormal TMJ sounds or not. The flowchart in Figure 29B, could potentially replace the need of having an OMS perform a physical exam on the patient.

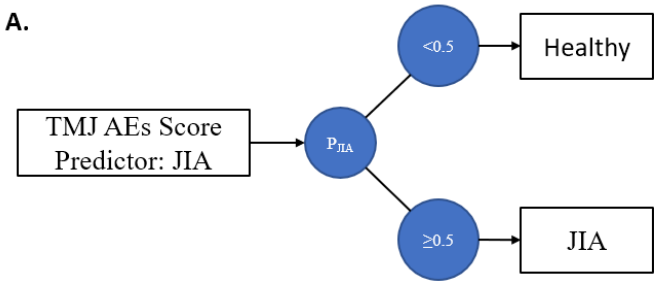

Figure 29. Flowchart of TMJ AE Screening in JIA. A subject's probability of having JIA is the average of the classification output for all cycles from that subject (A).

7.3.4 Results of JIA vs Healthy Control TMJ AE Comparison

7.3.4.1 Time Domain Qualitative Comparison

There are several differences that can be seen in the signals when compared against the patient with JIA with TMJ involvement. In the time domain, with the same rate of opening/closing the mouth the JIA patient's sound profile appears more chaotic, with

periodic large “clicks”. The difference in the patterns of the AEs of the other three groups are much more subtle. A comparison of a representative patient from each of the 4 experimental groups is depicted in Figure 30.

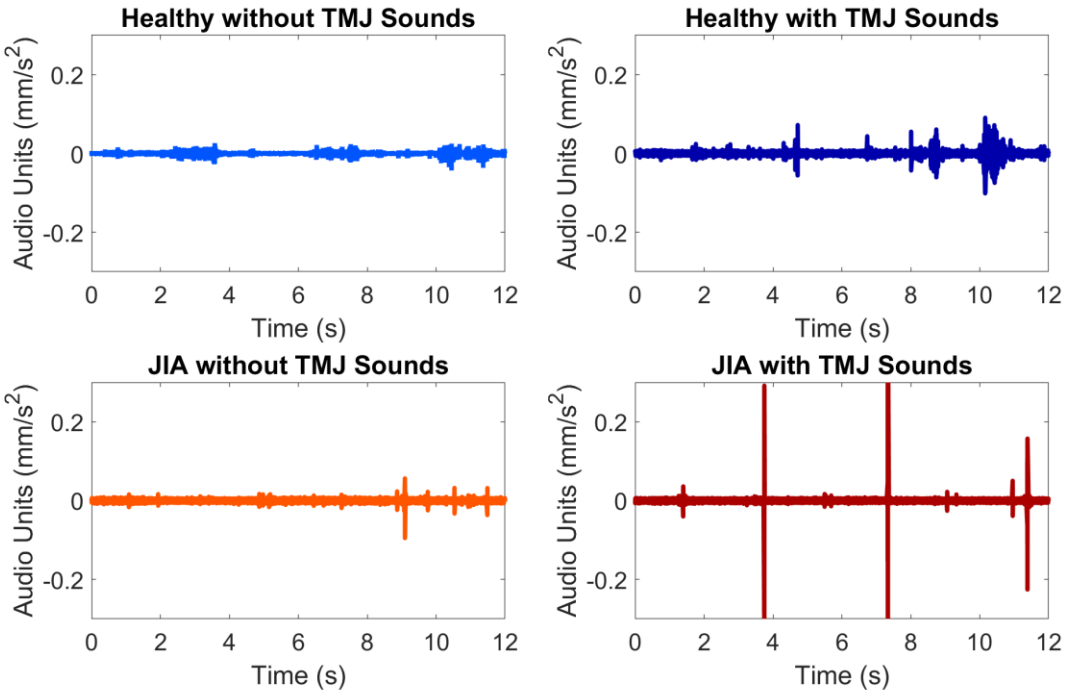

Figure 30 TMJ AEs from Four Subjects. Representative time domain signal of age and sex matched participants performing 3 open/close jaw movements**.**

7.3.4.2 Machine Learning Classification Accuracy

The goal of this work was to label the patients as having or not having JIA with or without TMJ involvement. As mentioned in 7.3.3 and seen in Figure 30, the differences in the AE patterns of the four groups were subtle. Four algorithms were tested to see which had the best classification accuracy. It was found that the bagged decision tree model, performed

best with an average accuracy of 68.6 ± 4.2%. The breakdown of the models' performances is shown below in Table 3.

Table 3. Accuracy of Algorithm Classification. The average classification accuracy was used to select the best performing algorithm. (Accuracy given in %)

	JIA v HC	Sounds vs No Sounds	Average
Cubic SVM	71.4	63.9	67.6 ± 5.3
Logistic Regression	71.3	54.9	63.1 ± 11.6
Weighted kNN	44.3	56.0	50.155 ± 8.2
Bagged Tree	71.6	66.6	68.6 ± 4.2

7.3.4.3 TMJ AEs to Screen for JIA

The features of all subjects with recorded TMJ AEs were used to train a bagged decision tree model, and the accuracy of classification was computed using LOSO-CV. This testing group including all subjects regardless of whether they had observed TMJ sounds in the clinic on physical exam. This model output the probability of a given cycle belonging to the positive class- 'JIA' in this case. These probabilities of all cycles belonging to each TMJ of each subject were averaged to give a "TMJ joint health score". A score of 1 indicates a TMJ with 100% probability of belonging to the JIA class. Likewise, a score of 0 indicates 100% probability of being healthy. These scores are plotted along the x-axis of Figure 31. With a threshold of 0.5, the accuracy of labelling patients as healthy was 79.2%, the accuracy of labelling patients as having JIA was 70.8%. This threshold could be adjusted in the future to prioritize sensitivity or specificity depending on the use-case. With a 0.5 threshold, the ROC curve had an area-under-the-curve of 70.9%. The scores and

classification of each subject are shown in Figure 31A. The computed ROC curve is shown in Figure 31B.

