

**DIGITAL BIOMARKER DISCOVERY FOR NON-INVASIVE
HEALTH MONITORING WITH ACOUSTIC AND VIBRATION
SIGNALS**

A Dissertation
Presented to
The Academic Faculty

by

Beren Semiz

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2020

COPYRIGHT © 2020 BY BEREN SEMIZ

DIGITAL BIOMARKER DISCOVERY FOR NON-INVASIVE HEALTH MONITORING WITH ACOUSTIC AND VIBRATION SIGNALS

Approved by:

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

Dr. David V. Anderson
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Dr. Wilbur A. Lam
Wallace Coulter Department of Biomedical
Engineering
Georgia Institute of Technology

Dr. Mozziyar Etemadi
Feinberg School of Medicine
Northwestern University

Dr. A. Fatih Sarioglu
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Date Approved: Nov.16, 2020

To my parents, Serap and Ender Semiz

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Prof. Omer T. Inan, for his support and guidance throughout my Ph.D. journey. When I moved to Atlanta, I was looking for an advisor and asking people around me about their experiences and recommendations. After discussions with several people, Prof. Inan's research stood out to me and I reached out to him to discuss his research projects and open positions. From the first moment I met him, I was sure that I want to join Prof. Inan's group. Not only his research projects were so exciting, but also his energy, positivity and sincerity were really assuring. Joining Inan Research Lab was one of the best decisions I have made in my life, and I learned a lot from Prof. Inan's expertise, experiences and creativeness. Moreover, he was a great mentor – he taught everything I need to know about pursuing an academic career and he was always available to help me whenever I need. This work would not have been possible without his guidance and supervision, and words cannot describe how grateful I am to work with him during these four years. As I always say, he is the best advisor one could think of, and I will be forever thankful for his support and mentorship.

In addition, I would like to thank Dr. A. Fatih Sarioglu, Dr. Wilbur A. Lam, Dr. David V. Anderson and Dr. Mozziyar Etemadi for taking time to serve on my dissertation committee and for their useful comments regarding my work. Their input and suggestions have not only improved my current work; but have also widened my vision about the potential directions the current work could take in the future. I am really lucky to have these distinguished researchers in my committee and to get the chance to learn from their thoughts and expertise.

I am grateful for the current and former Inan Research Lab members, Dr. Sinan Hersek, Dr. Mohsen Safaei, Dr. Nordine Sebkhi, Dr. Maziyar Pouyan, Dr. Hazar Ashouri, Dr. Andrew Carek, Dr. Daniel Whittingslow, Dr. Caitlin Teague, Dr. Nick Bolus, Dr. Jon Zia, Dr. Nil Gurel, Oludotun Ode, Mobashir Shandhi, Hyeon Ki Jeong, Hewon Jung, Samer Mabrouk, Venu Ganti, Sevda Gharehbaghi, Jacob Kimball, Brandi Nevius, Kristy Scott, Asim Gazi, Goktug Cihan Ozmen, Daniel Hochman, Arpan Bhavsar, and David Lin. I have been so lucky to work with these smart people, and I wish nothing but the best for them for realizing their dreams and future career goals. Specifically, I would like to thank Dr. Sinan Hersek for introducing me to Prof. Inan and for teaching me everything I need to know during my first years. Throughout my Ph.D., I have worked closely with many people, but specifically I would like to thank Mobashir Shandhi, Dr. Daniel Whittingslow and Dr. Andrew Carek for their help and collaboration in many different projects. Finally, I am grateful to Goktug Cihan Ozmen and Dr. Nil Gurel for their friendship and support, for our endless lunch and coffee chats, but most importantly for being there for me whenever I need.

Next, I would like to thank our collaborators from Emory University (Dr. Sampath Prahalad and Lori Ponder), University of California San Francisco (Dr. Liviu Klein, Dr. Leticia Blazquez-Arroyo, Dr. Van Selby, Dr. Georg Wieselthaler and Cynthia Partida), and Northwestern University (Dr. Mozziyar Etemadi, Dr. Shireen Ahmad, J. Alex Heller, Florencia G. Vicente and Jessica C. Johnson) for their help and collaboration in the projects we worked together. They have always been very approachable and thoughtful, and I feel so lucky to work with these brilliant physicians and scientists. This thesis has mainly built on these collaborative projects and I owe everything I have learned about human subject

studies and interdisciplinary research to these collaborations. I want to thank them again for believing in my research and for sharing their knowledge and expertise with me whenever I need.

I was so lucky to have a great group of people, who have become my family during these four years. While many of my friends pursuing graduate studies in other colleges/cities were complaining about how lonely Ph.D. life is, this was definitely not true for me. I would like to thank Baki Berkay Yilmaz, Berkay Yucel, Bige Deniz Unluturk, Duygu Deniz Umutlu, Fatma Karagoz, Goktug Cihan Ozmen, Gozde Tutuncuoglu, Guliz Ersoy Ozmen, Hakki Mert Torun, Murat Yildirim, Nil Gurel, Selin Cakmak, Sezen Yucel, Su Guvenir and many others who have made my days in Atlanta enjoyable. Thank you all for your friendship, support and love; and most importantly for being home away from home.

There is one special person who has been with me since the first day of my Ph.D. journey. I would like to thank my fiancé, M. Emre Gursoy, for his endless love and support. He is my partner-in-crime and my best friend, and words cannot describe how grateful I am to have him in my life. He stood by me against every challenge I had and always encouraged me to reach my highest potential. He is the best thing Atlanta has brought to my life and I look forward to the life we will spend together.

Last but not least, I would like to thank my parents, Serap and Ender Semiz, for their unconditional love and support. They have dedicated their life to me, and I am dedicating this thesis to them. Spending four years away from them was the hardest thing in my Ph.D. life, and I cannot imagine how they felt as it is always harder for those left

behind. I am grateful for their sacrifice and patience, and this dissertation would not have been possible without their unconditional love, support, and encouragement. They have made me the person I am today, and I work every day to make them proud.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	III
LIST OF TABLES.....	XI
LIST OF FIGURES.....	XII
SUMMARY	XV
CHAPTER 1: INTRODUCTION	1
1.1 Digital Biomarkers for Continuous Health Monitoring	1
1.2 Major Contributions	2
1.3 Scope and Organization of the Dissertation.....	3
CHAPTER 2: ACOUSTIC AND VIBRATION SIGNALS.....	5
2.1 Instrumentation for Acoustic and Vibration Measurements from the Body	5
2.1.1 Contact Microphones.....	5
2.1.2 Digital Stethoscopes	6
2.1.3 Accelerometers	8
2.2 Advantages and Potential Challenges of Wearable Acoustic and Vibration Measurements...	9
CHAPTER 3: WEARABLE ACOUSTIC MEASUREMENT FOR BIOMECHANICS	11

3.1.1	Using Knee Acoustical Emissions for Sensing Joint Health in Patients with Juvenile Idiopathic Arthritis.....	11
3.1.1	Introduction	11
3.1.2	Methods	14
3.1.2.1	Human Subject Protocol and Subject Demographics	14
3.1.2.2	Signal Processing and Feature Extraction.....	16
3.1.2.3	t-Distributed Stochastic Neighbor Embedding (t-SNE).....	18
3.1.2.4	Knee Audio Score Calculation.....	19
3.1.2.5	Feature Importance Ranking.....	22
3.1.2.6	Effect of Model Type.....	23
3.1.3	Results and Discussion	23
3.1.3.1	t-SNE Visualization of the Knee Acoustical Emissions	23
3.1.3.2	Knee Audio Score Assessment	25
3.1.3.3	Feature Importance Ranking	27
3.1.3.4	Effect of Model Type.....	28
3.1.4	Conclusion	29
3.2	Distinguishing between “Physiologic” and “Pathologic” Clicks using the Teager Energy Operator	31
3.2.1	Introduction	31
3.2.2	Methods	32
3.2.2.1	Human Subject Protocol	32
3.2.2.2	Pre-Processing and Acoustic Noise Suppression.....	33
3.2.2.3	Teager Energy Operator.....	34
3.2.2.4	Click Detection	34
3.2.2.5	Feature Extraction.....	35
3.2.2.6	t-Distributed Stochastic Neighbor Embedding (t-SNE).....	36
3.2.2.7	Random Forest Classification	36
3.2.3	Results and Discussion	37

3.2.3.1	t-distributed Stochastic Neighbor Embedding (t-SNE)	37
3.2.3.2	Random Forest Classification	38
3.2.4	Conclusion	39

CHAPTER 4: WEARABLE ACOUSTIC AND VIBRATION MEASUREMENT FOR CARDIOVASCULAR ASSESSMENT.....40

4.1	Detecting Suspected Pump Thrombosis in Left Ventricular Assist Devices via Acoustic Analysis	40
4.1.1	Introduction	40
4.1.2	Methods	43
4.1.2.1	Study Protocol and Subject Demographics.....	43
4.1.2.2	Study Design and Data Acquisition	45
4.1.2.3	Pre-processing and Feature Extraction	47
4.1.2.4	Feature Selection using Spearman Correlation Analysis	48
4.1.2.5	Threshold Selection and Thrombosis Score Calculation	51
4.1.2.6	Performance Comparison with Existing State-of-the-Art Methods.....	53
4.1.3	Results and Discussion	53
4.1.3.1	Feature Selection using Spearman Correlation Analysis	54
4.1.3.2	Threshold Selection and Thrombosis Score Calculation	55
4.1.3.3	Performance Comparison with Existing State-of-the-Art Methods.....	59
4.1.3.4	Limitations and Future Work.....	60
4.1.4	Conclusion	61
4.2	Non-Invasive Wearable Patch Utilizing Seismocardiography for Peri-Operative Use in Surgical Patients	63
4.2.1	Introduction	63
4.2.2	Study Design and Subject Demographics.....	65
4.2.2.1	Transesophageal Doppler (TED)	66

4.2.2.2	Wearable Patch	66
4.2.2.3	Pre-processing and Signal Segmentation.....	69
4.2.2.4	Feature Extraction	72
4.2.2.5	Regression Model and Validation.....	74
4.2.2.6	Elucidating Mechanistic Insight Regarding Stroke Volume Correlations to SCG.....	76
4.2.2.7	Comparison with State-of-the-Art	77
4.2.3	Results and Discussion	79
4.2.3.1	Machine Learning Based Estimation of Stroke Volume	79
4.2.3.2	Interpretation of Acceleration Directions and Features	80
4.2.3.3	Comparison with State-of-the-Art	86
4.2.3.4	Limitations and Future Work.....	88
4.2.4	Conclusion	90
CHAPTER 5: CONCLUSION AND FUTURE WORK.....		91
5.1	Future Work	92
5.1.1	Ensuring Reliability of the Measurements	92
5.1.2	Validating the Proposed Frameworks.....	93
5.1.3	Achieving Biomarker Phenotyping	94
5.2	Final Remarks.....	95
REFERENCES		96

LIST OF TABLES

Table 1: Subject Demographics, Pump Parameters and Blood Biomarkers	45
Table 2: Pump Parameters and Blood Biomarkers for Baseline and Thrombosis Recordings	46
Table 3: Number of Recordings for Each Subject.....	47
Table 4: Feature Descriptions of Different Feature Groups	49
Table 5: Performance Comparison with State-of-the-Art Methods	59
Table 6: Features Extracted from the Signals.....	72
Table 7: Experiments with Axis-Combinations	83
Table 8: Random Forest Feature Importance Ranking.....	84
Table 9: Contribution of the Acoustic and Vibration Components (SCG _{x,y,z,total} and ECG)	85
Table 10: Comparison with the Existing Literature	89

LIST OF FIGURES

Figure 1 : Acoustic and vibration analysis can potentially be leveraged in the design of wearable modalities for deriving digital biomarkers to assist in clinical decisions. The signals in the audible range (20Hz-20kHz) are referred as acoustic signals, whereas vibration signals fall under the infrasound group, which has a frequency content below 20 Hz.	7
Figure 2 : Axes (planes) of the body	8
Figure 3: An overview of the methods used to acquire acoustical emissions from the knee and the algorithms used to analyze the signals. (a) The experiment setup. (b) The signal analysis workflow. (c) Knee audio score calculation.	16
Figure 4: (a) The t-SNE visualizations of knee acoustical emission data from four control subjects and four subjects with JIA using all features. Representative signals and spectrograms from a subject with JIA (b), and two control subjects (c) and (d).....	25
Figure 5: (a) The right, left and average knee audio scores of the subjects with JIA and control subjects. (b) Knee audio scores for the subjects with JIA before and after treatment.	27
Figure 6: (a) Feature importance ranking showing the top fifteen features in terms of salient information provided for classification. (b)-(g) Distributions of the top six features for the control subjects and pre- and post-treatment subjects with JIA. The distributions were calculated using kernel density estimation.	28
Figure 7: (a) Time domain signal and corresponding spectrogram of a recording taken from a control subject with physiological clicks. (b) Time domain signal and corresponding spectrogram of a recording taken from a subject with JIA.....	32
Figure 8: Detection and classification pipeline. Knee acoustic emissions are filtered, and spectral noise suppression is applied. Using the Teager energy operator, change points (in this case "clicks") are detected. Using these click locations, 100ms long click segments are extracted from the signals. Clicks are windowed, and time and frequency domain features are extracted from each window. Using the random forest classifier and LOSO-CV, the clicks are categorized as "physiologic" or "pathologic".	35
Figure 9: (a) The signals are filtered using a 250Hz-20kHz band-pass filter. (b) Spectral noise suppression is applied for background noise. (c) Using the Teager energy operator, change point locations are detected. Using this information, 100ms long click segments are clipped from the band-pass filtered signals.	36
Figure 10: (a) t-SNE graph for the physiologic and pathologic clicks. (b) The ROC curve for the proposed model. The AUC is calculated to be 0.98.	39

Figure 11: System Overview. The digital stethoscope was placed on the patient's exposed chest over the mitral valve and recordings were obtained in cases of changing pump speed or clinical condition. The acquired recordings were windowed, and 60 acoustic features were determined using LDH-based correlation analysis. It was hypothesized that the features which have higher correlation with LDH values would perform better in thrombosis detection. After calculating the correlation coefficients for each LDH-feature pair, the selected features were fed to a logistic regression classifier which was validated using LOSO-CV. Each window was assigned an estimated class and the overall thrombosis score for any given recording was defined as the mean of these estimated values.44

Figure 12: Results. (a) The 15 features having the highest absolute correlation with LDH are presented. There is a great diversity in the feature types, thus each feature group has a relationship with the LDH values. (b) KDE plot for the most correlated feature, pump flow. (c) KDE plot for the most correlated acoustic feature, bandpower 16 (190 - 230 Hz band). (d) KDE plot for the least correlated feature, bandpower 21 (440 - 520 Hz band). (e) ROC curve when the classifier was trained using the top 15 features. The area under the curve (AUC) was calculated to be 0.92. The decision threshold was determined to be 0.15 as it is the optimum value for maximizing the TPR (~ 0.91) while minimizing the FPR (~ 0.13).57

Figure 13: The constructed model using the baseline and thrombosis recordings was fit into the 14 post-thrombolysis recordings taken from 4 subjects to investigate the pump conditions after treatment. The results revealed that, while pump power and hemolysis biomarkers may indicate that the thrombosis is resolved, there may still be other problems such as altered pump mechanics, increased friction, or even residual pump thrombosis that cannot be identified by these markers. Indeed, recurrent thrombosis episodes were observed in 3 of the 4 subjects following the treatment of thrombosis, and the 4 recordings which had scores above the threshold belonged to these 3 subjects.58

Figure 14: (a) Custom-made patch hardware. (b) A representative beat from the ECG, $SCG_{x,y,z,total}$ and $PCG_{sternum}$ signals (for PCG, only Z-axis is shown as the other axes are less commonly used). The portion within each R-R interval is called a *beat*, which is usually around $\sim 600 - 1000$ ms. The R-peak locations on the ECG signal are used to split the SCG into individual beats. The peaks and valleys on the SCG signals correspond to specific cardiac events.68

Figure 15: The block diagram for the data acquisition and signal processing. The wearable patch was mounted on the mid-sternum while the patient was undergoing a surgery. The ECG and $SCG_{x,y,z}$ signals were recorded throughout the surgery while the SV values were being acquired with the TED. In addition to the $SCG_{x,y,z}$ signals, the total acceleration magnitude, SCG_{total} , was calculated. As the TED was averaging and storing the SV values every 30 seconds, the ECG and $SCG_{x,y,z,total}$ signals were split into 30 seconds-long segments. Within each segment, the R-peak locations were found on the ECG and used to detect the individual beats on the $SCG_{x,y,z,total}$. The ensemble average of the detected beats was calculated, so per segment there was one single ensemble averaged frame for each axis. In total 65 features were extracted for the analysis. Then a random forest regression model was trained to predict the corresponding SV values.69

Figure 16: Cubic spline interpolation was performed on the pre-incision SV and HR vectors to increase the number of instances. From the interpolated SV and HR curves, equally spaced (3 seconds) data points were obtained to increase the frequency of the SV and HR readings. Similarly, the patch signals were segmented using a 30 seconds long moving window with 90% overlap (step size = 3 seconds).....71

Figure 17: A simplified representation of how LOSO-CV was performed. In each fold, intra-operative data from one subject was left out S_i . The model was trained using the remaining S_n where $n \neq i$ and all B_n . The model was then tested on S_i and the predicted SV values were stored. This step was repeated for each subject.....75

Figure 18: (a) The reference vs. predicted SV values are presented. 65 features were extracted and used to train a random forest regression model. This resulted in a correlation coefficient (r-value) of 0.81 and MedAE of 7.56 mL. (b) The Bland-Altman graph for the reference and predicted SV values. Horizontal line in the middle represents the bias line, whereas the area between the upper and lower horizontal lines represent the 95% confidence interval, (mean \pm 2*standard deviation).82

Figure 19: Pearson correlation coefficients between different feature pairs from $SCG_{x,y,z,total}$ signals.86

SUMMARY

A biomarker is a parameter that can be used to objectively quantify a physiological or pathophysiological process. Biomarkers used in medicine are commonly derived from blood, saliva, urine or other bodily fluids, and in many cases are used to inform medical decisions. There is a new emerging class of biomarkers, *digital biomarkers*, which are measures collected through connected digital tools, generally across multiple hardware and software layers. This work describes the use of wearable acoustic and vibration measurements to derive digital biomarkers, which can be used together with existing medical information to assist in clinical decisions.