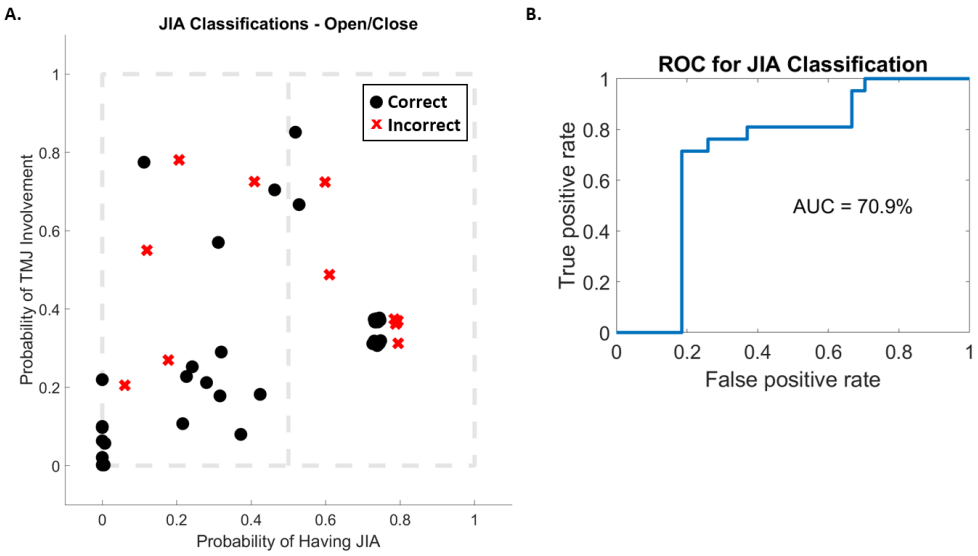

Figure 31. TMJ AE JIA Screening Accuracy. (A) Individual subject predictions. (B) ROC curve of JIA classification. AUC = 70.9%.

7.3.4.4 Difference in Healthy TMJ Sounds from JIA Sounds

When a patient presents to clinic and is complaining of or has observed TMJ sounds, it is difficult to determine if those sounds are pathologic, or from perhaps a benign anatomical variation. To test if AE analysis could be used to differentiate between those two cases, a second bagged tree model was trained using patients that had observed TMJ sounds regardless of whether or not they had JIA. In this way, we can determine if AEs from patients with JIA differ from those from healthy patients. Again, the probability of JIA was used as a TMJ health score. It was found that the cycles from healthy controls had an average TMJ health score of 0.56±0.25 and the cycles from JIA subjects had an average of

0.73±0.27. When tested using a 2-tailed, unmatched t-test, these score distributions were significantly different ($p<0.0001$).

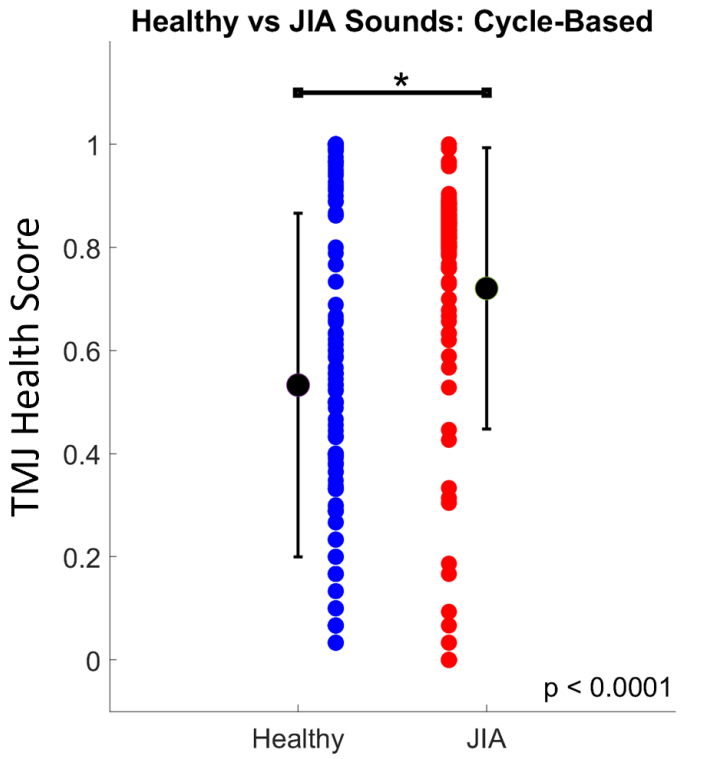

Figure 32. TMJ AEs of JIA Patients with Jaw Sounds vs Healthy Patients with Jaw Sounds.

7.3.4.5 Feature Ranking in JIA vs Healthy TMJ AEs

The 49 features were ranked based on the average information gained at each split of the bagged tree as described in 4.2.4. The top 13 features are presented in Figure 33. The RMS power and Energy were the top ranked time-domain features. Spectral slope, spread and entropy, were the three highest spectral features.

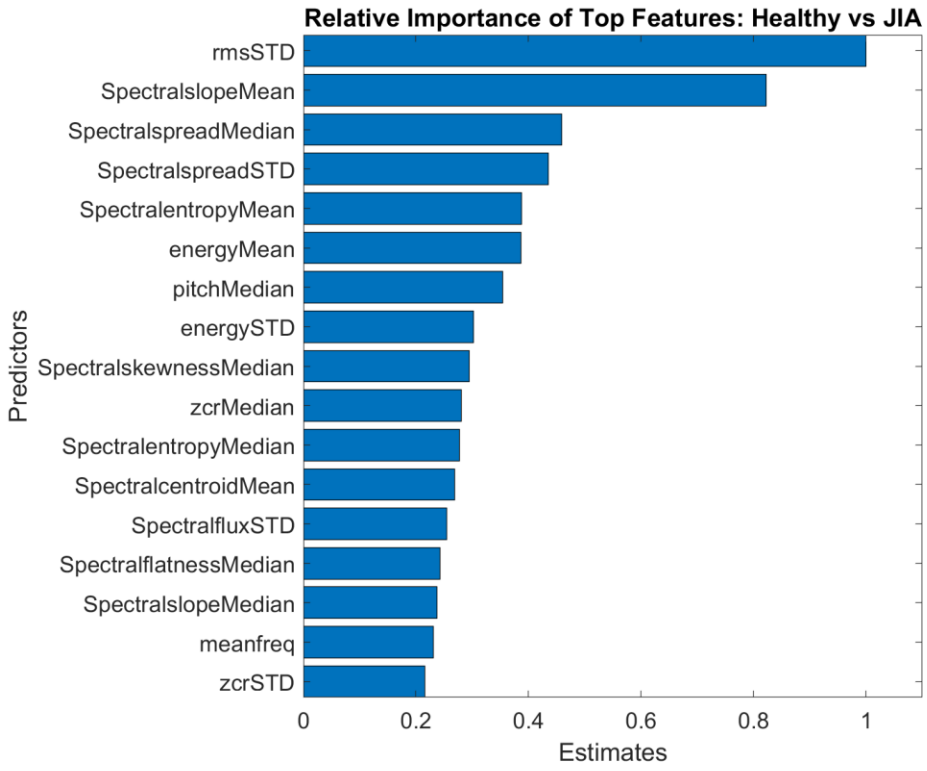

Figure 33. Relative Feature Importance Ranking for Distinguishing JIA from Healthy AEs in the TMJ.

7.3.5 Discussion of JIA TMJ AE Results

This study sought to answer two key questions: 1) can AEs from the TMJ be used to screen for JIA, and 2) is there a difference in pathologic and non-pathologic TMJ sounds. To answer these questions the AEs from 51 subjects. 28 healthy controls (8 with observed TMJ sounds), and 23 subjects with JIA (14 had TMJ sounds) were first recorded using our custom setup described in 7.2.2. Each subject opened and closed their mouths 10 times while their AEs were recorded. These AEs were first qualitatively compared in Figure 30. It was seen that they had much less variability in the AE patterns than was previously seen in the knee AE recordings in a similar population with JIA. However, the patients with JIA

and TMJ sounds still appeared to have a pattern filled with more large vertical spikes or "clicks".

To more quantitatively compare the AE patterns, 49 features (described in Table 1) were computed for each cycle of movement. These features were used to train a machine learning classification algorithm, specifically a bagged decision tree. First, this model was trained on the subjects regardless of whether they had observed TMJ sounds. Using LOSO-CV, it was shown that AEs can reasonably predict JIA in a subject Figure 31. These results were less compelling than the near 100% accuracy we previously found when recording the AEs from the knees of patients with JIA (see Figure 22). However, since JIA is a systemic disease, perhaps AE assessment of both the knees and the TMJ could be used in tandem to further improve the sensitivity of this test. The 70.8% accuracy may not be high enough for being used as the sole basis for a JIA diagnosis. However, in the future, AE assessment may be used in addition to a thorough physical exam and patient history. The combined accuracy of these three techniques should help an OMS decide whether further testing/imaging is necessary to diagnose a patient. In this way, AE assessment may serve as an indicator of the need for further workup.

Assuming a patient is being seen for a jaw complaint, and that they have noticeable jaw sounds on physical exam, the question then becomes are these sounds pathologic and harmful or the result of a benign anatomical variant. This question was addressed in this study by comparing the AEs of patients with and without JIA that did have TMJ sounds. To determine if the AEs from these two groups had characteristics that could be used to separate them, again a bagged decision tree model was trained using the recorded AEs from these two subgroups from our larger database of recorded AEs. This cohort had 17 subjects

(9 of whom had JIA). The output joint health index of the sounds was significantly different between the two groups. This indicates that AE assessment could be used to determine if a patient's TMJ sounds are pathologic or not.

The work presented in this section shows that TMJ AE assessment is a promising if nascent technique. Before clinical adoption is possible, more patients should be recruited to assess the generalizability of the results seen in our work. Other potential causes of TMJ sounds should also be recorded and assessed to determine the full diagnostic capabilities of this technique. In order to classify the patients in this study as having JIA or not, and having TMJ sounds or not, we relied on the physical exam and history results acquired by the treating physician. In the future, we may have better classification accuracy using AEs if the groups are better stratified using gold-standards of diagnosis such as MRI imaging. Finally, the hardware used to record the AEs should ideally be made more portable and user-friendly as well. Currently, it is all powered by a laptop and able to be moved from one room to another, but the software is mostly command-line and prompt based. This primitive user interface is functional but requires a significant amount of training before it could easily be used by a non-expert.

CHAPTER 8. CONCLUSIONS AND FUTURE WORK

8.1 Conclusion

Musculoskeletal injuries and rheumatic conditions are prevalent with nearly 18 million injury-related patient visits occurring each year in the United States. In addition, 1 in 4 people experience some form of arthritis. Few objective biomarkers exist to quantify the health of the joints. AE assessment of the joints would potentially allow for improved detection, management, and tracking of these conditions.

In this thesis, we developed techniques for recording and analyzing AEs from two joints: the knee and the TMJ. We first developed a cadaver model of acute knee injury and quantified the specific characteristics of AEs that change following injury. Those changes were described using a collection of signal features. The feature set from this study formed the basis for our proposed machine-learning classification algorithms that output easily interpretable joint health scores. We then translated this technology into the clinic and recorded pediatric populations with JIA.

In the JIA studies, the joint health score in both the knees and the TMJ were able to differentiate between the AEs of healthy patients and those with the condition. Longitudinal monitoring of the knees' AEs was also capable of quantifying successful treatment. This study demonstrated that AE analysis is a convenient technique for non-invasively quantifying joint health status that could be performed both in and outside the clinic. In summary, the work presented in this thesis presents a significant step forward in

the understanding of the nature and potential applications of joint AE sensing. With further research, this modality could potentially improve the lives of the children suffering from JIA, and all those recuperating from MSK injury.

8.2 Future Work

Various future research directions could stem from this work. One of the most immediate paths is to determine if the alterations to the AEs seen in the cadaver model on acute knee injury is preserved *in vivo*. In fact, recruitment is ongoing in this avenue in our lab. The current number of subjects in that study is insufficient to make any significant claims, but preliminary findings look promising. The recording setup developed in this work uses a combination of accelerometers and IMUs to capture joint AEs. In the future, the location and number of these sensors could be adjusted to optimize signal quality and localize MSK injuries. Additionally, AEs from other joints should be researched to see if they can similarly provide clinically relevant information. In this work, we saw that machine learning models can predict a subject's disease status. However, these models only compared a healthy cohort to a cohort with the pathology of interest. Whether an arthritis-induce AE could be distinguished from an injury-induced AE is yet to be determined. As the dataset continues to grow a comprehensive joint health scoring algorithm could be developed – one that outputs the probability of a patient having arthritis or any number of MSK injuries and pathologies. This algorithm would form the basis of an ideal screening tool for when any patient with an MSK complaint is seen by a primary care healthcare worker.