Acoustic and vibration signals carry information that is in many cases complementary to electrophysiology or movement, but the signals are not fundamentally well understood. This leads to a limited one-to-one correspondence between signal characteristics and important health parameters. To that end, this work aims to investigate this correspondence through signal processing, data-driven feature discovery and statistical techniques for deriving accurate and clinically relevant digital biomarkers. In addition, acoustic and vibration signals exhibit substantial inter-subject and intra-subject variability, thus their use in classical diagnostic approaches have not been successful in the past. Rather than focusing on adapting these signals as diagnostic tools, this work aims to derive and employ new algorithms to detect and track the relative changes in health, e.g. exacerbation in clinical state and/or response to a specific treatment, for a given subject over time.

The first part of this dissertation discusses how wearable acoustic measurements can be leveraged in biomechanics, specifically in joint health assessment, to derive

clinically useful digital biomarkers. The first work presents the use of knee acoustical emissions captured through miniature sensors to derive a clinically relevant joint health score for the evaluation of juvenile idiopathic arthritis. Then, a novel click detection and classification algorithm leveraging the Teager Energy Operator is presented to detect the clicks in joint sound signals and distinguish between physiologic and pathologic clicks.

The second part of this dissertation studies the wearable acoustic and vibration measurements for cardiovascular assessment in two different applications. The first application involves pump thrombosis detection in left ventricular assist devices based on analyzing the operating sounds of the pump with machine learning algorithms. The second application is the non-invasive estimation of stroke volume based on wearable seismocardiogram and phonocardiogram measurements taken from the sternum.

Overall, this dissertation presents novel frameworks leveraging wearable vibration and acoustic measurements for knee joint and cardiovascular health assessment. Once verified and validated through large studies, such systems can potentially assist in clinical decisions and improve the management of various diseases and injuries outside the physical confines of the clinic.

CHAPTER 1: INTRODUCTION

1.1 Digital Biomarkers for Continuous Health Monitoring

Biomarkers are measurable and quantifiable indicators used in evaluating “normal vs. pathologic processes or biological responses to therapeutic interventions” [1, 2]. They are commonly derived from blood, saliva, urine or other bodily fluids, and in many cases are used to inform medical decisions as they relate to symptoms and surrogate or clinical endpoints [3, 4]. The diagnoses and treatment titration for any disease or injury are mostly achieved with these conventional biomarkers; however, their derivation requires frequent hospital visits and expensive laboratory tests [3]. Thus, there is a compelling need for novel modalities which can digitize the derivation of such indicators, and employ continuous and non-invasive health monitoring outside the physical confines of the clinic.

The concept of *remote health monitoring* has existed for decades. In the early 70s, the World Health Organization (WHO) introduced the term **telematics**, which relies on the fusion of informatics and telecommunication to enable data transfer across different systems [3]. In the 90s, the idea of leveraging the information and communication technologies for disease diagnosis, treatment titration and patient monitoring was adopted. These discussions led to the definition of **e-Health**, where *the Internet* and related technologies are used to deliver health-related information and services to enable patients’ active involvement in treatments [3, 5]. With the ongoing establishment of mobile devices and patient monitoring devices, WHO defined the term **mobile-Health** (i.e. m-Health) around 2010. Today, **digital health** is the term used to describe digitization of the

healthcare industry, including through point-of-need diagnostic tools and smartphone-linked wearable devices which can provide pervasive and unimpeded sensing [3, 6, 7].

Within the umbrella of digital health exists a new class of biomarkers, ***digital biomarkers***, which are measures collected through connected digital tools, generally across multiple hardware and software layers [1]. Such digital biomarkers comprise (i) *behavioral data* — such as consumer, social media and online search data, and (ii) *physiological data* — which is the main focus of this dissertation.

Recent technological advancements have made it possible to embed various sensors into electronic devices which enable unobtrusive and non-invasive collection of digitized physiological signals. These captured signals are converted into actionable metrics through appropriate algorithms to enable continuous assessment of phenotypic indications [8]. For the caregivers and patients, some of these digital biomarkers are easier to interpret as they are the digitized versions of the well-known parameters, such as heart rate. However, health-related outcomes have a great range, varying from explaining disease states to predicting drug or intervention responses [1, 8]. Although most of these digital biomarkers are still being validated, they may, in the future, provide exhaustive information about the clinical status of the patients to inform diagnostics and dose treatments [9]. This dissertation describes the use of wearable acoustic and vibration measurements to derive novel digital biomarkers, which can be used together with existing medical information to assist in clinical decisions.

1.2 Major Contributions

- Designed and implemented an algorithm for deriving a *knee audio score*, which places the knee joint on a gradient from 0 (healthy) to 1 (affected), for knee acoustic emission signals captured with wearable sensors for patients with juvenile idiopathic arthritis.
- Designed and implemented a novel analysis framework which (i) detects the clicks heard from the knee joint using *Teager Energy Operator*, and (ii) classifies the detected clicks as *physiologic* or *pathologic* to provide an objective guide for knee health assessment.
- Demonstrated that when multiple features are fused together with machine learning, (i) higher accuracy and sensitivity are achieved in thrombosis detection in left ventricular assist devices (LVADs) compared to the state-of-the-art methods, and (ii) recurrent thrombosis cases, which are not identified by pump parameters or hemolysis markers, can be detected in advance.
- Designed and implemented, for the first time, a machine learning based algorithm for mapping wearable seismocardiogram (SCG) and sternal phonocardiogram (PCG) features to stroke volume in patients undergoing major surgery, and thereby elucidated the salient features of SCG and sternal PCG signals corresponding to stroke volume.

1.3 Scope and Organization of the Dissertation

The remainder of this dissertation is organized as follows:

Chapter 2 provides a short background on acoustic and vibration signals, and discusses the different acquisition methods to collect acoustic and vibration signals from

the body. Following this, the advantages and possible challenges of using acoustic and vibration signals in clinical applications are discussed.

Chapter 3 discusses how wearable acoustic measurements can be leveraged in biomechanics, specifically in joint health assessment, to derive clinically useful digital biomarkers. The first sub-section of Chapter 4 summarizes how knee acoustical emissions captured through miniature sensors can be benefitted to derive a clinically relevant joint health score for knee health evaluation [10-12]. In the second sub-section, a novel click detection and classification algorithm leveraging Teager Energy Operator is presented to detect the clicks in joint sound signals and distinguish between physiologic and pathologic clicks [13].

Chapter 4 studies the wearable acoustic and vibration measurements for cardiovascular assessment in two different applications. This first work shows that the operating sounds of LVADs may provide substantial information regarding pump thrombosis when combined with machine learning algorithms [14-16]. The second project discusses the use of wearable SCG in non-invasive stroke volume (SV) estimation and suggests that SV estimation benefits from both the acoustic and vibrational characteristics of the cardiovascular system [17].

Chapter 5 summarizes the concluding remarks and discusses the current challenges and future directions to achieve more robust and accurate detection and analysis frameworks. To that aim, several research questions and possible solutions are presented.

CHAPTER 2: ACOUSTIC AND VIBRATION SIGNALS

According to theory, sound is a form of energy originating from the vibration of particles. These vibrations require a medium to travel (air, liquid or solid), and are transmitted through a series of compressions and rarefactions, and thus they are mechanical by nature [18]. The signals in the audible range (20 Hz to 20kHz) are referred as acoustic signals. On the other hand, vibration signals fall under the infrasound group, which has a frequency content below 20 Hz. In some cases, these vibration signals are also referred to as low-frequency sound (Fig. 1) [18, 19].

As acoustic and vibration analysis can provide a non-invasive method to assist in acquiring useful physiological information from the body, they can potentially be leveraged in the design of wearable sensing modalities [20]. Such modalities can provide continuous monitoring during daily life regardless of time, environmental setting and stressors, and thus are of great importance for the digitization of healthcare [21].

2.1 Instrumentation for Acoustic and Vibration Measurements from the Body

2.1.1 *Contact Microphones*

The acoustical energy originating from the underlying physiological events travels to the surface of the skin where it can be acquired as *sound* [20]. However, as there is a great impedance mismatch between the air and fluid-filled tissue, a small amount of energy propagates to the air and most of the acoustical energy is reflected back to the tissue [22].

Theoretically, contact microphones can provide higher signal quality for on-body measurements than microphones capturing airborne sounds; contact microphones are less susceptible to background noise and can acquire the original non-attenuated signal directly from the surface of the body [20]. Thus, although air microphones (such as electret or microelectromechanical systems (MEMS) microphones) can be used to acquire the sounds emitted to air from the body, most studies leverage contact microphones (piezoelectric systems, stethoscope, etc.) in clinical or lab settings [20, 23]. Some example studies include heart sound analysis [24], joint health assessment [23], and laryngeal disorder detection [25] studies.

2.1.2 Digital Stethoscopes

Auscultation stands for the recognition of the signals in the audible range (20Hz-2kHz) from the body surface as a part of physical examination [26]. Auscultation is one of the most important steps in cardiovascular examination: specifically, in heart sound analysis, peripheral vascular system examination and blood pressure measurement. In the clinic, the ubiquitous tool for auscultation is a stethoscope, which transmits the sounds acquired from the chest piece to the examiner's ears through an air-filled hollow. René Laennec invented the first monaural stethoscope consisting of a hollow wooden tube in 1800s, and around the same time Golding Bird proposed a different version of the monaural stethoscope having a flexible tubing. Since Arthur Laered invented the binaural stethoscope, a number of advances have been made in the design of conventional stethoscopes [26]. The current binaural version consists of a chest piece having a diaphragm and a bell for transmitting high and low frequency sounds, respectively.

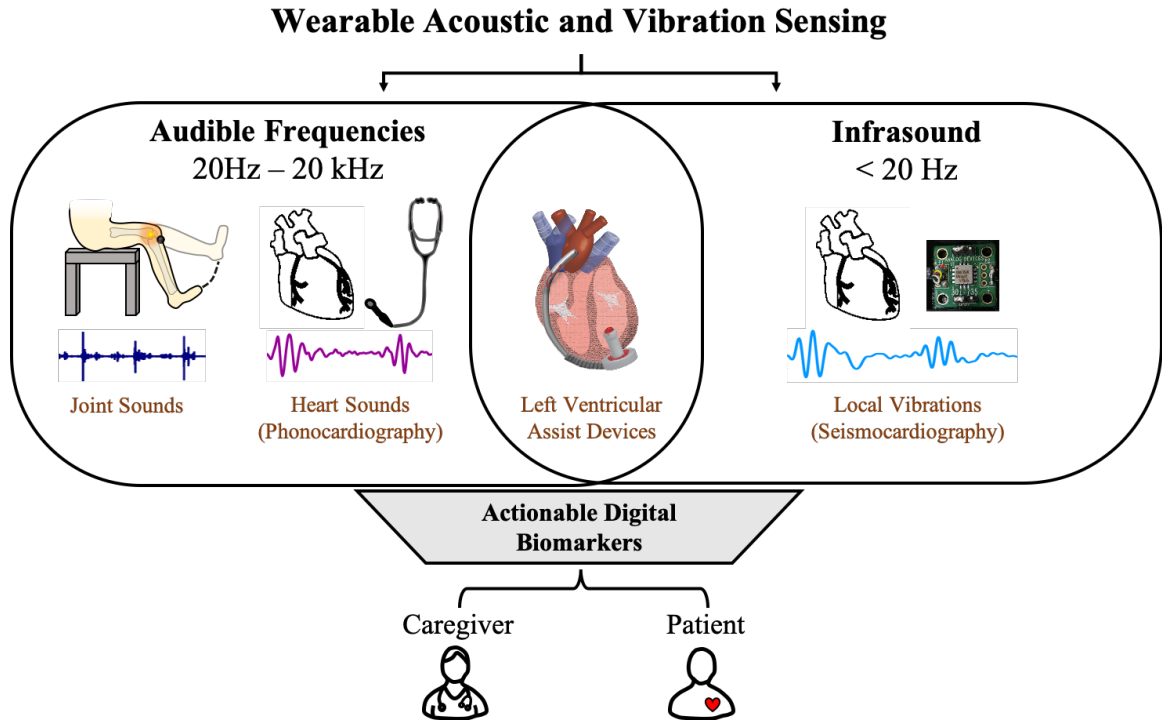


Figure 1 : Acoustic and vibration analysis can potentially be leveraged in the design of wearable modalities for deriving digital biomarkers to assist in clinical decisions. The signals in the audible range (20Hz-20kHz) are referred as acoustic signals, whereas vibration signals fall under the infrasound group, which has a frequency content below 20 Hz.

As the sensitivity of the human ear is variable, auscultation with conventional stethoscopes requires good listening training and clinical experience [27]. To overcome this challenge, *digital stethoscopes* emerged. Digital stethoscopes convert the acquired sound to electronic signals and consist of three main functionalities: data acquisition, pre-processing and signal processing. In the data acquisition step, the sound is captured, filtered, buffered, amplified and converted to a digital signal. The pre-processing step is responsible for removing artifacts and normalization, and signal processing step conducts feature extraction and classification for clinical decision making. The transducers used in

the digital stethoscopes may be of great variety, such as microphones, piezoelectric sensor, etc. Moreover, current digital stethoscopes provide different frequency response options to adapt to different auscultation purposes (lung, heart and other parts of the body) [26, 27].

2.1.3 Accelerometers

Today, the most commonly used sensor type in wearable system design is an accelerometer, which can be mounted on the skin using, for example, textiles [28], adhesives [29], or plastic mounting [30]. Accelerometers can respond to both acceleration from movement and acceleration from gravity [31]. In addition, they are very sensitive to physical vibrations [23], thus they have been widely used in various different studies ranging from monitoring cardiac-induced vibrations [32-34] to gait and motion analysis [35, 36].