On a broader scale, there are still several substantial research avenues to explore before AE analysis would be considered for clinical adoption. Using the cadaver model, more exploration should be done on the nature, origin, and possible confounding factors of AE production. Principle among these are the effects of kinematics (e.g. velocity and range of motion) on the AE profile. The effects of joint swelling on AEs was explored briefly, and it was shown that it had no impact on the AEs. However, this was on a small number of cadaver knees with a rather basic method of testing. A more thorough could reveal new information about swelling's effects. It would seem as though with more swelling one would see less interaction of the articulating surface and thus fewer spikes in the AEs. This was not observed in the six legs tested which indicates one of three things: 1) the injected saline volume was inadequate at maintaining separation of the joint perhaps due to loss of tissue integrity, 2) swelling does in fact not impact AE production, or 3) there is an alternate mechanism of AE production occurring. With that final possibility, it raises the point that the mechanism of joint sound production should be further researched and explained. In particular, many of our patients with JIA were not expected to have progressed to the stage of having cartilage erosions and bone spurs. They would however have synovial thickening. If this was the only change to the joint, increased bone-on-bone contact would not occur and be creating the sounds as hypothesized in Chapter 5. The mechanism of AE production in JIA needs to be further researched so that the full capabilities of AEs for monitoring JIA could be assessed.

Another hurdle to clinical adoption that must be addressed is the ease of use of the AE recording software and hardware. Currently, the recording software is still largely command-line based and the sensors do not have any packaging. In future

implementations, the code should be written in a way to maximize user-friendliness – perhaps with a graphical user interface. This would allow for faster training on the software. On the hardware, thought should be placed on providing a guide for appropriately adhering the accelerometers on the ideal locations of the joint being recorded. The design of the packaging, installation technique, and methods for increasing ease of use remain important areas for improvement.

The work presented in this thesis is a substantial step in the understanding of AEs, but there are several limitations that should be addressed in the future. Principle among these is the size of the datasets. The next feasible step in this work is to considerably increase study recruitment, perhaps by disseminating the technology or partnering with other institutions and hospital systems. We will not understand the full diagnostic capabilities of AEs until recordings from a wide range of MSK conditions are available. To begin collecting such a dataset, we should leverage one of the key benefits of AE analysis – its portability. Hardware form-factors should be developed that will allow for more frequent recordings outside of the clinic. With that data, we could better understand the potential for AEs to be used as a longitudinal biomarker of rehabilitation and treatment efficacy. While we do not envision AE assessment replacing imaging in the workup of MSK patients, it could potentially complement the clinical assessment and help provide a more complete understanding of a patient's joints.

REFERENCES

[1] A. D. Woolf and K. Akesson, "Understanding the burden of musculoskeletal conditions. The burden is huge and not reflected in national health priorities.," *BMJ*, vol. 322, no. 7294, pp. 1079–80, May 2001.

[2] American Academy of Orthopaedic Surgeons, "Information about Musculoskeletal Conditions." [Online]. Available: https://www.aaos.org/CustomTemplates/Content.aspx?id=6406&ssopc=1. [Accessed: 07-Sep-2019].

[3] M. Majewski, H. Susanne, and S. Klaus, "Epidemiology of athletic knee injuries: A 10-year study," *Knee*, vol. 13, no. 3, pp. 184–188, Jun. 2006.

[4] J. M. Hootman, C. A. Macera, B. E. Ainsworth, C. L. Addy, M. Martin, and S. N. Blair, "Epidemiology of musculoskeletal injuries among sedentary and physically active adults," *Med. Sci. Sport. Exerc.*, vol. 34, no. 5, pp. 838–844, May 2002.

[5] A. Petersson, "What you can and cannot see in TMJ imaging - an overview related to the RDC/TMD diagnostic system," *J. Oral Rehabil.*, vol. 37, no. 10, pp. 771–778, Oct. 2010.

[6] T. A. Larheim, A. S. Doria, E. Kirkhus, D. A. Parra, C. J. Kellenberger, and L. Z. Arvidsson, "TMJ imaging in JIA patients—An overview," *Semin. Orthod.*, vol. 21, no. 2, pp. 102–110, Jun. 2015.

[7] A. P. Winterstein, T. A. McGuine, K. E. Carr, and S. J. Hetzel, "Comparison of IKDC and SANE Outcome Measures Following Knee Injury in Active Female Patients," *Sport. Heal. A Multidiscip. Approach*, vol. 5, no. 6, pp. 523–529, Nov. 2013.

[8] L. Atallah, G. G. Jones, R. Ali, J. J. H. Leong, B. Lo, and G.-Z. Yang, "Observing Recovery from Knee-Replacement Surgery by Using Wearable Sensors," in *2011 International Conference on Body Sensor Networks*, 2011, pp. 29–34.

[9] G. Cooper *et al.*, "Inertial sensor-based knee flexion/extension angle estimation," *J. Biomech.*, vol. 42, no. 16, pp. 2678–2685, Dec. 2009.

[10] C. Glaros, D. I. Fotiadis, A. Likas, and A. Stafylopatis, "A wearable intelligent system for monitoring health condition and rehabilitation of running athletes," in *4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine, 2003.*, pp. 276–279.

[11] B. J. Munro, T. E. Campbell, G. G. Wallace, and J. R. Steele, "The intelligent knee sleeve: A wearable biofeedback device," *Sensors Actuators B Chem.*, vol. 131, no. 2, pp. 541–547, May 2008.

[12] H. Saito, T. Watanabe, and A. Arifin, "Ankle and Knee Joint Angle Measurements during Gait with Wearable Sensor System for Rehabilitation," Springer, Berlin, Heidelberg, 2009, pp. 506–509.

[13] L. D. Toffola, S. Patel, M. Y. Ozsecen, R. Ramachandran, and P. Bonato, “A wearable system for long-term monitoring of knee kinematics,” in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, 2012, pp. 188–191.

[14] M. L. Chu, I. A. Gradisar, M. R. Railey, and G. F. Bowling, “Detection of knee joint diseases using acoustical pattern recognition technique,” *J. Biomech.*, vol. 9, no. 3, pp. 111–114, Jan. 1976.

[15] C. B. Frank, R. M. Rangayyan, and G. D. Bell, “Analysis of knee joint sound signals for non-invasive diagnosis of cartilage pathology,” *IEEE Eng. Med. Biol. Mag.*, vol. 9, no. 1, pp. 65–68, Mar. 1990.

[16] L.-K. Shark, H. Chen, and J. Goodacre, “Discovering differences in acoustic emission between healthy and osteoarthritic knees using a four-phase model of sit-stand-sit movements.,” *Open Med. Inform. J.*, vol. 4, pp. 116–25, Jul. 2010.

[17] C. N. Teague *et al.*, “Novel Methods for Sensing Acoustical Emissions From the Knee for Wearable Joint Health Assessment,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 8, pp. 1581–1590, Aug. 2016.

[18] W. E. Blodgett, “Auscultation of the Knee Joint,” *Bost. Med. Surg. J.*, vol. 146, no. 3, pp. 63–66, Jan. 1902.

[19] R. A. Mollan, G. C. McCullagh, and R. I. Wilson, “A critical appraisal of auscultation of human joints.,” *Clin. Orthop. Relat. Res.*, no. 170, pp. 231–7, Oct. 1982.

[20] S. Krishnan, R. M. Rangayyan, G. D. Bell, and C. B. Frank, “Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology,” *IEEE Trans. Biomed. Eng.*, vol. 47, no. 6, pp. 773–783, Jun. 2000.

[21] R. M. Rangayyan, S. Krishnan, G. D. Bell, C. B. Frank, and K. O. Ladly, “Parametric representation and screening of knee joint vibroarthrographic signals,” *IEEE Trans. Biomed. Eng.*, vol. 44, no. 11, pp. 1068–1074, 1997.

[22] M. Leslie, “Knee osteoarthritis management therapies.,” *Pain Manag. Nurs.*, vol. 1, no. 2, pp. 51–7, Jun. 2000.

[23] L. S. Lohmander, A. Östenberg, M. Englund, and H. Roos, “High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury,” *Arthritis Rheum.*, vol. 50, no. 10, pp. 3145–3152, 2004.

[24] O. T. Inan *et al.*, “Wearable knee health system employing novel physiological biomarkers,” *J. Appl. Physiol.*, vol. 124, no. 3, pp. 537–547, Mar. 2017.

[25] E. Cannizzaro, S. Schroeder, L. M. Müller, C. J. Kellenberger, R. K. Saurenmann, and J. Goldenberg, “Temporomandibular Joint Involvement in Children with Juvenile Idiopathic Arthritis,” *J. Rheumatol.*, vol. 38, no. 3, pp. 510–515, Mar. 2011.

[26] S. Abramowicz, S. Kim, S. Prahalad, A. F. Chouinard, and L. B. Kaban, “Juvenile arthritis: current concepts in terminology, etiopathogenesis, diagnosis, and

management," *Int. J. Oral Maxillofac. Surg.*, vol. 45, no. 7, pp. 801–812, Jul. 2016.

[27] K. L. Moore, A. M. R. Agur, A. F. Dalley, K. L. Moore, and others, *Essential Clinical Anatomy*. Wolters Kluwer Health, 2015.

[28] A. D. Woolf and B. Pfleger, "Burden of major musculoskeletal conditions," *Bull. World Health Organ.*, vol. 81, pp. 646–656, 2003.

[29] L. S. Bickley, P. G. Szilagyi, and B. Bates, *Bates' Guide to Physical Examination and History Taking*. Lippincott Williams & Wilkins, 2009.

[30] A. J. Rao, B. J. Erickson, G. L. Cvetanovich, A. B. Yanke, B. R. Bach, and B. J. Cole, "The Meniscus-Deficient Knee: Biomechanics, Evaluation, and Treatment Options.," *Orthop. J. Sport. Med.*, vol. 3, no. 10, p. 2325967115611386, Oct. 2015.

[31] M. D. Wechalekar and M. D. Smith, "Utility of arthroscopic guided synovial biopsy in understanding synovial tissue pathology in health and disease states.," *World J. Orthop.*, vol. 5, no. 5, pp. 566–73, Nov. 2014.

[32] S. Abramowicz, H. K. Susarla, S. Kim, and L. B. Kaban, "Physical Findings Associated With Active Temporomandibular Joint Inflammation in Children With Juvenile Idiopathic Arthritis," *J. Oral Maxillofac. Surg.*, vol. 71, no. 10, pp. 1683–1687, Oct. 2013.

[33] R. E. Petty *et al.*, "International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001.," *J. Rheumatol.*, vol. 31, no. 2, pp. 390–2, Feb. 2004.

[34] S. Prahalad and D. N. Glass, "A comprehensive review of the genetics of juvenile idiopathic arthritis," *Pediatr. Rheumatol.*, vol. 6, no. 1, p. 11, Dec. 2008.

[35] A. Ravelli and A. Martini, "Juvenile idiopathic arthritis," *Lancet*, vol. 369, no. 9563, pp. 767–778, Mar. 2007.

[36] K. Murray, S. D. Thompson, and D. N. Glass, "Pathogenesis of juvenile chronic arthritis: genetic and environmental factors.," *Arch. Dis. Child.*, vol. 77, no. 6, pp. 530–4, Dec. 1997.

[37] S. L. Masters, A. Simon, I. Aksentijevich, and D. L. Kastner, "Horror Autoinflammaticus : The Molecular Pathophysiology of Autoinflammatory Disease," *Annu. Rev. Immunol.*, vol. 27, no. 1, pp. 621–668, Apr. 2009.

[38] C. A. Wallace, "Current management of juvenile idiopathic arthritis," *Best Pract. Res. Clin. Rheumatol.*, vol. 20, no. 2, pp. 279–300, Apr. 2006.

[39] P. Woo, "Systemic juvenile idiopathic arthritis: diagnosis, management, and outcome," *Nat. Clin. Pract. Rheumatol.*, vol. 2, no. 1, pp. 28–34, Jan. 2006.

[40] F. H. M. Prince, M. H. Otten, and L. W. A. van Suijlekom-Smit, "Diagnosis and management of juvenile idiopathic arthritis.," *BMJ*, vol. 341, p. c6434, Dec. 2010.

[41] CLeveland Clinic, "Juvenile Idiopathic Arthritis | Cleveland Clinic," 2017. [Online]. Available: https://my.clevelandclinic.org/health/articles/juvenile-idiopathic-arthritis. [Accessed: 07-Sep-2019].

[42] J. M. E. C. for C. D. and C. John M. Eisenberg Center for Clinical Decisions and

Communications Science, *DMARDs for Juvenile Idiopathic Arthritis: A Review of the Research for Parents and Caregivers*. Agency for Healthcare Research and Quality (US), 2005.

[43] G. Gartlehner, R. A. Hansen, B. L. Jonas, P. Thieda, and K. N. Lohr, "Biologics for the treatment of juvenile idiopathic arthritis: a systematic review and critical analysis of the evidence," *Clin. Rheumatol.*, vol. 27, no. 1, pp. 67–76, Nov. 2007.

[44] B. P. Boden, F. T. Sheehan, J. S. Torg, and T. E. Hewett, "Noncontact anterior cruciate ligament injuries: mechanisms and risk factors.," *J. Am. Acad. Orthop. Surg.*, vol. 18, no. 9, pp. 520–7, Sep. 2010.