Depending on the application, either uni-axial or tri-axial accelerometers can be leveraged. Although a uni-axial accelerometer is capable of acquiring the vibrations in the dorso-ventral direction, tri-axial accelerometers can provide information regarding all three directions: dorso-ventral, lateral and head-to-foot (Fig. 2) [37]. Several

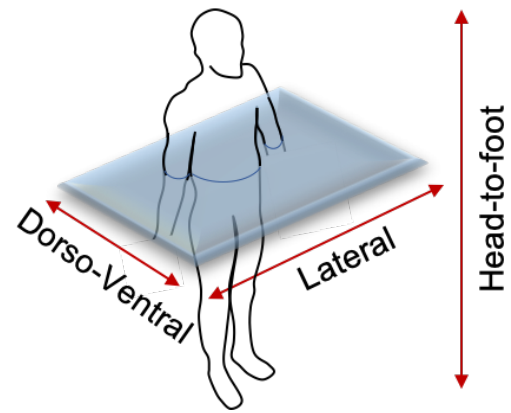


Figure 2 : Axes (planes) of the body

uni-axial and tri-axial accelerometer types include piezo-electric [38] and MEMS [39] accelerometers. With the advancements in sensor technology, accelerometers have become lighter, with low noise and power consumption, and thus can be readily integrated into

wearable systems. These improvements also increased the quality of the signals acquired; therefore, among similar technologies, accelerometers are more widely used in remote health monitoring applications [37, 40].

2.2 Advantages and Potential Challenges of Wearable Acoustic and Vibration Measurements

Acoustic and vibration signals carry information that is in many cases complementary to electrophysiology, biomechanics and movement. As they originate from the body, they have direct relationships to the underlying physiological events. However these signals are not fundamentally well understood, since most of this information has been obtained through empirical studies, and there has been no mechanistic model in the literature that describes the origin of these signals from a physiological standpoint [41]. Therefore, one of the main challenges with these signals is to decide whether the signal-of-interest is a good “*fit-for-purpose*” [1]. Unfortunately, there is a limited one-to-one correspondence between signal characteristics and important health parameters. To that end, the first aim of this work is to investigate this correspondence through signal processing, feature exploration and statistical techniques for deriving more accurate digital biomarkers.

In addition, physiological signals are highly subject-specific as each patient’s physiology reacts uniquely to medical interventions, treatment agenda and changes in environmental conditions [34]. Therefore, they can be great tools for individualized treatment through person-centered feedback systems [42]. Although their subject-specific nature is a great tool for individualized treatment, it also poses a major obstacle to the

development of accurate diagnostic tools. Similarly, repeatability of these measurements for a single subject is challenging as well, since the morphology of the physiological signals varies depending on the patient's body position (supine or standing), physical activity level (sleeping, resting, exercising) and sensor location or misplacement [33]. Thus, rather than focusing on adapting these signals as diagnostic tools, this work aims to derive and employ new algorithms to detect and track the *relative* changes in health, e.g. exacerbation in clinical state and/or response to a specific treatment, for a given subject over time.

Based on these motivating factors, the use of wearable acoustic and vibration signals in biomechanical and cardiovascular health assessment is discussed through several different applications in the following chapters. In each work, signal processing, feature exploration and statistical techniques are leveraged to investigate the correspondences between the acquired signals and targeted clinical conditions. Following this, digital biomarkers are derived for each application using machine learning-based models, and the changes in these biomarkers are monitored to track the relative changes in patients' clinical status.

CHAPTER 3: WEARABLE ACOUSTIC MEASUREMENT FOR BIOMECHANICS

The previous chapter presented a short background on acoustic and vibration signals, and discussed the different acquisition methods to collect acoustic and vibration signals from the body. One example of such acoustic signals is the waveform derived from measurements of *knee acoustical emissions*.

In this chapter, a novel hardware setup and signal analysis algorithm is proposed to contribute to the field of acoustical emission signal analysis and non-invasive joint health monitoring. To acquire the acoustical emissions from the knee joint, a small piezoelectric accelerometer is attached to the medial side of the patella. The recorded signals are then analyzed using a novel algorithm that computes a knee audio score, which places the recorded joint along a gradient from healthy to a joint with arthritis. In the second part, a novel click detection and classification algorithm leveraging the Teager Energy Operator is presented to detect the clicks in joint sound signals and distinguish between physiologic and pathologic clicks.

3.1.1 Using Knee Acoustical Emissions for Sensing Joint Health in Patients with Juvenile Idiopathic Arthritis

3.1.1 Introduction

Juvenile idiopathic arthritis (JIA) describes a clinically heterogeneous group of arthritides. It is the most common rheumatic condition in children and one of the more common chronic illnesses of childhood affecting more than 50,000 children in the United

States [43, 44]. The cause and pathogenesis of JIA are still poorly understood, but associations with various genetic and environmental factors have been made. There are seven defined categories of JIA, each with their own distinct presentation, clinical signs, symptoms, and clinical course. Identifying these categories coupled with recent studies into the genetic contributors of JIA have allowed for more precise treatment protocols; however, there is still significant variability within each of these categories. Each patient's clinical course is unique. Often, patients with JIA experience cyclical periods of active disease and remission. These unpredictable flare-ups, coupled with the highly variable causes and presentations, have made predicting the long-term prognosis of JIA difficult. This difficulty in predicting prognosis exacerbates the already difficult process of selecting an ideal treatment regimen [45, 46].

The pathophysiologic changes to the joints caused by JIA can lead to progressive joint destruction. The long-term sequela of JIA is severe and includes chronic pain, joint immobility, unstable articulation, and even disability. Fortunately, if JIA is detected and treated properly early in its presentation, the long-term consequences can be largely prevented [47, 48]. One of the joints most commonly affected by JIA is the knee [48]. Unfortunately, there are very few quantitative means for readily assessing the disease status in affected knees.

There are both invasive and non-invasive procedures for knee-health evaluation. Various imaging technologies such as computed tomography (CT), musculoskeletal ultrasound, magnetic resonance imaging (MRI), and fMRI are non-invasive evaluation methods, however they fail to provide early diagnosis, are prohibitively expensive for continuous monitoring, and are inconvenient to perform [49]. One of the more common invasive procedures for assessing knee health is arthroscopy. Arthroscopy provides detailed information, but it is a small surgical procedure, cannot be used on highly degenerated knees and is both cost and time intensive [50, 51]. In a chronic condition as

variable as JIA, there is a compelling need for a quantitative, unobtrusive, and affordable method for assessing joint health.

The use of acoustics -- recording the sounds that the joints make during movement -- may provide such a method [20]. In the case of the knee, vibrations are emitted from the mid-patellar region during active movements such as flexion and extension [52]. These vibrations can be measured on the surface of the skin as sound. These so called vibroarthrographic signals have been proposed as a possible diagnostic tool for early diagnosis of joint disorders [23]. Since Blodgett pioneered the technique in 1902, a number of advances have been made in the field of vibroarthrography [53]. Vibration arthrometry was used to show substantial differences in injured and healthy joints [54]. Later, the power profile was found to be significantly different between joints that were healthy, had rheumatoid arthritis and degenerative arthritis. It was also discovered that the majority of the signals occurred in the range of 20-20,000 Hz [55]. The field of vibroarthrography made a large leap forward with the development and application of piezoelectric accelerometers. This type of sensor is sensitive to physical vibrations (such as those seen on the skin during joint movement) and has less chance of registering external noises as compared to electret or other microphones sensing airborne pressure waves [56, 57].

Accurately recording the joint's acoustical emissions is only part of what is required for the development of a successful joint health monitoring suite. Those signals must also be analyzed and given physiologic context, such that this technique can be applied on patients. To analyze the recorded signals, different signal processing and machine learning techniques have been used. The techniques used thus far include wavelet decomposition [58], time-frequency analysis [59], Fourier analysis [60], autoregressive modelling [61], statistical parameter investigation and neural networks [62]. There is a significant need for, and lack of understanding about, which signal analysis technique is ideal for a given acoustical emission signal, and more particularly for a given disorder, such as JIA.

In order to contribute to the field of acoustical emission signal analysis and non-invasive joint health monitoring, a novel hardware setup and signal analysis algorithm have been developed. In the hardware setup, a small piezoelectric accelerometer is attached to the medial side of the patella to acquire the acoustical emissions from the knee joint. The recorded signals are then analyzed using a novel algorithm that computes a knee audio score, which places the recorded joint along a gradient from healthy to an involved joint with arthritis. In this work, both the hardware setup for recording knee acoustical emissions and the developed machine learning algorithm for classifying the knees were detailed. The proposed methods were validated via a human subject study involving four healthy control subjects and four subjects with JIA. Finally, to investigate the ability of this technique for quantifying disease progression and treatment response---essential components for the monitoring of JIA--- data acquired from the subjects with JIA before and after treatment was presented.

3.1.2 Methods

3.1.2.1 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. Ten subjects participated in this study. However, due to a microphone failure, two control subjects had to be excluded from the data analysis and thus data is presented from only eight subjects. Although the dataset was small, the number of instances was increased using appropriate window and step sizes, which made it possible to train different machine learning algorithms as previously done in [63-65]. Four of the subjects were diagnosed with JIA by a pediatric rheumatologist and four of the subjects were healthy controls with no history of JIA or acute knee injuries. The group with JIA consisted of three females (height= 157.1 ± 8.8 cm,

weight=48.9±12.3 kg, and age=14.7±2.1 years old), and one male (height=175.7 cm, weight=65.3 kg, age=17 years old). The healthy control group consisted of five females (height=141.7±10 cm, weight=34.1±3.6 kg, and age=9.6±1.8 years old) and one male (height=167.6 cm weight=54.5 kg, age=11 years old). In order to monitor the changes in knee acoustical emissions during the course of treatment, data were acquired from the subjects with JIA a second time, 3-6 months after initial measurements.

The data acquisition set up for each subject is shown in Fig.3(a). To record the sounds produced by the joints, a uniaxial analog accelerometer (3225F7, Dytran Instruments Inc. Chatsworth, CA) was attached 2 cm medial to the patellar tendon using KinesioTex tape (Kinesio Tex Gold, Kinesio, Albuquerque, NM). This accelerometer has a broad bandwidth (2Hz-10kHz), high sensitivity (100 mV/g), low noise floor (700 μg_{rms}), miniature size and low weight (1 gram). To ensure strong contact between the accelerometer and the subject's knee, the accelerometer was additionally wrapped in MEDca adhesive tape. The medial patellar location was selected due to the relatively unimpeded route (only a thin layer of muscle, tendon, and fat) to the articulating surface of the knee (where inter-joint friction is theorized to produce the recorded vibrations) [66, 67].

To record the knee acoustical emissions, the subjects performed ten unloaded knee flexion/extension exercises, while seated on a height-adjustable stool to prevent foot contact with the ground. The signals from the accelerometer were sampled at 108 kHz and recorded using a data acquisition module (USB-4432, National Instruments Corporation, Austin, TX). The exercise and recording protocol was repeated for both knees for all

subjects. The recorded signals were analyzed using Matlab (MathWorks, Natick, MA) and Python (Python Software Foundation, Beaverton, OR).

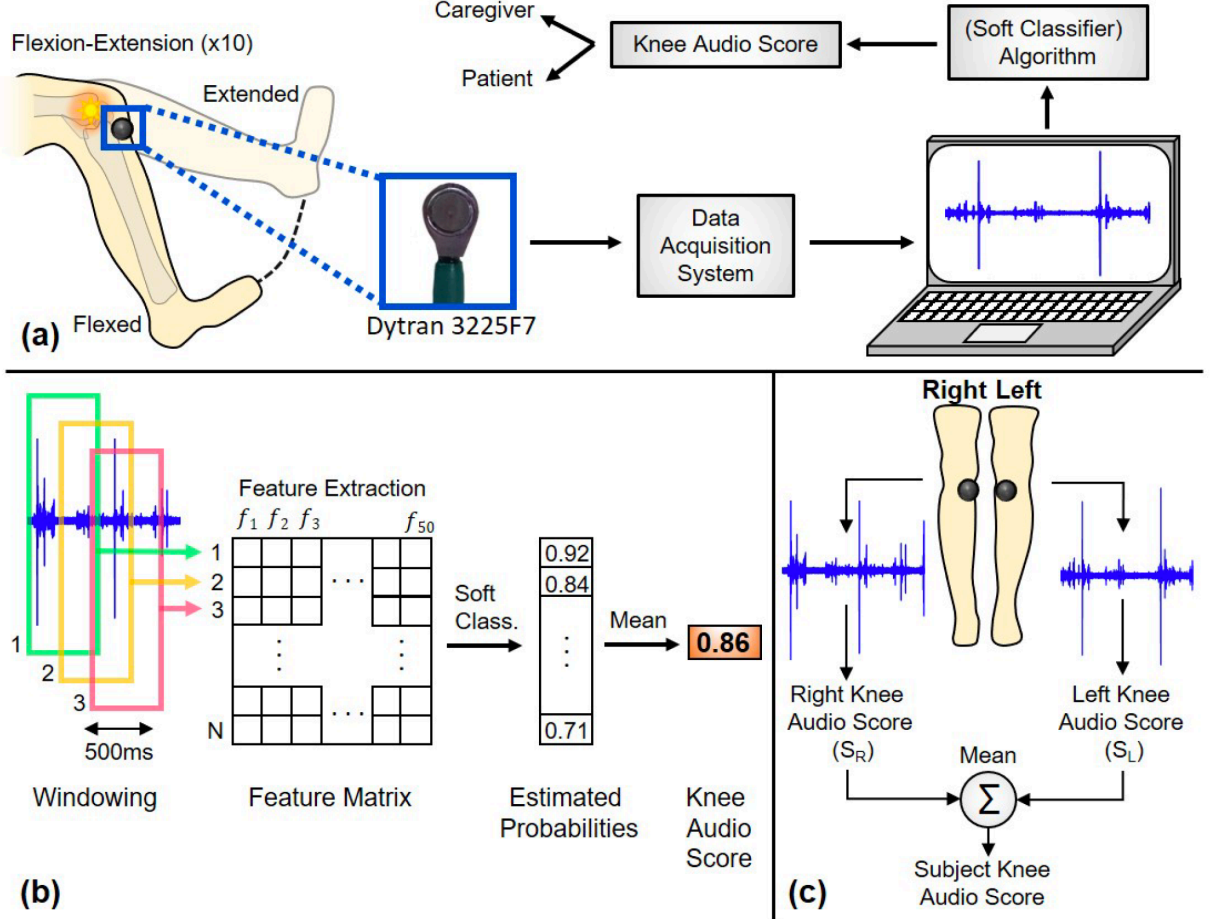


Figure 3: An overview of the methods used to acquire acoustical emissions from the knee and the algorithms used to analyze the signals. (a) The experiment setup. (b) The signal analysis workflow. (c) Knee audio score calculation.

3.1.2.2 Signal Processing and Feature Extraction

The joint acoustical emissions were analyzed in the time and frequency domains.

Fig.3(a) shows a representative plot of the time domain signal from one subject with JIA.

The acoustical emissions from knee joints have high energy and a short duration (between

10-20 ms). One unique characteristic of these signals is the audible “clicks” that have a spike-like appearance in the time domain plot. Additionally, these signals have high bandwidth frequency content reaching up to 20 kHz, which is fully expected for acoustical emissions [20, 55, 68].

Fig.3(b) details 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-20kHz bandwidth. Artifacts related to sensor placement and repositioning often observed at the beginning and end of the recordings are removed manually. These signals are then separated into 500ms-long frames with 50% overlap, resulting in $N=102$ frames. The frame length of 500ms ensures that a single frame comprises an adequate sample of the variable signal, including both a combination of silent segments of the signal and those where clicks are observed. Additionally, for this dataset, using 500ms-long frames allows for multiple joint sound signatures to be present within a given frame [68]. The 50% overlap is selected to increase the number of instances used in processing. An overlap greater than 50% could decrease the processing speed. Fifty signal features are extracted from each frame and stored in the matrices \mathbf{X}_R and \mathbf{X}_L for the right and left knees, respectively. The rows of these matrices represent a single signal frame, and the columns represent the 50 features extracted (see Feature Matrix in Fig.3(b)).