[45] K. M. Sutton and J. M. Bullock, "Anterior cruciate ligament rupture: differences between males and females.," *J. Am. Acad. Orthop. Surg.*, vol. 21, no. 1, pp. 41–50, Jan. 2013.

[46] W. M. Wind, J. A. Bergfeld, and R. D. Parker, "Evaluation and treatment of posterior cruciate ligament injuries: revisited.," *Am. J. Sports Med.*, vol. 32, no. 7, pp. 1765–75, Oct. 2004.

[47] M. S. Schulz, K. Russe, A. Weiler, H. J. Eichhorn, and M. J. Strobel, "Epidemiology of posterior cruciate ligament injuries.," *Arch. Orthop. Trauma Surg.*, vol. 123, no. 4, pp. 186–91, May 2003.

[48] K. E. Dehaven, "Diagnosis of acute knee injuries with hemarthrosis," *Am. J. Sports Med.*, vol. 8, no. 1, pp. 9–14, Jan. 1980.

[49] P. Bansal, D. J. Deehan, and R. J. H. Gregory, "Diagnosing the acutely locked knee.," *Injury*, vol. 33, no. 6, pp. 495–8, Jul. 2002.

[50] P. Abdon, A. Lindstrand, and K. G. Thorngren, "Statistical evaluation of the diagnostic criteria for meniscal tears.," *Int. Orthop.*, vol. 14, no. 4, pp. 341–5, 1990.

[51] K. G. Shea and J. L. Carey, "Management of anterior cruciate ligament injuries: evidence-based guideline.," *J. Am. Acad. Orthop. Surg.*, vol. 23, no. 5, pp. e1-5, May 2015.

[52] P. D. Fabricant and M. S. Kocher, "Management of ACL Injuries in Children and Adolescents.," *J. Bone Joint Surg. Am.*, vol. 99, no. 7, pp. 600–612, 2017.

[53] L. M. Kruse, B. Gray, and R. W. Wright, "Rehabilitation after anterior cruciate ligament reconstruction: a systematic review.," *J. Bone Joint Surg. Am.*, vol. 94, no. 19, pp. 1737–48, Oct. 2012.

[54] V. K. Tjong, M. L. Murnaghan, J. M. Nyhof-Young, and D. J. Ogilvie-Harris, "A qualitative investigation of the decision to return to sport after anterior cruciate ligament reconstruction: to play or not to play.," *Am. J. Sports Med.*, vol. 42, no. 2, pp. 336–42, Feb. 2014.

[55] A. Bedi, V. Musahl, and J. B. Cowan, "Management of Posterior Cruciate Ligament Injuries: An Evidence-Based Review.," *J. Am. Acad. Orthop. Surg.*, vol. 24, no. 5, pp. 277–89, May 2016.

[56] C.-J. Wang, H.-S. Chen, T.-W. Huang, and L.-J. Yuan, "Outcome of surgical

reconstruction for posterior cruciate and posterolateral instabilities of the knee.," *Injury*, vol. 33, no. 9, pp. 815–21, Nov. 2002.

[57] M. J. Strobel, A. Weiler, M. S. Schulz, K. Russe, and H. J. Eichhorn, "Arthroscopic evaluation of articular cartilage lesions in posterior-cruciate-ligament-deficient knees.," *Arthroscopy*, vol. 19, no. 3, pp. 262–8, Mar. 2003.

[58] J. R. Giffin, K. J. Stabile, T. Zantop, T. M. Vogrin, S. L.-Y. Woo, and C. D. Harner, "Importance of tibial slope for stability of the posterior cruciate ligament deficient knee.," *Am. J. Sports Med.*, vol. 35, no. 9, pp. 1443–9, Sep. 2007.

[59] J. R. Giffin, T. M. Vogrin, T. Zantop, S. L. Y. Woo, and C. D. Harner, "Effects of increasing tibial slope on the biomechanics of the knee.," *Am. J. Sports Med.*, vol. 32, no. 2, pp. 376–82, Mar. 2004.

[60] K. A. Turman and D. R. Diduch, "Meniscal repair: indications and techniques.," *J. Knee Surg.*, vol. 21, no. 2, pp. 154–62, Apr. 2008.

[61] C. C. Rankin, D. M. Lintner, P. C. Noble, V. Paravic, and E. Greer, "A biomechanical analysis of meniscal repair techniques.," *Am. J. Sports Med.*, vol. 30, no. 4, pp. 492–7, Jul. 2002.

[62] C. E. Henning, M. A. Lynch, and J. R. Clark, "Vascularity for healing of meniscus repairs.," *Arthroscopy*, vol. 3, no. 1, pp. 13–8, 1987.

[63] A. Shieh, T. Bastrom, J. Roocroft, E. W. Edmonds, and A. T. Pennock, "Meniscus tear patterns in relation to skeletal immaturity: children versus adolescents.," *Am. J. Sports Med.*, vol. 41, no. 12, pp. 2779–83, Dec. 2013.

[64] C. Palazzo, J.-F. Ravaud, A. Papelard, P. Ravaud, and S. Poiraudeau, "The Burden of Musculoskeletal Conditions," *PLoS One*, vol. 9, no. 3, p. e90633, Mar. 2014.

[65] H. Yoshida, T. Sano, R. Kataoka, K. Takahashi, and Ken-ichi Michi, "A Preliminary Investigation of a Method of Detecting Temporomandibular Joint Sounds.," *J. Orofac. Pain*, vol. 8, no. 1, pp. 73–79, 1994.

[66] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust.*, vol. 27, no. 2, pp. 113–120, Apr. 1979.

[67] R. GutenbergB, "SeismicityoftheEarthandAssociatedPhenomena," *NewYork London: HafnerPub*, 1954.

[68] H. K. Jeong, D. Whittingslow, and O. T. Inan, "b-Value: A Potential Biomarker for Assessing Knee-Joint Health Using Acoustical Emission Sensing," *IEEE Sensors Lett.*, vol. 2, no. 4, pp. 1–4, Dec. 2018.

[69] R. E. Ricklefs, "A Graphical Method of Fitting Equations to Growth Curves," *Ecology*, vol. 48, no. 6, pp. 978–983, Nov. 1967.

[70] J.-P. Gabriel, F. Saucy, and L.-F. Bersier, "Paradoxes in the logistic equation?," *Ecol. Modell.*, vol. 185, no. 1, pp. 147–151, Jun. 2005.

[71] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *J. Biomed. Inform.*, vol. 35, no. 5–6, pp. 352–359, Oct. 2002.

[72] C. G. Broyden, “Quasi-Newton Methods and their Application to Function Minimisation,” *Math. Comput.*, vol. 21, no. 99, p. 368, Jul. 1967.

[73] M. Jansche and Martin, “Maximum expected F-measure training of logistic regression models,” in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT ’05*, 2005, pp. 692–699.

[74] S. R. Safavian and D. Landgrebe, “A survey of decision tree classifier methodology,” *IEEE Trans. Syst. Man. Cybern.*, vol. 21, no. 3, pp. 660–674, 1991.

[75] L. Breiman, “Bagging predictors (Technical Report 421),” *Univ. California, Berkeley*, 1994.

[76] T. Giannakopoulos and A. Pikrakis, *Introduction to audio analysis : a MATLAB approach*. .

[77] G. Peeters, “A large set of audio features for sound description,” 2004.

[78] S. Hersek *et al.*, “Wearable Vector Electrical Bioimpedance System to Assess Knee Joint Health.,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 10, pp. 2353–2360, 2017.

[79] J. M. Hootman, C. G. Helmick, K. E. Barbour, K. A. Theis, and M. A. Boring, “Updated Projected Prevalence of Self-Reported Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation Among US Adults, 2015-2040,” *Arthritis Rheumatol.*, vol. 68, no. 7, pp. 1582–1587, Jul. 2016.

[80] T.-F. Lee, W.-C. Lin, L.-F. Wu, and H.-Y. Wang, “Analysis of Vibroarthrographic Signals for Knee Osteoarthritis Diagnosis,” in *2012 Sixth International Conference on Genetic and Evolutionary Computing*, 2012, pp. 223–228.

[81] M. P. Sinan Hersek Caitlin Teague, Michael Sawka, Mindy Stafford, Geza Kogler, Paul Wolkoff, Omer Inan *et al.*, “Acoustical Emission Analysis by Unsupervised Graph Mining: A Novel Biomarker of Knee Health Status,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 6, pp. 1291–1300, Jun. 2018.

[82] M. J. Medvecky and F. R. Noyes, “Surgical Approaches to the Posteromedial and Posterolateral Aspects of the Knee,” *JAAOS - J. Am. Acad. Orthop. Surg.*, vol. 13, no. 2, pp. 121–128, 2005.

[83] W. R. Post, S. R. Akers, and V. Kish, “Load to failure of common meniscal repair techniques: effects of suture technique and suture material.,” *Arthroscopy*, vol. 13, no. 6, pp. 731–6, Dec. 1997.

[84] T. Fukubayashi and H. Kurosawa, “The Contact Area and Pressure Distribution Pattern of the Knee: A Study of Normal and Osteoarthrotic Knee Joints,” *Acta Orthop. Scand.*, vol. 51, no. 1–6, pp. 871–879, Jan. 1980.

[85] H. R. Schumacher, “Aspiration and injection therapies for joints,” *Arthritis Rheum.*, vol. 49, no. 3, pp. 413–420, Jun. 2003.

[86] E. L. Brainerd *et al.*, “X-ray reconstruction of moving morphology (XROMM): precision, accuracy and applications in comparative biomechanics research,” *J. Exp. Zool. Part A Ecol. Genet. Physiol.*, vol. 9999A, no. 5, p. n/a-n/a, Jun. 2010.

[87] B. J. Knörlein, D. B. Baier, S. M. Gatesy, J. D. Laurence-Chasen, and E. L. Brainerd, "Validation of XMALab software for marker-based XROMM," *J. Exp. Biol.*, vol. 219, no. 23, pp. 3701–3711, 2016.

[88] A. M. Ahmed and D. L. Burke, "In-Vitro of Measurement of Static Pressure Distribution in Synovial Joints—Part I: Tibial Surface of the Knee," *J. Biomech. Eng.*, vol. 105, no. 3, pp. 216–225, Aug. 1983.

[89] T. J. Fairbank, "Knee Joint Changes after Menisectomy," *J. Bone Joint Surg. Br.*, vol. 30-B, no. 4, pp. 664–670, Nov. 1948.

[90] A. Matsas, N. Taylor, and H. McBurney, "Knee joint kinematics from familiarised treadmill walking can be generalised to overground walking in young unimpaired subjects.," *Gait Posture*, vol. 11, no. 1, pp. 46–53, Feb. 2000.

[91] D. Périé and M. C. Hobatho, "In vivo determination of contact areas and pressure of the femorotibial joint using non-linear finite element analysis.," *Clin. Biomech. (Bristol, Avon)*, vol. 13, no. 6, pp. 394–402, Sep. 1998.

[92] I. D. McDermott and A. A. Amis, "The consequences of meniscectomy," *J. Bone Joint Surg. Br.*, vol. 88-B, no. 12, pp. 1549–1556, Dec. 2006.

[93] M. Englund, E. M. Roos, and L. S. Lohmander, "Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls.," *Arthritis Rheum.*, vol. 48, no. 8, pp. 2178–87, Aug. 2003.

[94] S. Henry, R. Mascarenhas, D. Kowalchuk, B. Forsythe, J. J. Irrgang, and C. D. Harner, "Medial Meniscus Tear Morphology and Chondral Degeneration of the Knee: Is There a Relationship?," *Arthrosc. J. Arthrosc. Relat. Surg.*, vol. 28, no. 8, pp. 1124-1134.e2, Aug. 2012.

[95] B. Bai, H. Shun, Z. X. Yin, Z. Liao, and N. Chen, "Changes of contact pressure and area in patellofemoral joint after different meniscectomies," *Int. Orthop.*, vol. 36, no. 5, pp. 987–991, May 2012.

[96] J. F. Prinz and K. W. Ng, "Characterization of Sounds Emanating from the Human Temporomandibular Joint," *Arch. Oral Biol.*, vol. 41, no. 7, pp. 631–639, 1996.

[97] M. R. Torry, M. J. Decker, R. W. Viola, D. D. O'Connor, and J. R. Steadman, "Intra-articular knee joint effusion induces quadriceps avoidance gait patterns.," *Clin. Biomech. (Bristol, Avon)*, vol. 15, no. 3, pp. 147–59, Mar. 2000.

[98] D. C. Whittingslow, H.-K. Jeong, V. G. Ganti, N. J. Kirkpatrick, G. F. Kogler, and O. T. Inan, "Acoustic Emissions as a Non-invasive Biomarker of the Structural Health of the Knee," *Ann. Biomed. Eng.*, pp. 1–11, Jul. 2019.