The features extracted are categorized as either “time domain”, “spectral”, “MFCC” or “bandpower” features. The time domain features are the zero crossing rate (ZCR), energy, and energy entropy ($f_1 - f_3$). The most important characteristic of these signals is the audible “clicks” that have a spike-like appearance in the time domain plot, and this unique, consistent *time domain* pattern of the signals result in distinctive time

domain features. On the other hand, the high bandwidth frequency characteristics of the joint sounds result in them having distinctive *spectral* features, which are the spectral centroid, spectral spread, spectral flux, spectral entropy and spectral roll-off ($f_4 - f_8$). The “MFCC” features are composed of the 13 mel-frequency cepstrum coefficients ($f_9 - f_{21}$), which have the ability to separate joint sound signatures from background noise as previously shown in [68]. The “bandpower” features consist of the signal power in 29 distinct frequency bands, between 30 logarithmically spaced frequencies in the range of 250Hz-20kHz ($f_{22} - f_{50}$). These extracted features are detailed in [68, 69].

3.1.2.3 t-Distributed Stochastic Neighbor Embedding (t-SNE)

The features extracted from the 500ms signal frames from all subjects are analyzed using machine learning techniques. For visualizing the ability of the feature set to distinguish between the control group and group with JIA, dimensionality reduction is performed using t-Distributed Stochastic Neighbor Embedding (t-SNE) [70]. Each feature of the data set represents one dimension. For this analysis, the dimensionality of the data is reduced from fifty to two dimensions for ease of visualization. Then, a scatter plot of the data was constructed with the two axes representing the two t-SNE dimensions and each point representing one 500ms signal frame. Two colors are used to categorically label the data points --- JIA in red and healthy controls in blue. If a particular feature set has the ability to distinguish between control subjects and subjects with JIA, the groups would form two separate clusters in the scatter plot.

The dimensionality reduction technique that t-SNE employs attempts to maintain the distances of points based on their probabilities of being neighboring data points.

Assuming that x_i is a point in the high dimensional space, x_i chooses x_j as its neighbor with the conditional probability $p_{i|j}$; likewise, in the low dimensional space, this probability can be represented as $q_{i|j}$. The t-SNE method aims to find the best low-dimensional data representation for minimizing the mismatch between the probabilities that two points are neighbors in high dimensional space ($p_{i|j}$) and low dimensional space ($q_{i|j}$) [70]. These probabilities represent a similarity metric between the two points. The t-SNE method is the preferred dimensionality reduction technique because, as shown in the literature, it minimizes local distortions and preserves the details within the data structure better than competing techniques, such as principal components analysis (PCA) or isometric feature mapping (ISOMAP) [71].

In order to visualize the dataset, first the matrices \mathbf{X}_R and \mathbf{X}_L from the right and left knees of all subjects are concatenated to form a matrix \mathbf{X}_{all} . As this matrix has features with different physical units, all of the columns are standardized to zero mean and unity variance. This allows all features to be weighted equally during dimensionality reduction. The labels corresponding to the rows of \mathbf{X}_{all} are stored in a vector y . A given entry of y is labeled 0 if the corresponding frame belongs to a healthy subject and 1 if it belongs to a subject with JIA. Using all features, a scatter plot is plotted. The points of the scatter plot are colored according to the vector y to investigate if there is separation between the two different groups.

3.1.2.4 Knee Audio Score Calculation

A *knee audio score* is calculated to place the knees on a gradient ranging from 0 to 1 where a score of 0 represents a healthy, unaffected knee and 1 represents an involved

knee with arthritis. To calculate the knee audio score from the features extracted from the acoustical emissions, a classification model is trained. Signals acquired from the subjects in the training set are separated into frames and fifty features are extracted from each frame as explained previously. These features are stored in the matrix $\mathbf{Z}_{\text{train}}$ and their corresponding labels are stored in the vector w_{train} . $\mathbf{Z}_{\text{train}}$ and w_{train} can then be used to train the classification model with the relationship between $\mathbf{Z}_{\text{train}}$ and w_{train} . After training, the model can be used to predict the label (healthy or JIA) of an incoming frame. The selected classification method supports “soft classification”, or the ability to estimate the *probability* that a given frame belongs to a subject with JIA [72]. Those probabilities are used to calculate the knee audio score from a recording. The knee audio score is the mean of the probabilities across all the frames from a given subject's recording. It is expected to obtain a higher score for subjects with JIA compared to control subjects. It is also hypothesized that this score will decrease with treatment.

The classification model is a necessary component of the knee sound analysis algorithm. The relationships observed in this type of data are non-linear and high-dimensional making more traditional classification techniques (which often apply linear methods) ineffective. In this project, Extreme Gradient Boosting (XGBoost) classification, which is a relatively new machine learning algorithm, is used [73]. XGBoost is an implementation of the gradient boosting machine learning algorithm [74] which falls under a category of learning algorithms called ensemble methods. In this type of algorithm, to predict a variable, multiple estimators are simultaneously used rather than the typical use of a single estimator [75]. In the XGBoost algorithm, many decision trees are iteratively

trained. This iterative training allows the model to predict the residual errors from the previous iteration and improve over time.

The assessment of the performance of the algorithm for classifying healthy vs. involved joints with arthritis is performed using leave-one-subject-out cross-validation (LOSO-CV) [76]. In each fold of the cross-validation, one subject is omitted and an XGBoost classifier is trained using the data from both knees of the remaining seven subjects. The trained model is then used to classify the signal frames of the excluded subject's knee acoustical emissions. The classified frames are stored in the vectors $y_{R,i}$ and $y_{L,i}$ for the right and left knees of the i^{th} subject. The same classification model is used to estimate the probability of JIA for each frame. These probabilities are stored in the vectors $p_{R,i}$ and $p_{L,i}$ for the right and left knees of the i^{th} subject.

The audio scores for the subject's knees are calculated by averaging the contents of the $p_{R,i}$ and $p_{L,i}$ for the right and left knees separately. An average knee audio score is also calculated for each subject by averaging the audio scores of the right and left knees (Fig.3(c)). A given subject is classified into the JIA group if the average knee audio score is greater than 0.5. For subjects with JIA, this process is repeated to calculate the knee audio scores for the post-treatment recordings. The cross-validation is completed by calculating knee audio scores for all eight subjects, excluding one subject per fold. Note that the post-treatment data for the JIA group is not used for training, as the ground truth labels for this data is not known certainly. The generalizability of the model is assessed by calculating the accuracy of the algorithm in labelling each frame. In addition, the accuracy

of the algorithm in predicting whether a subject belongs to the JIA or control groups is calculated.

3.1.2.5 Feature Importance Ranking

As was shown in Section 3.1.2.4, the knee audio score was calculated using the XGBoost classifier using fifty audio features. The relative weighting of each of those features in the model needs to be calculated to better understand which features are most relevant for the classification. All gradient boosting trees, including the XGBoost classifier train decision trees which can be used to rank the features according to their relative importance in the generated classification algorithm. Typically, the nodes of a tree divide using less important features while the initial node divides on the most important feature. The importance of features obtained from all the trees in the model are averaged resulting in the final relative feature importance scores [77]. These scores are applied to this particular data set helps to discern which features are most important for properly classifying joints with JIA.

Next, the XGBoost classifier is applied to evaluate which audio features were most relevant for distinguishing between the control subjects and subjects with JIA. The 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 and the resulting model is used to generate relative feature importance scores. No testing set is required to quantify feature importance as the aim is not to evaluate how well the model generalizes. After ranking the fifty features, how the top six features, which contributed the majority of the classification

strength, differed between the control subjects and subjects with JIA as well as the manner in which they changed post-treatment is analyzed.

3.1.2.6 Effect of Model Type

In order to evaluate the sensitivity of the results to the type of classifier used, a neural network is trained using the same audio features. The estimated class probability output of the neural network is used to score each audio frame as described in Section 3.1.2.4. Varying number of hidden layers (1 and 2), number of neurons in the hidden layers (16, 32, 64), and number of epochs (20, 50, 100, 200) are tested to classify the recorded acoustic emission data. All the activation functions in the hidden layers are chosen to be rectified linear unit (ReLU) activation functions. The final layer's activation function is chosen to be a sigmoid to get probability estimates as outputs. The network is trained using a binary cross-entropy loss function via a RMSprop optimizer [78]. Then, the best combination is found and the cross-validated accuracy values of the proposed approach and the neural network approach are compared.

3.1.3 *Results and Discussion*

3.1.3.1 t-SNE Visualization of the Knee Acoustical Emissions

The data from the four controls and the four subjects with JIA was visualized using t-SNE as described in Section 3.1.2.3. The t-SNE output using all features is plotted in Fig.4(a). In this plot, it is observed that the healthy subjects form two different clusters. The representative signals in the time domain and their spectrograms from each of the clusters were plotted in Fig.4(a). To determine the cause of the two distinct clusters formed

by the healthy controls' knee sounds, the time-domain plots and spectrograms of a representative subject from these two clusters were compared (Fig.4(c) vs. Fig.4(d)). The plot in Fig.4(c) is of one of the healthy controls from the cluster that overlapped with the JIA data clusters in the t-SNE plots. It can be seen that the signal shown in this plot has a number of clicks. Fig.4(d) is from a healthy control from the cluster with no overlap of the JIA data clusters in the t-SNE. This plot has no sizeable clicks. It can be inferred from the difference in these sound profiles that the healthy data with some clicks led to the overlap in the t-SNE analysis with the arthritis data that included much more prevalent and louder clicks. On the other hand, the healthy subjects with no clicks form the t-SNE clusters that are distinct from the JIA group [79, 80]. This is simply a minor issue that will not really have an impact on the classification process, but it is an interesting scientific observation worth mention.

The subjects with JIA (Fig.4(b)) show periodic high energy clicks in each flexion-extension cycle. These clicks have frequency components that reach up to 20 kHz. The most important characteristic of these signals is the audible “clicks” that have a spike-like appearance in the time domain plot. This unique, consistent “time domain” pattern of the signals belonging to the group with JIA results in them having distinctive time-domain features. On the other hand, the high bandwidth “frequency characteristics” of these signals result in distinct spectral features which increases the separation ability of our feature set. More work is needed to determine the origin of these high frequency clicks, but it is hypothesized that they occur due to the degraded articulating surface creating increased friction in the joints of patients with JIA. The degraded articulating surface of the knees

develops from the chronic inflammation of the synovial membrane and reduced joint space [81].

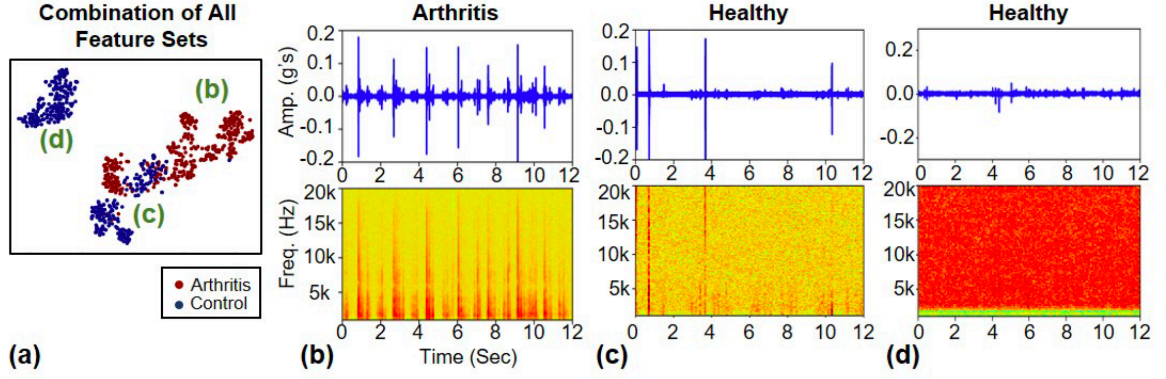


Figure 4: (a) The t-SNE visualizations of knee acoustical emission data from four control subjects and four subjects with JIA using all features. Representative signals and spectrograms from a subject with JIA (b), and two control subjects (c) and (d).

3.1.3.2 Knee Audio Score Assessment

Fig.5(a) shows the knee audio scores of the right legs, left legs and the average of both for all subjects. A knee audio score of 1.0 corresponds to an involved joint with arthritis; a score of 0.0 corresponds to a healthy knee. A subject's score was calculated by training the model on all of the other subjects (LOSO-CV). The average left knee audio score was 0.022 ± 0.025 and 0.96 ± 0.044 for the control group and group with JIA, respectively ($p < 0.05$ using Mann-Whitney U Test). The average right knee audio score was 0.147 ± 0.17 and 0.83 ± 0.26 for the control group and group with JIA, respectively ($p < 0.05$ using Mann-Whitney U Test). The average knee audio score was 0.085 ± 0.099 and 0.89 ± 0.012 for the control group and group with JIA, respectively ($p < 0.05$ using Mann-Whitney U Test). The calculated knee audio scores were closer to 1.0 for the subjects with

JIA while the scores were closer to 0.0 for the control subjects as expected. It can be observed that the right and left knee audio scores for both JIA and control subjects differed by 0.2 (20% of the score's range). Therefore, using this method, asymmetries in symptoms could potentially be detected and monitored as well. This capability would be useful in tracking compensation when one of the knees is more symptomatic than the other by comparing the calculated knee audio scores.

The subjects with average audio scores higher than 0.5 were classified as JIA and lower than 0.5 as healthy. This led to a cross-validated accuracy of 100% for the dataset. Additionally, each signal frame was classified as JIA or healthy (as explained in Section 3.1.2.4) which led to a cross-validated accuracy of 92.3%.

Next, in an attempt to quantify the response of the subjects with JIA to treatment, the knee audio scores were calculated for the data recorded pre-treatment and 3-6 months post-treatment (Fig.5(b)). Again, the data statistics for the right leg, left leg and averaged knee audio scores are plotted. The right knee audio score for these subjects decreased from 0.83 ± 0.26 to 0.27 ± 0.42 ($p > 0.05$ using Wilcoxon Test). The left knee audio score for these subjects decreased from 0.96 ± 0.044 to 0.24 ± 0.34 ($p < 0.05$ using Wilcoxon Test). The average knee audio score for these subjects decreased from 0.89 ± 0.012 to 0.25 ± 0.20 ($p < 0.05$ using Wilcoxon Test). Of note, each of the post-treatment audio scores fell below the selected 0.5 classification threshold and thus were classified as healthy. The before and after treatment scores were statistically significant except the right knee scores, where statistical significance was not present as one subject's score increased post-treatment.

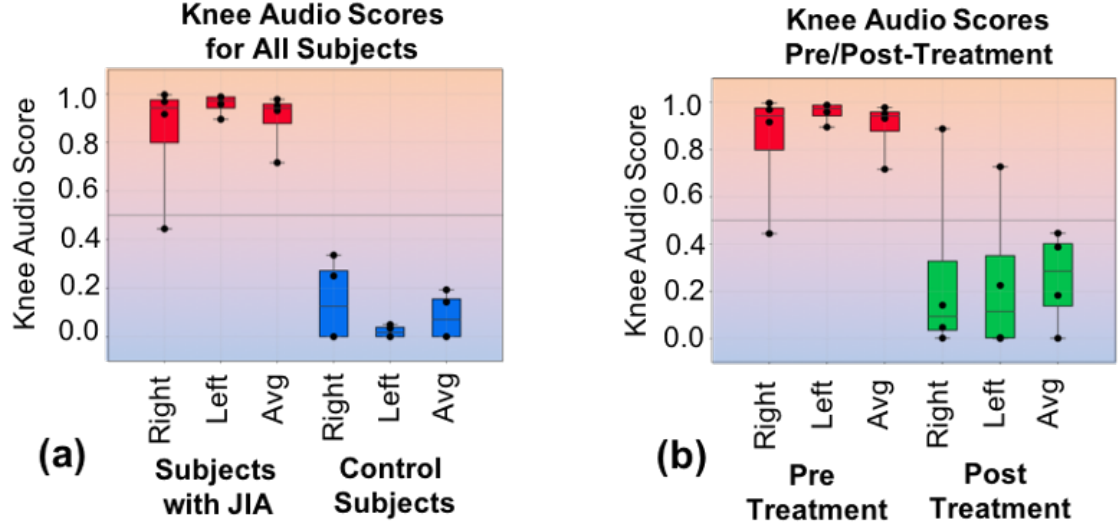


Figure 5: (a) The right, left and average knee audio scores of the subjects with JIA and control subjects. (b) Knee audio scores for the subjects with JIA before and after treatment.