[99] R. Hogg, J. McKean, and A. Craig, "Introduction to Mathematical Statistics (6th Edition)," *All Books Monogr. by WMU Authors*, Jan. 2005.

[100] R. Hemke *et al.*, "Frequency of joint involvement in juvenile idiopathic arthritis during a 5-year follow-up of newly diagnosed patients: implications for MR imaging as outcome measure," *Rheumatol. Int.*, vol. 35, no. 2, pp. 351–357, Feb. 2015.

[101] R. Hemke, N. Tzaribachev, A. M. Barendregt, J. Merlijn van den Berg, A. S. Doria, and M. Maas, “Imaging of the knee in juvenile idiopathic arthritis.,” *Pediatr. Radiol.*, vol. 48, no. 6, pp. 818–827, 2018.

[102] C. L. Deal *et al.*, “The United States rheumatology workforce: Supply and demand, 2005–2025,” *Arthritis Rheum.*, vol. 56, no. 3, pp. 722–729, Mar. 2007.

[103] Arthritis Foundation, “Increase Access to Pediatric Rheumatologists.” [Online]. Available: https://www.arthritis.org/advocate/our-policy-priorities/access-to-care/increase-access-to-pediatric-rheumatologists/. [Accessed: 12-Sep-2019].

[104] B. Semiz, S. Hersek, D. C. Whittingslow, L. A. Ponder, S. Prahalad, and O. T. Inan, “Using Knee Acoustical Emissions for Sensing Joint Health in Patients with Juvenile Idiopathic Arthritis: A Pilot Study,” *IEEE Sens. J.*, vol. 18, no. 22, 2018.

[105] D. A. Winter, *Biomechanics and motor control of human movement*. Wiley, 2009.

[106] L.-S. Ording Muller, P. Humphries, and K. Rosendahl, “The Joints in Juvenile Idiopathic Arthritis.,” *Insights Imaging*, vol. 6, no. 3, pp. 275–84, Jun. 2015.

[107] R. K. Saurenmann *et al.*, “Prevalence, risk factors, and outcome of uveitis in juvenile idiopathic arthritis: A long-term followup study,” *Arthritis Rheum.*, vol. 56, no. 2, pp. 647–657, Feb. 2007.

[108] R. K. Saurenmann *et al.*, “Epidemiology of juvenile idiopathic arthritis in a multiethnic cohort: Ethnicity as a risk factor,” *Arthritis Rheum.*, vol. 56, no. 6, pp. 1974–1984, Jun. 2007.

[109] M. L. Chu, I. A. Gradisar, M. R. Railey, and G. F. Bowling, “An electro-acoustical technique for the detection of knee joint noise.,” *Med. Res. Eng.*, vol. 12, no. 1, pp. 18–20, 1976.

[110] H.-K. Jeong, M. B. Pouyan, D. C. Whittingslow, V. Ganti, and O. T. Inan, “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, 2018.

[111] A. Ravelli and others, “Toward an understanding of the long-term outcome of juvenile idiopathic arthritis,” *Clin. Exp. Rheumatol.*, vol. 22, no. 3, pp. 271–275, 2004.

[112] Y. Wu, S. Krishnan, and R. M. Rangayyan, “Computer-aided diagnosis of knee-joint disorders via vibroarthrographic signal analysis: a review.,” *Crit. Rev. Biomed. Eng.*, vol. 38, no. 2, pp. 201–24, 2010.

[113] A. Steindler, “Auscultation of Joints,” *J. Bone Jt. Surg.*, vol. 19, no. 1, pp. 121–136, 1937.

[114] R. Brackin and S. Road, “Process and Apparatus for Analyzing Joint Disorders,” US3181528, 1961.

[115] R. Mollan, “Orthopedic Diagnostic Procedures and Apparatus Therefor,” US4,437,437, 1984.

[116] D. J. S. Thomas Gay, Charles N. Bertolami, “Method and Apparatus for the Acoustic

Detection and Analysis of Joint Disorders," US4,836,218, 1989.

[117] P. Łyżwa, M. Kłaczyński, and P. Kazana, "Vibroacoustic Methods of Imaging in Selected Temporomandibular Joint Disorders During Movement," *Diagnostyka*, vol. 19, no. 3, pp. 109–117, 2018.

[118] S. E. Widmalm, W. J. Williams, and C. Zheng, "Time frequency distributions of TMJ sounds," 1991.

[119] S. E. Widmalm, W. J. Williams, R. L. Christiansen, S. M. Gunn, and D. K. Park, "Classification of temporomandibular joint sounds based upon their reduced interference distribution," 1996.

[120] S. E. Widmalm, Y. Dong, B. X. Li, M. Lin, L. J. Fan, and S. M. Deng, "Unbalanced lateral mandibular deviation associated with TMJ sound as a sign in TMJ disc dysfunction diagnosis," *J. Oral Rehabil.*, vol. 43, no. 12, pp. 911–920, 2016.

[121] A. Akan and R. Başar Ünsal, "Time–frequency analysis and classification of temporomandibular joint sounds," *J. Franklin Inst.*, vol. 337, no. 4, pp. 437–451, 2000.

[122] M. de Onis, A. W. Onyango, J. Van den Broeck, W. C. Chumlea, and R. Martorell, "Measurement and Standardization Protocols for Anthropometry Used in the Construction of a New International Growth Reference," *Food Nutr. Bull.*, vol. 25, no. 1_suppl_1, pp. S27–S36, Jan. 2004.

[123] L. G. Farkas, J. C. Posnick, and T. M. Hreczko, "Anthropometric Growth Study of the Head," *Cleft Palate-Craniofacial J.*, vol. 29, no. 4, pp. 303–308, Jul. 1992.

[124] World Health Organization, *WHO Child Growth Standards - Methods and development*. 2007.

[125] T. K. Koo and M. Y. Li, "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research.," *J. Chiropr. Med.*, vol. 15, no. 2, pp. 155–63, Jun. 2016.

[126] G. Oster and J. S. Jaffe, "Low Frequency Sounds from Sustained Contraction of Human Skeletal Muscle," 1980.

[127] T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "Comparing Boosting and Bagging Techniques With Noisy and Imbalanced Data," *IEEE Trans. Syst. Man, Cybern. - Part A Syst. Humans*, vol. 41, no. 3, pp. 552–568, May 2011.

[12 8] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.