3.1.3.3 Feature Importance Ranking

The fifteen most important features from the feature importance ranking are shown on Fig6(a). Features from all sets (time, spectral, MFCC and band power) can be seen within the top fifteen features, showing the importance of a diverse feature set. The Kernel Density Estimated distributions of the top six features are shown for the control, pre-treatment JIA and post-treatment JIA data [77]. It can be seen in Fig.6(b) that control subjects have higher Zero Crossing Rate (ZCR) than the subjects with JIA as the signals from the control group are closer to the noise floor. Fig.6(c) shows that the energy distribution of the signal frames is slightly narrower for the control subjects compared to the subjects with JIA. This can be attributed to the fact that the subjects with JIA have a more variable set of signal frames ranging from silent to ones that contain clicks. The healthy subjects mainly contain low energy (silent) signal frames. A similar observation can be made for the spectral spread in Fig.6(d) for the same reasons. In Fig.6(f), the subjects

with arthritis have lower spectral rolloff on average than the controls. This might be a similar phenomenon as white noise having a high bandwidth since the healthy subjects have signals that are closer to the noise floor and thus a higher spectral rolloff. The differences in MFCC-13 and band power-26 seen in Fig.6(e) and (g) are due to the differences in the frequency content of signals from the two groups. The differences seen in the feature distributions (Fig.6(b)-(g)) between the pre-treatment JIA and control groups demonstrates the ability of separating these groups based on those features. The post-treatment data is distributed similarly to the healthy control subjects indicating that the treatment was successful and the patients with JIA returned to their healthy baseline.

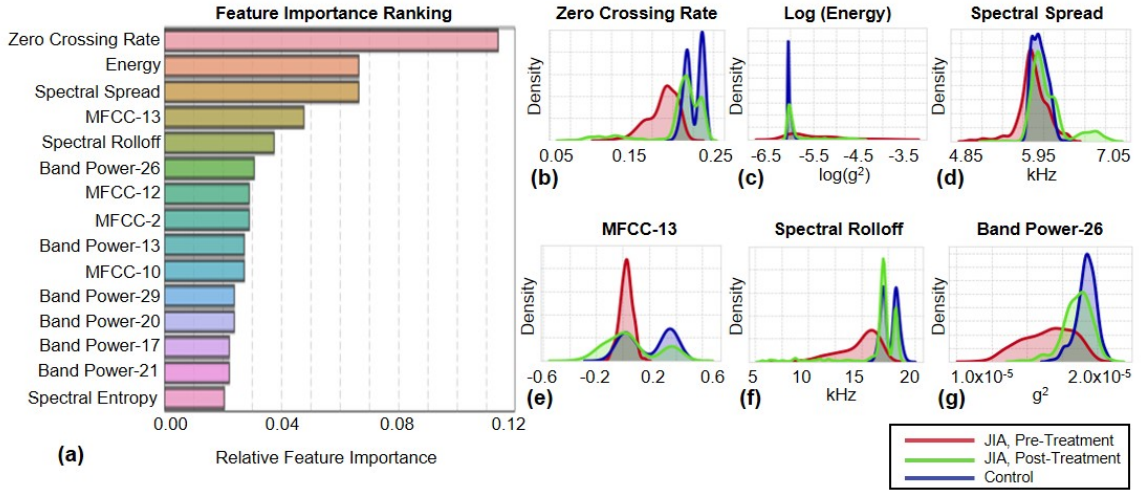


Figure 6: (a) Feature importance ranking showing the top fifteen features in terms of salient information provided for classification. (b)-(g) Distributions of the top six features for the control subjects and pre- and post-treatment subjects with JIA. The distributions were calculated using kernel density estimation.

3.1.3.4 Effect of Model Type

The best result we obtained using a neural network was 72.9% cross-validated accuracy with 2 layers, 32 neurons and 100 epochs. As seen, a great discrepancy was

observed between the result of the proposed algorithm (92.3%) and the one we calculated using neural networks (72.9%). Although using neural networks is an appropriate approach for audio and speech processing applications, it is not the best fit for this study because of the limited size of the dataset. Nevertheless, as more subjects are recruited for the future studies, there will be a chance to apply deep neural networks properly as well.

3.1.4 Conclusion

In this work, the recording and analysis of acoustical emissions from the knee joints of healthy subjects and subjects with JIA were described. It was demonstrated that acoustical emissions acquired from the knee joint can potentially be used as a quantitative metric in the assessment and monitoring of JIA.

First, knee acoustical emission data acquired from control subjects and subjects with JIA was visualized using t-SNE plots. The ability of different feature sets to distinguish between the two groups was explored. Following this analysis, an algorithm to compute a knee audio score was designed. This algorithm utilizes a soft classification model based on gradient boosted trees. It was shown that this algorithm can accurately distinguish control subjects and subjects with JIA for the dataset using LOSO-CV. Also, how the knee audio score of four subjects with JIA changes before and after treatment was demonstrated. Finally, to understand the contribution of individual audio features to the algorithm, feature ranking was performed. This analysis revealed that elements from all the feature sets that contributed to some degree in the knee audio score calculation.

This work provides, for the first time, a framework for extracting a clinically-relevant and actionable joint health score from acoustical emission signals measured with

miniature sensors that can be embedded into a wearable system for home use. The approach leverages several key innovative concepts, including the adaptation of a gradient boost regression algorithm as a soft classifier, and the use of t-SNE as a means of discovering features of relevance from the high-dimensional acoustical data. While the studies performed were with a small group of patients, the results are already statistically significant and suggest that the approach holds merit and should be tested and validated further in larger populations. The concept of delivering a joint health score to the physician or patient to assist in clinical decisions can be extended beyond JIA to assistive rehabilitation following musculoskeletal injury; such a joint health score can, for example, be used to determine when an athlete rehabilitating an acute injury can resume certain activities or intensities of activities.

In future studies how different methods of interfacing the sensors to the skin affects the acoustical emission signals will be investigated. Furthermore, the feasibility of integrating the sensors to a knee brace or sleeve for continuous monitoring will be explored. In addition, the underlying physiology and source of the knee acoustical emissions using cadaver models will be studied. Most importantly, the statistically significant, but still preliminary, results from this work will be validated in a larger dataset of patients and controls. Upon completion of rigorous testing and validation in a larger population of subjects, the proposed approach can potentially be used to deliver a joint health score to the physician or patient to assist in clinical decisions.

3.2 Distinguishing between “Physiologic” and “Pathologic” Clicks using the Teager Energy Operator

3.2.1 Introduction

In the previous section, a novel hardware setup and signal analysis algorithm was developed to assess knee-joint health. The proposed algorithm was validated with the measurements taken from 4 control subjects with no history of knee injury or pathology and 4 subjects diagnosed with JIA with affected knees. Several time and frequency domain features were extracted from the acoustic emission signals and fed into a soft classifier which is based on gradient boosting trees. Using LOSO-CV, each subject was given a *knee audio score* ranging from 0 to 1, with 0 being a healthy knee and 1 being an involved joint. Although the cross-validated accuracy of the signal frames was high (92.3%), there remained an important observation and a potential drawback for the proposed algorithm. There were clicks that occurred in the healthy controls’ joint sounds (Fig. 7(a)) that appeared to resemble the clicks seen in the joint sounds of patients with JIA (Fig. 7(b)). The previously proposed algorithm did not focus on distinguishing the source of these clicks or classifying the clicks as physiologic versus pathologic. These physiologic clicks may arise from bursting of bubbles in the synovial fluid, ligament snapping, or physiologic synovium catching [82]. Although these physiologic clicks are generally not associated with joint or synovium pathology, they still differ from the silent baseline knee recordings in terms of their time and frequency domain characteristics. Therefore, to prevent any false

positives during the analysis of joint sound recordings, these physiologic clicks should be detected precisely and be distinguished from the pathologic clicks.

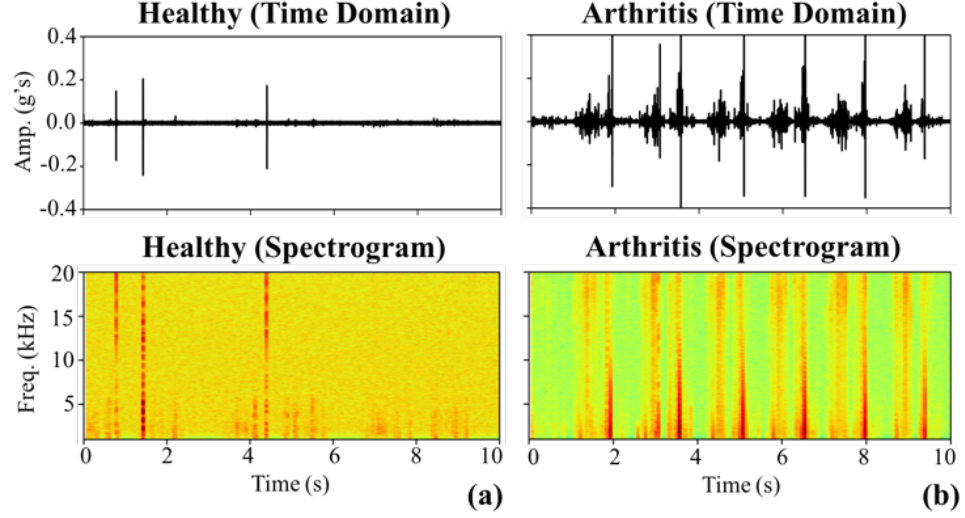


Figure 7: (a) Time domain signal and corresponding spectrogram of a recording taken from a control subject with physiological clicks. (b) Time domain signal and corresponding spectrogram of a recording taken from a subject with JIA.

It was hypothesized that these physiologic and pathologic clicks have different characteristics and could be differentiated with an automated system allowing for more precise classification of knee health status. This work presents a novel algorithm to distinguish between these pathologic clicks and physiologic clicks. First, the desired clicks are located and extracted from the band-pass filtered signals using a detection pipeline (Fig.8). These clicks are then classified as “physiologic” or “pathologic” using a random forest classification algorithm.

3.2.2 Methods

3.2.2.1 Human Subject Protocol

This study was conducted under a protocol approved by the Georgia Institute of Technology and Emory University Institutional Review Boards. Knee acoustic emissions from 4 control subjects and 4 subjects with JIA were acquired as described in [10]. The recorded signals were analyzed using Matlab (MathWorks, Natick, MA) and Python (Python Software Foundation, Beaverton, OR). In this work, the recordings from the control subjects are compared against the recordings taken from patients with JIA prior to starting treatment. Post-treatment data is not included as the true nature of the clicks (pathologic vs. physiologic) could not be certainly known at this point.

3.2.2.2 Pre-Processing and Acoustic Noise Suppression

The signals are pre-processed using a digital finite impulse response (FIR) band-pass filter with a 250Hz-20kHz bandwidth to reduce interface noise. Sound artifacts at the beginning and end of the recordings are removed manually (Fig.9(a)). After this preprocessing step, the spectral noise suppression algorithm developed by Ephraim and Malah [83] is employed to reduce the background noise. Each recording is segmented with Hanning windows and the corresponding short time Fourier transform (STFT) is computed to calculate the signal spectral power. Assuming that the noise is stationary, the power spectrum of the noise is computed using the small silent portions of the recordings where no clicks are present. For each frame, a posteriori (ratio of the noisy speech spectrum and the noise spectrum) and a priori (ratio of the clean speech spectrum and the noise spectrum) signal-to-noise ratio (SNR) values are calculated using the decision-directed approach described in [83]. The signal gain is updated after each frame and this gain vector is used to clean the actual STFT. Finally, the noise-reduced signal is reconstructed using inverse STFT and overlap-add methods (Fig.9(b)).

3.2.2.3 Teager Energy Operator

The Teager energy operator is a non-linear operator derived from the energy of a simple oscillator using the physics of motion. It represents the running estimate of the signal energy by operating on three sequential samples of the signal at a given time point [84]. In the discrete-time case, it is defined as:

$$\Psi_{TE}[n]=x[n]^2 - x[n + 1]x[n - 1] \quad (1)$$

This operator is generally used to detect instantaneous changes in signals such as discontinuities, changes in signal amplitude and/or changes in frequency. Additionally, it suppresses the background activity and discards soft changes [85]. Knee acoustic emissions usually have high energy and short duration [10, 55], therefore, the Teager operator appears to be well-suited for detecting the instantaneous changes in these signals: specifically, the “clicks” in the recordings.

3.2.2.4 Click Detection

After computing the Teager energy operator, the locations of these change points are stored for each recording using a simple peak detection algorithm. First, the envelope of the Teager operator is generated, then the peaks which are greater than 20% of the range of the signal value are selected as the potential click locations. This threshold is selected heuristically to include as many clicks as possible while ignoring small fluctuations in the signal. Using these clicks locations, the portions within *click location* $\pm 50ms$ are extracted from the main signals for each detected click. This 100ms window-length is selected to include the main click and its smaller, subsequent clicks (Fig.9(c)). A total of 53 clicks (30

from the involved knee recordings and 23 from the control recordings) are stored in matrix \mathbf{Y} for further processing and classification.

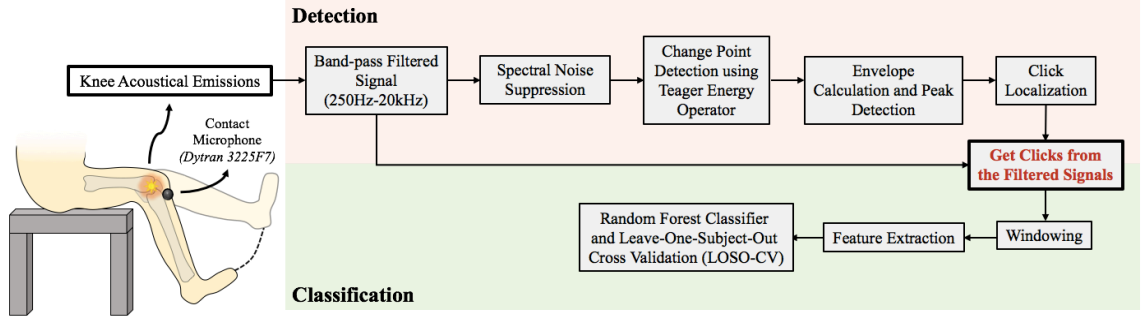


Figure 8: Detection and classification pipeline. Knee acoustic emissions are filtered, and spectral noise suppression is applied. Using the Teager energy operator, change points (in this case "clicks") are detected. Using these click locations, 100ms long click segments are extracted from the signals. Clicks are windowed, and time and frequency domain features are extracted from each window. Using the random forest classifier and LOSO-CV, the clicks are categorized as "physiologic" or "pathologic".

3.2.2.5 Feature Extraction

Each 100ms-long click in matrix \mathbf{Y} is segmented into 10ms windows using 50% overlap to increase the number of instances for the classification algorithm. This resulted in 954 frames in total (18 frames from each click), and time and frequency domain features are extracted from each frame. The wide bandwidth (250Hz to 20kHz) of the joint sounds causes these signals to have distinctive spectral features, whereas the spike-like appearances in the time domain result in unique time domain features. Thus, 8 total features are extracted: 3 in the time domain and 5 in the frequency domain. In the time domain, the zero crossing rate, energy, and energy entropy are calculated. In the spectral domain, the spectral flux, spectral spread, spectral entropy, spectral roll-off, and spectral centroid are calculated. The performance of these features and the importance of having a diverse

feature set in knee acoustic emission analysis were previously discussed in [10]. These features are stored under the 954x8 matrix \mathbf{X} .

3.2.2.6 t-Distributed Stochastic Neighbor Embedding (t-SNE)

To visualize the ability of the feature set to distinguish between the physiologic and pathologic clicks, dimensionality reduction using t-distributed stochastic neighbor embedding (t-SNE) is applied [70]. In the current problem, each feature represents one dimension. Using t-SNE, the 8-dimensional data (with each dimension corresponding to one of the 8 features) is mapped onto 2-dimensional space for better visualization. After dimensionality reduction, a scatter plot with two axes, where each point is representing a 10ms signal frame, is constructed. Two colors are used to categorically label the data points - pathologic clicks in red and physiologic clicks in green.

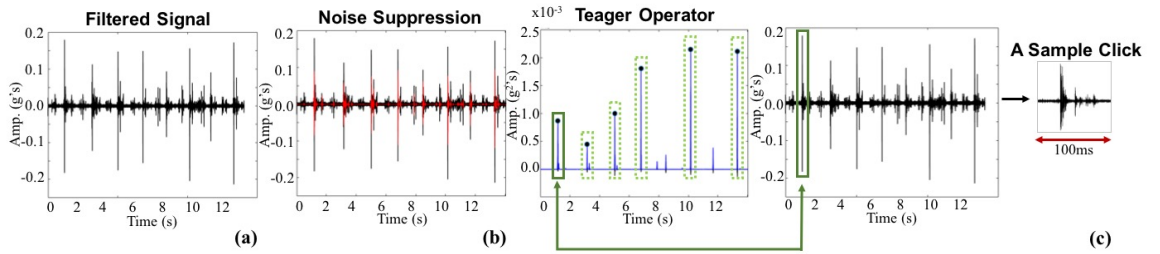


Figure 9: (a) The signals are filtered using a 250Hz-20kHz band-pass filter. (b) Spectral noise suppression is applied for background noise. (c) Using the Teager energy operator, change point locations are detected. Using this information, 100ms long click segments are clipped from the band-pass filtered signals.

3.2.2.7 Random Forest Classification

As previously explained, the clicks clipped from the filtered signals are windowed and 8 features are extracted from each frame. These features are stored in matrix \mathbf{X} . The pathologic signals and physiologic signals are given the labels 1 and 0, respectively, and

these labels are stored in 53×1 vector w . Random forest classification with 100 estimators and maximum depth value of 7 is employed to classify 954 frames from 53 physiologic or pathologic clicks. Random forest classifier is used, since it can achieve high accuracy values without causing any overfitting. Random forests are sets of trees with the same distribution where each tree depends on the values of an independently sampled random vector [86]. After several trees are generated, each tree casts a unit vote for the most popular class in the input data – in this case for “physiologic” and “pathologic” classes.

The performance assessment is performed using LOSO-CV. In each fold, all click frames from one subject are left out, and the model is trained on the remaining click frames from 7 subjects. Then the model is tested on the click frames of the subject being left out. This procedure is completed for all 8 subjects and the predicted classes of the frames are stored in 53×18 matrix Z , where 53 is the number of clicks and 18 is the number of frames belonging to each click. These frame scores are then averaged to get the final score of each click and stored in 53×1 vector h . The clicks are classified as “pathologic” if the calculated score is greater than 0.5, an as “physiologic” if less than 0.5. Finally using the predicted scores in vector h and the actual classes in vector w , the performance metrics (accuracy, sensitivity and precision) are calculated and the corresponding receiver operating characteristics (ROC) curve is plotted.

3.2.3 *Results and Discussion*

3.2.3.1 t-distributed Stochastic Neighbor Embedding (t-SNE)

The data from 4 subjects with JIA and 4 control subjects is visualized as described in Section 3.2.2.6. In Fig.10(a) the corresponding t-SNE graph for the physiologic (green

dots) and pathologic (red dots) clicks is presented. The frames from these two click types construct separate clusters in 2-dimensional space based on their time and spectral domain features. This separation supports the ability of the feature set to distinguish between these two groups, and supports the hypothesis about these physiologic and pathologic clicks having different characteristics. Physiologic clicks may be caused by cavitation in the synovial fluid, ligaments snapping, etc. [82]. The causes of clicks in JIA are yet unknown, but are most likely attributed to the chronic inflammatory state of the synovium and cartilage degradation [79-81]. Since these clicks have different origins, it is reasonable that there would be differences in their time and frequency information.

3.2.3.2 Random Forest Classification

As explained in Section 3.2.2.7, 53 physiologic and pathologic click frames are classified using random forest classifier and the model is validated using LOSO-CV. This led to cross validated accuracy, sensitivity and precision of 94.3%, 93.3% and 96.6%, respectively. Similarly, the ROC curve is plotted (Fig.10(b)), and the corresponding area under curve (AUC) is calculated to be 0.98. In biomedical applications, along with high accuracy, a high sensitivity value is preferred in the design of a screening test/technology, since this metric corresponds to the “detection rate” of the algorithm. In this study, a detection rate of 93.3% was obtained, which shows that this approach can detect and interpret the changes in joint sound signals. Thus, this algorithm could potentially be used to track and evaluate the knee health status of both healthy and diseased knees. In the future, this click-scoring pipeline could be extended beyond JIA to assist in rehabilitation and recovery following musculoskeletal injuries. Additionally, this signal differentiating capability could be utilized in remote monitoring or telemedicine frameworks. By using

miniature sensors embedded in a wearable system, the joints could be continuously monitored, and joint status changes or disease exacerbations could be detected or maybe even predicted. This early detection would improve patients' quality of life and reduce surgeries and healthcare costs.

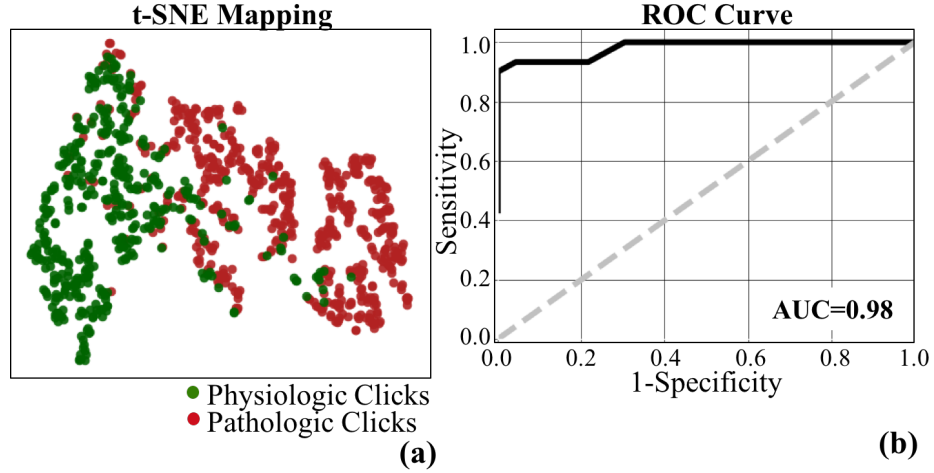


Figure 10: (a) t-SNE graph for the physiologic and pathologic clicks. (b) The ROC curve for the proposed model. The AUC is calculated to be 0.98.

3.2.4 Conclusion

In this work, a new algorithm for detecting and classifying physiologic and pathologic clicks from the knees of subjects with JIA and healthy control subjects was presented. An automated algorithm such as the one presented in this work can potentially be used to assist clinical decision making and provide knee-health tracking in wearable systems for home-monitoring. In future studies, these findings should be validated with larger datasets and attempt to leverage not only supervised learning algorithms, but also unsupervised ones to ensure the generalizability of the proposed approach.

CHAPTER 4: WEARABLE ACOUSTIC AND VIBRATION MEASUREMENT FOR CARDIOVASCULAR ASSESSMENT

Previous chapter investigated how acoustical emissions acquired from the knee joint can potentially be used as a quantitative metric in the assessment and monitoring of JIA. Similarly, recent studies have shown that wearable acoustic and vibration signals may provide substantial information regarding cardiovascular health, and enable continuous and non-invasive cardiovascular health assessment and monitoring under various settings and stressors [21].

The first work in this chapter discusses pump thrombosis detection in LVADs based on determining salient acoustic and pump features indicative of pump thrombosis and evaluating the added value of acoustic features in thrombosis detection using machine learning algorithms. The second project investigates how the combination of vibration and acoustic signals, specifically the SCG and PCG, captured through a wearable patch can be leveraged to enable continuous and non-invasive SV monitoring.

4.1 Detecting Suspected Pump Thrombosis in Left Ventricular Assist Devices via Acoustic Analysis

4.1.1 Introduction

Pump thrombosis occurs in up to 10-13% of left ventricular assist device (LVAD) recipients within the first year post implant [87, 88]. Thromboembolic events may lead to pump failure requiring replacement at substantial rates (~5-7%) which has been previously reported in large trials [89-91]. LVADs stimulate the coagulation cascade resulting in

thrombus formation and, when exposed to the shearing force of blood flow, patients are predisposed to thromboembolic complications [92]. Pump thrombosis may result in hemodynamic derangement, stroke, and death [90, 92-94]. Presence of hemolysis {elevated plasma lactate dehydrogenase (LDH) or plasma free hemoglobin (pfHb) levels} and degradation of device performance (e.g., increased power) can indicate pump thrombosis, which is further investigated using echocardiogram ramp studies [95, 96].

Resultant end-organ dysfunction requires the exchange of the pump through an additional surgical procedure, which is expensive, and can contribute to higher allosensitization that is correlated with worse heart transplant graft survival outcome if associated with blood product use [97]. Moreover, device exchange does not preclude recurrence of thrombosis if the origin is biological rather than mechanical [93]. Therefore, device thrombosis should be mitigated with earlier diagnosis, before triggering any heart failure (HF) exacerbation and additional surgeries. In particular, outpatient monitoring is needed as up to 15% of patients are readmitted due to device thrombosis and subsequent complications increase healthcare costs and decrease patients' quality of life [98, 99]. Remote monitoring of pump parameters and hemodynamics should be achieved to allow real-time communication between caregivers and patients [100]. However, without having the required blood-derived biomarkers, it is not feasible to evaluate the normal vs. pathologic processes, or biological responses to therapeutic interventions. Nevertheless, such monitoring can be achieved by deriving “digital biomarkers” [1], which are measured through home-based sensors, wearable devices and implants, to support continuous measurement outside the physical confines of the clinic. Unfortunately, there is no technology currently available commercially nor in the research domain that has

demonstrated the ability to characterize and monitor hemodynamics and pump functionality of VAD recipients at home.

In clinical settings, the operating sounds of these artificial blood pumps has been studied for assisting in thrombosis detection [101, 102]. Several studies, both in vitro [102-107] and in vivo [101, 102, 104, 106, 108-111], have evaluated acoustic analysis as a non-invasive method for thrombosis detection in axial and centrifugal pumps. These studies mostly leveraged harmonic frequency analysis and focused on detecting changes in the acoustic features following pump thrombosis. For example, Kaufmann *et al.* [101] asserted that the most intense harmonic is the fourth harmonic in the centrifugal pump recordings. Based on this assertion, they calculated the intensities of the first four harmonics and normalized the intensities of the first three to the fourth. They suggested that an increase in first and second harmonic intensities and the existence of the third harmonic indicate pump thrombosis. In addition, Yost *et al.* [104] and Castagna *et al.* [110] performed similar harmonic analysis on axial pump recipients. Yost *et al.* suggested that the normalized harmonic intensities decrease as the pump starts to develop thrombosis. On the other hand, Castagna *et al.* proposed that an increase in the normalized power of the first two harmonics and a decrease in the third harmonic indicate development of pump thrombosis.

To the best of our knowledge, no validated method has been reported so far. Moreover, the reliance on a single feature of pump acoustics - harmonic content - as an indicator of pump thrombosis may not be generalizable to all datasets and all pump types. In this work, there were two objectives: (1) determining salient acoustic and pump features indicative of pump thrombosis; and (2) evaluating the added value of acoustic features in

thrombosis detection. Ultimately, this approach may enable detection and diagnosis of pump thrombosis with improved accuracy in clinical and/or home settings.

4.1.2 Methods

4.1.2.1 Study Protocol and Subject Demographics

This study was conducted under a protocol approved by the University of California San Francisco (UCSF) and Georgia Institute of Technology Institutional Review Boards and all patients provided written informed consent. Thirteen adult centrifugal pump recipients were enrolled in the study. At the time of enrolment - the first post-implant clinic visit when the patient was deemed clinically optimized - a baseline sound recording was obtained and baseline pump parameters and laboratory data were collected. All the LVAD patients implanted at UCSF are followed using a remote management system (Abbott Alere Home Monitoring, Livermore, CA or Acticare Health, Livermore, CA) that records daily pump parameters (speed, power, flow, pulsatility index, peak, trough, alarms), vital signs (blood pressure, heart rate, weight) and point of care CoaguCheck INR (Roche Diagnostics, Indianapolis, IN). In case of abnormalities in the pump parameters or HF symptoms, patients are contacted and hemolysis biomarkers (LDH, pfHb) are promptly obtained to establish the diagnosis of suspected pump thrombosis. During the study, new sound recordings were obtained in cases where there was a change in pump speed or clinical condition. All subjects were recipients of HeartWare HVAD (Medtronic Framingham, MA) device for either bridge-to-transplantation or destination therapy indications.

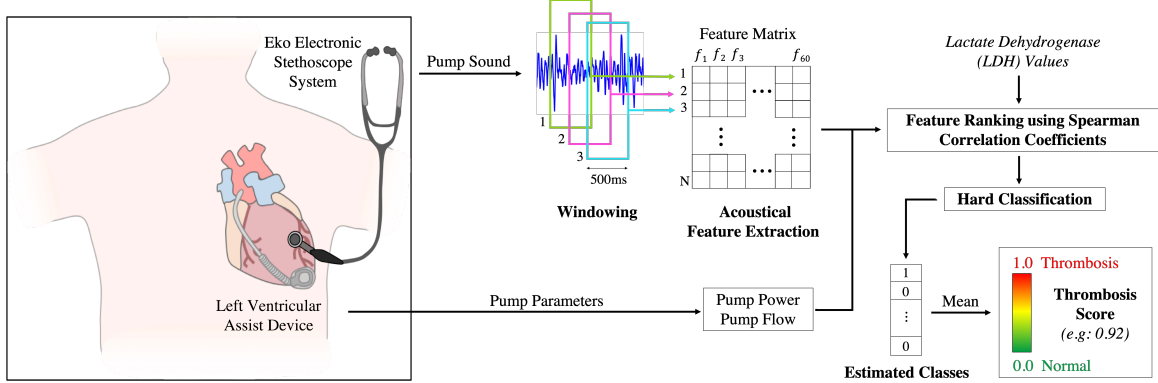


Figure 11: System Overview. The digital stethoscope was placed on the patient's exposed chest over the mitral valve and recordings were obtained in cases of changing pump speed or clinical condition. The acquired recordings were windowed, and 60 acoustic features were determined using LDH-based correlation analysis. It was hypothesized that the features which have higher correlation with LDH values would perform better in thrombosis detection. After calculating the correlation coefficients for each LDH-feature pair, the selected features were fed to a logistic regression classifier which was validated using LOSO-CV. Each window was assigned an estimated class and the overall thrombosis score for any given recording was defined as the mean of these estimated values.

Patients were followed longitudinally for the outcome of suspected pump thrombosis, defined as abnormal pump parameters (power elevations above manufacturer's baseline or an absolute power ≥ 10 Watts), hemolysis (LDH > 500 IU/L, pFhb > 20 mg/dL), development of unexplained HF or clinical hemolysis (hemoglobinuria), or poor left ventricular unloading during an echocardiographic ramp study. Due to the enhanced surveillance of the patients treated at UCSF, suspected pump thrombosis was determined based on abnormal pump parameters and hemolysis. In patients hospitalized for suspected pump thrombosis, clinical data and sound recordings were obtained at admission, prior to and after administration of thrombolytic or anticoagulation therapy, and every 24 hours until laboratory and pump parameters normalized. Thrombus resolution was defined as normalization of pump power and LDH with no clinical evidence of hemolysis for at least 24 hours.

Table 1: Subject Demographics, Pump Parameters and Blood Biomarkers

SUBJECT DEMOGRAPHICS	
Age, y	57.4 ± 11.3
Male, %	84.6
Weight, kg	86.7 ± 20.1
Height, cm	179.1 ± 10.5
Number of Recordings per Subject	3.2 ± 4.2
BASELINE BLOOD AND PUMP MARKERS	
Speed, rpm	2738.8 ± 169.1
Flow, L/min	4.4 ± 1.0
Power, watts	4.6 ± 0.9
Pulsatility, L	3.5 ± 1.1
Systolic Blood Pressure, mmHg	98.4 ± 9.7
Diastolic Blood Pressure, mmHg	69.9 ± 7.6
Mean Arterial Pressure, mmHg	79.3 ± 7.4
Heart rate, bpm	79.2 ± 13.3
Hemoglobin, g/dL	12.1 ± 2.4
Hematocrit, %	36.0 ± 6.7
Partial Thromboplastin Time, sec	38.6 ± 15.0
INR	2.3 ± 0.7

4.1.2.2 Study Design and Data Acquisition

All sounds during the study period were collected and stored using the Eko Electronic Stethoscope System (Eko Devices, Inc. Berkeley, CA). With the patient in the

supine position in a quiet room, the diaphragm of the stethoscope was placed on the patient's exposed chest over the mitral valve for 15 seconds (Fig. 11). The stethoscope settings were as follows: 40x gain, 4kHz sampling rate, 20Hz - 2kHz bandwidth.

Table 2: Pump Parameters and Blood Biomarkers for Baseline and Thrombosis Recordings

	Baseline	Thrombosis
Lactate Dehydrogenase (LDH), IU/L	291.4 \pm 80.4	866.7 \pm 382.9
Plasma Free Hemoglobin (pfHb), mg/dL	80%, < 8	50%, < 8
	20%, 12.0 \pm 5.6	50%, 133.4 \pm 64.8
INR	2.2 \pm 0.6	2.3 \pm 0.8
Abnormal power by device manufacturer standards, %	N/A	81.8
Treatment with both tissue plasminogen activator and IV heparin, %	N/A	27.0
Treatment with IV heparin exclusively, %	N/A	54.5

Each recording was assigned to one of three groups for analysis: “normal” for the baseline recording obtained at enrolment, “thrombosis” for recordings obtained when a patient met the above criteria for suspected thrombosis, and “post-thrombosis” for recordings obtained after initiation of medical treatment until normalization of pump parameters. For the post-thrombosis analysis, the recordings which were taken when the pump parameters and biomarkers returned to normal were used. The procedure by which recordings were processed and analyzed was as follows: (1) windowing and feature extraction; (2) feature selection using correlation analysis, (3) threshold selection and thrombosis score calculation.

Table 3: Number of Recordings for Each Subject

Sub1	1 baseline
Sub2	1 baseline, 3 thrombosis. 2 post-thrombolysis
Sub3	1 baseline
Sub4	3 baseline
Sub5	4 baseline, 3 thrombosis, 8 post-thrombolysis
Sub6	4 thrombosis, 3 post-thrombolysis
Sub7	1 baseline
Sub8	1 baseline
Sub9	1 thrombosis, 1 post-thrombolysis
Sub10	1 baseline
Sub11	1 baseline
Sub12	1 baseline
Sub13	1 baseline
TOTAL	41 recordings (16 baseline, 11 thrombosis, 14 post-thrombolysis)

4.1.2.3 Pre-processing and Feature Extraction

In this study, there were 41 recordings from 13 subjects (16 normal (baseline), 11 thrombosis and 14 post-thrombolysis). The average number of days post-implantation the recordings were obtained was 372, ranging between 11 and 1374. Information about subject demographics, pump parameters and blood biomarkers can be found in Tables 1 and 2. Detailed information about the recording types is presented in Table 3. In this work, the focus was on the generalization across recordings rather than generalization across

subjects as it is more important to demonstrate whether the method is consistent regardless of when the recording is performed for a particular patient. Thus, the signals were segmented into 500 ms frames with 90% overlap to increase the number of instances. From each frame, 60 acoustic features were extracted. In the recent studies, the best indicators for detecting various pump states were shown to be the spectral content of the signals, as these pump states may have different frequency characteristics [101, 104, 112]. Contrarily, it was hypothesized that having a more diverse feature set would enhance the classifier performance, thus various different features were explored [10, 113]. The feature set included the common audio-processing features and features used in previous LVAD studies [101, 104, 110]. The features were categorized as “temporal”, “spectral”, “MFCC”, “bandpower”, “wavelet”, “non-linearity” or “harmonic” features. Detailed information about these features can be found in Table 4. Extracted features were stored in a matrix, where each row represents a single frame and the columns represent the 60 features extracted. In addition to the 60 acoustical features, also the reference pump parameters (pump power and pump flow) were tested as features, which resulted in 62 features in total (Fig.11).

4.1.2.4 Feature Selection using Spearman Correlation Analysis

In the previous section, the extracted features were presented; however as in many other cases, these features are selected ad hoc with a black box approach. Therefore, the redundant features were eliminated while keeping the informative ones; because although having more features can give extra power to ML algorithms, focusing on the intrinsic dimension is usually more desirable for preventing overfitting and ensuring robustness [114-116]. In addition, a mapping between blood-derived biomarkers and acoustical

features should be quantified for evaluating the pathological processes or tracking therapeutic interventions accurately outside of the clinic, as the ideal scenario is recording and analyzing the pump sounds of the subjects outside the clinic in the absence of blood tests or any additional diagnostic tool. Based on this motivation, it was hypothesized that finding the features which are highly correlated with the blood biomarkers may be beneficial to exclude the redundant features and include the most related ones to construct a global feature set, which can be used for all subjects in thrombosis detection without any need for blood work [15].

Table 4: Feature Descriptions of Different Feature Groups

Feature Group	Features
Temporal	Zero Crossing Rate (ZCR), Energy, Energy Entropy
Spectral	Spectral Centroid, Spectral Spread, Spectral Flux, Spectral Entropy and Spectral Roll-off
MFCC	13 Mel-frequency Cepstrum Coefficients
Bandpower	Signal power in 29 distinct frequency bands, between 30 logarithmically spaced frequencies in the range of 20Hz - 2kHz.
Wavelet	3-level Daubechies8 wavelet transform was computed and the bandpower of the 3rd level approximation and detail coefficients and 2nd level detail coefficients were calculated. [113]
Non-linearity	Fundamental frequency, total harmonic distortion (odd harmonics), total harmonic distortion (even harmonics)
Harmonic	First, second, third and fourth harmonic amplitudes

As previously discussed, changes in hemolysis biomarker values ($\text{LDH} > 500 \text{ IU/L}$, $\text{pfHb} > 20 \text{ mg/dL}$) indicate pump thrombosis, therefore it was hypothesized that if the acoustical features which have similar trends with the hemolysis biomarkers could be determined, the classifier might have a higher performance in thrombosis detection. For

this study, LDH values were preferred over pfHb values as the reference, since it would not be possible to use some of the pfHb values (such as pfHb < 8 as in Table 2). Also, recent studies proposed LDH as the primary [90, 117-119] and the most specific [120, 121] indicator of thrombosis in LVADs.

The Spearman correlation analysis [122] was used to investigate the relationship between the LDH values and the acoustical features. Spearman correlation was chosen instead of Pearson correlation as it does not make any assumptions regarding the frequency distribution of the variables or does not require a linear relationship between them. It was desired to treat the features as independent agents, because although some features (such as harmonics) are based on the same fundamental frequency, their behavior may be different during thrombosis development and there is no consensus across different studies [101, 104, 110]. The Spearman correlation coefficient ρ was calculated for each LDH - feature pair (e.g.: LDH and ZCR, LDH and Energy, etc.) which resulted in 62 correlation coefficients varying between -1 and 1. Then the absolute value of these correlation values were taken and ranked in descending order. The cut-off correlation coefficient was determined as 0.3, since any value below 0.3 is considered negligible in the literature, especially in medical research [123-125]. Thus, it was decided to use the features which have absolute correlation coefficient above 0.3 (when rounded to two significant digits). Based on this investigation, 15 features were included in the final feature set. Note that the target labels of the normal and thrombosis groups were not included in this investigation, so the analysis was not biased.

After ranking the features based on their correlation with LDH values, the kernel density estimate (KDE) plots were plotted for the best and worst performing features to

visualize the differences in their probability density functions [126]. It was expected to see distinguishable density functions between normal and thrombosis classes if a feature has high correlation with LDH biomarker. After determining the top 15 features, the analysis continued with threshold selection and thrombosis score calculation.

4.1.2.5 Threshold Selection and Thrombosis Score Calculation

The end goal was determined as deriving a thrombosis score to place the recordings on a scale ranging from 0 (normal) to 1 (thrombosis). To calculate the thrombosis score, the classifier was trained twice: using top *15 features* and all *62 features*, to test the effectiveness of the correlation-based feature ranking. The classifier was validated using LOSO-CV. In each cross-validation fold, all recordings from one subject were left out and the logistic regression classifier was trained using the recordings from the remaining subjects. The model was then tested on the recordings from the subject being left out. This procedure was completed for all subjects. Note that the post-thrombosis data was not used for training, as the ground truth labels for this data were not known certainly - in many cases, thrombosis can be recurrent even following treatment.

Logistic regression was the preferred method due to several reasons. First, it is easy to implement and interpret - for example it does not require heavy scaling or tuning. This was the primary motivation, as an application like thrombosis detection is a collaborative work between medical doctors, researchers and patients; therefore, comprehensibility is an important consideration. Secondly, it is widely used in biostatistical applications where binary responses occur quite frequently, such as patients having heart disease or not, etc. [126]. Therefore, for a case like thrombosis detection, it is well suited.

In more detail, let \mathbf{R} denote the training data. Each instance within our training set corresponds to a 500ms-long frame in the form (\mathbf{x}_i, y_i) , where $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,62})$ are the 62 features, and y_i is the thrombosis label (0 or 1). The training dataset \mathbf{R} is used to build the logistic regression classifier M which predicts the class of an incoming frame using its features. i.e., $M: \mathbf{x} \rightarrow \text{predicted thrombosis label}$. In each fold, \mathbf{R} includes data from all subjects, except the subject whose recordings are being tested. At prediction time, 302 unlabeled frames from each recording are provided to M , and M predicts their labels as 0 or 1 (hard classification). This prediction phase was completed for all recordings of the subject which had been left out in the training phase. In the following folds, the same pipeline was followed for all subjects. The *thrombosis score* of each recording was defined as the mean of the predicted classes of all 302 frames within that recording.

Using the thrombosis scores obtained from LOSO-CV and the actual class labels, the ROC curve was plotted to investigate how two models (with 15 features and with all features) behave for different threshold values. After showing the superior performance of the condensed feature set, the best threshold value for the current application was determined using the ROC curve plotted. Since ROC represents the probability of detection (true positive rate) vs. probability of false alarm (false positive rate) for different threshold values, a threshold was selected such that the probability of thrombosis detection was maximized while the probability of false alarm was minimized. Therefore, the optimum threshold was selected for the classification problem, which helped to discard the sub-optimal models. The generalizability of the classifier was then assessed by calculating the accuracy, sensitivity, and precision values based on the threshold selected. In the final part of the scoring procedure, the constructed model using the baseline and thrombosis

recordings was fit into the post-thrombolysis data to investigate the pump conditions after heparin and/or tissue plasminogen activator (tPA) treatments.

4.1.2.6 Performance Comparison with Existing State-of-the-Art Methods

To benchmark the performance of the proposed approach - which integrates multiple acoustic features using machine learning approaches rather than focusing on one or two individual features - the reference parameters (pump features obtained from the controller) and state-of-the-art algorithms from the existing literature were employed, and the results of these approaches were compared to the proposed work. First, the classification was performed using the reference parameters: only the pump power and only the pump flow values as the features. As these parameters are directly associated with abnormal pump functionality, the performance metrics of these reference variables and the proposed method were compared.

Second, Kaufmann *et al.* [101] asserted that the most intense harmonic is the fourth harmonic in centrifugal pumps; therefore based on their methods, the intensities of the first four harmonics were calculated for each recording. They suggested that an increase in the first and second normalized harmonic intensity and the existence of the third harmonic suggest pump thrombosis. The effectiveness of their features was investigated both with and without a machine learning approach - first the values of the harmonic intensities for normal and thrombosis recordings were compared, then the classifier performance, which was trained using only these normalized intensities as features, was reported.

4.1.3 *Results and Discussion*

4.1.3.1 Feature Selection using Spearman Correlation Analysis

As explained in Section 4.1.2.4, the Spearman correlation coefficient was calculated for each LDH-feature pair and these correlation values were used as a feature ranking metric. The hypothesis was that the features which have higher correlation with LDH values would perform better in thrombosis detection. Also, the redundant features were eliminated for the sake of robustness and simplicity. The 15 features having the highest absolute correlation coefficient are reported in Fig. 12(a). Each color represents a different feature group (e.g.: green: temporal, orange: spectral, etc.) As seen in the bar graph, there is a great diversity in the colors, thus each feature group has a relationship with the LDH values. Also highly correlated features do not come from a single group, which makes the investigation of a diverse feature set rather than relying solely on one feature group (like harmonics) even more desirable.

To test whether the highly correlated features are indeed better in distinguishing between normal and thrombosis groups, the KDE plots were plotted and the differences in distributions were observed (Fig. 12(b-d)). The distributions of the top feature (pump flow, $\rho=0.84$), the top acoustical feature (bandpower 16, ~ 195 -230 Hz, $\rho =0.48$) and the least correlated feature (bandpower 21, ~ 440 -520 Hz, $\rho=0.012$) were plotted. In the graphs, the thrombosis recordings were plotted with red and normal recordings with green. If a specific feature has the ability to distinguish between two groups, it was expected to see less overlapping distributions. As expected, it is visually clear that the features having higher correlations with LDH have less overlapping distributions compared to the ones having lower correlation. Thus, it was justified that if a feature behaves similar to the actual blood

biomarker, it will be more significant in thrombosis detection; therefore a mapping between our “acoustical biomarkers” and actual “blood biomarkers” is indeed achievable.

4.1.3.2 Threshold Selection and Thrombosis Score Calculation

As explained in Section 4.1.2.5, the algorithm was run twice: 1) Using *the top 15 features*, 2) Using *all 62 features*. The ROC curves were plotted for both experiments to investigate whether the correlation-based feature selection algorithm helps to increase classifier performance and eliminate the redundant features independent of the threshold used. When all 62 features were used, the area under the ROC curve was calculated to be 0.764, whereas for the top 15 features, this value was 0.921. This justified that the Spearman correlation coefficients between the features and LDH values can be used to rank and select the most prominent features. Also, another important point is that having more features is not always beneficial for the model and redundant features should be eliminated. A similar approach can be applied to many other biosignals to focus on the most relevant features while eliminating the unnecessary and deviating ones for a more robust analysis. Based on the ROC curve plotted using the 15 features in Fig.12(e), the threshold was determined as 0.15, since it was the best value which maximizes the TPR (probability of detection) while minimizing the FPR (probability of false alarm). Using this threshold, the accuracy, sensitivity and precision were calculated to be 88.9%, 90.9% and 83.3%, respectively.

Additionally, the constructed model using the baseline and thrombosis recordings was fit into the post-thrombolysis data to investigate the pump conditions after heparin and/or tPA treatments. The model was tested on 14 post-thrombolysis recordings taken

from 4 HVAD subjects; however, although the pump parameters and hemolysis biomarkers returned to normal, the scores were varying from 0 to 1, rather than being closer to 0, with 4 of the 14 recordings being above the thrombosis threshold and 10 being below the threshold (Fig.13).

The results from the post-thrombolysis recordings from 4 subjects revealed that, while pump power and hemolysis biomarkers may indicate that the thrombosis is resolved, there may still be other problems such as altered pump mechanics, increased friction, or even residual pump thrombosis that cannot be identified by these markers. Indeed, for 3 of the 4 subjects, recurrent thrombosis episodes were observed following the treatment of thrombosis, and the 4 recordings which had scores above the threshold belonged to these 3 subjects (Subjects A, B, D in Fig. 13). When the clinical logs of these 3 subjects were investigated further, it was found out that these subjects had indeed been diagnosed with recurrent thrombosis right after the recordings having higher scores. The ability to predict recurrent thrombosis by supplementing pump power and blood biomarker data with acoustical signatures measured from the pump should be evaluated extensively in future studies. Note that no direct confirmation of pump thrombosis was available in the study as no devices were exchanged.

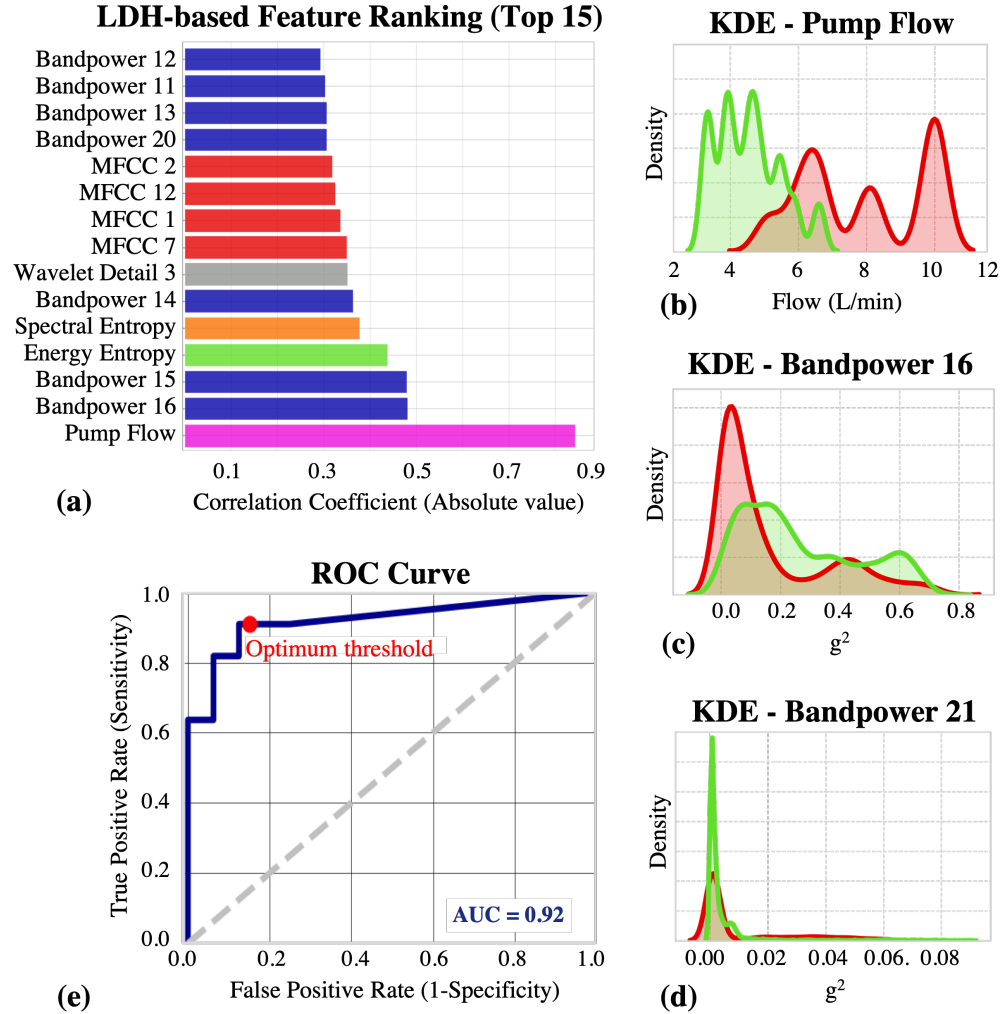


Figure 12: Results. (a) The 15 features having the highest absolute correlation with LDH are presented. There is a great diversity in the feature types, thus each feature group has a relationship with the LDH values. (b) KDE plot for the most correlated feature, pump flow. (c) KDE plot for the most correlated acoustic feature, bandpower 16 (190 - 230 Hz band). (d) KDE plot for the least correlated feature, bandpower 21 (440 - 520 Hz band). (e) ROC curve when the classifier was trained using the top 15 features. The area under the curve (AUC) was calculated to be 0.92. The decision threshold was determined to be 0.15 as it is the optimum value for maximizing the TPR (~ 0.91) while minimizing the FPR (~ 0.13).

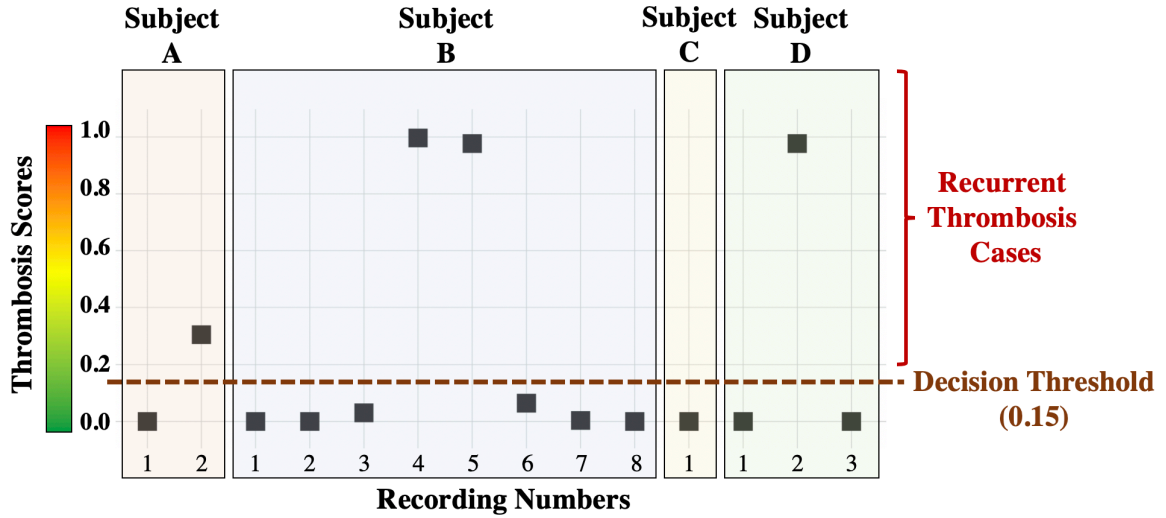


Figure 13: The constructed model using the baseline and thrombosis recordings was fit into the 14 post-thrombolysis recordings taken from 4 subjects to investigate the pump conditions after treatment. The results revealed that, while pump power and hemolysis biomarkers may indicate that the thrombosis is resolved, there may still be other problems such as altered pump mechanics, increased friction, or even residual pump thrombosis that cannot be identified by these markers. Indeed, recurrent thrombosis episodes were observed in 3 of the 4 subjects following the treatment of thrombosis, and the 4 recordings which had scores above the threshold belonged to these 3 subjects.

Pump thrombosis decision is currently based on the presence of hemolysis and degradation of device performance, which is further investigated using echocardiogram ramp studies [95, 96]. In these studies, left ventricle end-diastolic diameter is monitored in response to changing LVAD speed, which may potentially diagnose pump thrombosis or other obstructions within the pump [127]. However, there is a need for a convenient remote monitoring system for LVAD recipients for early detection of thrombus formation, therefore patients should be able to monitor their LVAD status at home continuously. In particular, the patients at home could use the microphone function of their smartphones and transmit these sounds electronically to the cloud for analysis. Machine learning algorithms could then be employed to identify abnormal recordings from baseline, initiating a warning for patients and caregivers before thrombosis exacerbation and

determining the need for clinical pump or blood tests. This way, early diagnosis and optimum treatment recommendations could be achieved, and this could help preventing HF aggravation and additional surgeries.

4.1.3.3 Performance Comparison with Existing State-of-the-Art Methods

Using only pump power values, accuracy, sensitivity and precision values were 62.9%, 27.3% and 60.0%, respectively. Using only pump flow values, the corresponding values were 81.4%, 72.7% and 80.0%. This demonstrates that mechanical power or blood flow characteristics are not sensitive nor specific enough to detect the suspected pump thrombosis episodes by themselves. On the other hand, flow showed higher performance in thrombosis detection compared to power probably due to frequently adapted hematocrit settings, which affected the flow calculation but not power consumption. Thus, it can be said that although one of the most prominent clinical signs for thrombosis development is elevated pump power, the combination of acoustic features and pump flow values yields higher thrombosis detection rate and accuracy as shown in Table 5.

Table 5: Performance Comparison with State-of-the-Art Methods

	Accuracy	Sensitivity	Precision
Pump power	62.9 %	27.3 %	60.0 %
Pump flow	81.4 %	72.7 %	80.0 %
Harmonic analysis (Kaufman <i>et al.</i>) [101]	74.1 %	63.6 %	70.0 %
Our method	88.9 %	90.9 %	83.3 %

Furthermore, when the approach outlined by Kaufmann *et al.* [101] was applied on the current dataset, the fourth harmonic was the most intense harmonic in only 17 of 41 recordings. Nevertheless, the first, second and third harmonic intensities were normalized to the fourth harmonic intensity as the authors explained. It was found that the average intensity for the normalized first harmonic to be 2.46 ± 3.52 and 3.06 ± 2.43 for the control and thrombosis groups, respectively. For the second harmonic, these values were 0.43 ± 0.57 and 2.01 ± 2.70 for the control and thrombosis groups, respectively. Indeed, these intensities increase as the pump starts to develop thrombosis as seen in previous work; however, high inter-recording variability resulted in lower than expected performance for classifying thrombosis. As the next step, the model was evaluated using the harmonic features and all classification results were reported in Table 5. However, it is important to note that Kaufmann *et al.* averaged 30 spectra in their work, whereas in this comparison the harmonics were computed from single 500 ms frames, which is likely noisier. Thus, this should not be considered as a direct comparison.

4.1.3.4 Limitations and Future Work

Currently, there is no clinical gold standard existing for certain comparisons. Therefore, the findings in this work may serve as the preliminary evidence for developing an extensive clinical study in the future. This study was limited by a relatively small sample size, and the study was conducted at a single clinical site; accordingly, future studies will be needed to assess the reproducibility of the methods. Moreover, as the number of subjects in the training set is increased, the generalizability and prospective validation of the proposed algorithm will be improved as will the determination of key acoustical features

associated with pump thrombosis. In addition, although the proposed algorithm may eventually help to decrease the frequency of hospital visits and blood tests, recording the sound signals and processing them requires additional time and computational power, compared to bedside testing.

While this work leverages hard classification using logistic regression due to their high interpretability in the medical domain, the integration of soft-classification algorithms and calibration curve implementation into the current study could be investigated in future work. Similarly, the effect of recording length on feature sensitivity (e.g. sensitivity to harmonics) should be investigated in future studies. In this work, the bandpower features were calculated within fixed frequency bands. In future work, the possibility of obtaining higher accuracy if the frequency bands were adapted based on the pump speed will be investigated. Lastly, LDH was used as the reference blood-biomarker to employ correlation-based feature ranking and any other condition that may have increased LDH was excluded. In the future studies, the correlation between the acoustic features and other blood-biomarkers will be investigated as well.

4.1.4 Conclusion

It was demonstrated that the acoustical signatures of LVADs, combined with machine learning algorithms, can improve the detection accuracy of suspected thrombosis outcomes. Additionally, the scoring of post-thrombolysis recordings suggested residual pump thrombosis which is not identified by pump parameters or hemolysis markers. Additionally, patients who have normal pump operations but abnormal levels of hemolysis markers could benefit from acoustic analysis to detect acoustic signs consistent with

thrombosis. As there is a need for a convenient remote monitoring system for VAD recipients, patients should be able to monitor their VAD status at home. Once the proposed algorithm is evaluated prospectively and demonstrates improvement in clinical management of pump thrombosis, the use of acoustical patterns can potentially enable detection of suspected pump thrombosis when biomarkers and pump parameters are non-diagnostic.

4.2 Non-Invasive Wearable Patch Utilizing Seismocardiography for Peri-Operative Use in Surgical Patients

The previous section investigated the use of acoustic signatures to derive clinically useful digital biomarkers for thrombosis detection in LVADs. This section presents a novel framework for non-invasive SV estimation through wearable SCG and PCG measurements taken from the sternum.

4.2.1 Introduction

Peri-operative fluid management has been shown to impact clinical outcomes including reducing the duration of hospital and/or intensive care unit (ICU) stays. Achieving and maintaining optimum fluid status is clinically challenging, especially since physiologic alterations associated with surgery may complicate interpretation of routine parameters. Clinicians often rely on additional hemodynamic parameters, such as stroke volume (SV), to guide peri-operative fluid management [128-135]. Several studies have demonstrated the advantages of fluid management by administering crystalloids or colloids to maintain SV within 10% of baseline [136-138].

In the last decade, hemodynamic monitoring at the bedside has been evolved remarkably both in the operating room and ICU. The most significant progression has been the declining use of the pulmonary artery catheter and increasing adoption of continuous, real time and less invasive monitoring techniques [139]. Today, a commonly used tool for SV estimation in the operating room is the Transesophageal Doppler (TED) which measures blood flow velocity in the descending thoracic aorta [136-138]. Although the TED is generally considered to be sufficiently accurate for fluid management intra-

operatively, it has some disadvantages. It requires general anesthesia or deep sedation and therefore must be removed on emergence. In addition, it requires periodic repositioning and constant supervision by an experienced professional. Surgical patients often experience significant alterations in volume status post-operatively, however the TED is not a valid monitor in those situations for these reasons [128, 140]. There is therefore a clinical need for a non-invasive alternative to the TED, that could be used for monitoring SV in patients during the entire peri-operative period.

Although multiple technologies for non-invasive SV prediction have been described in the literature, none is suitable for use in the pre-, intra- and post-operative settings. Impedance cardiography (ICG) involves placing 4-8 electrodes on the thorax, sending a small electrical current into the body and measuring the subsequent voltage drop to compute thoracic impedance changes resulting from aortic ejection of blood. While ICG has been demonstrated to be accurate for SV estimation in some settings, the accuracy is limited in patients with low cardiac output, subjects with high body mass index, and can be confounded by other fluids in the thoracic cavity (e.g., edema) [141-144]. Volume-clamping based continuous blood pressure measurement technologies have also been implemented with empirical algorithms for estimating SV, however their accuracy has typically been shown to be low in large studies [145-148]. Ballistocardiography (BCG) based techniques, quantifying whole-body movements in response to the blood movement in the vascular tree, have been investigated as well for SV estimation; however, such techniques require cumbersome beds or tables and are thus not well-suited to clinical settings and cannot be applied readily to ambulatory settings [21, 149].

In this study, the primary aim was to develop an alternative to TED, which can track a patient's SV continuously and non-invasively both during and after surgery. Previously a wearable patch was developed [150], and studied for ambulatory monitoring in heart failure patients. It was found that it accurately reflects the changes in clinical status [32, 151]. In this proof of concept study, the feasibility of estimating SV with wearable patch that measures the seismocardiogram (SCG) and electrocardiogram (ECG) signals in patients undergoing major gynecologic oncology surgery was investigated. To achieve this, an end-to-end signal processing and regression-based prediction algorithm using the features extracted from the patch signals was presented. The TED was used as a reference standard against which the estimates from the wearable patch could be compared. With such a mapping function, the wearable patch is expected to enable non-invasive and continuous hemodynamic monitoring and guide fluid management throughout the entire peri-operative phase.

4.2.2 Study Design and Subject Demographics

This project was conducted under a protocol approved by the Northwestern University Institutional Review Board and all subjects provided written consent. A total of 12 female subjects (Age: 52.0 ± 10.1 , Weight: $79.2 \pm 23.2\text{kg}$, Height: $162.6 \pm 5.4\text{ cm}$) undergoing major intra-abdominal gynecological oncology surgery participated in the study. The peri-operative care of all subjects was based on the Institutional Enhanced Recovery (ERAS) Protocol for Gynecologic Oncology patients. The protocol included pre-operative hydration and multimodal analgesia. The wearable patch was applied pre-operatively, to the anterior chest of the subject at the mid-sternal level. All patients received an opioid sparing total intravenous anesthetic. The TED was inserted following tracheal

intubation. The intravenous fluid therapy protocol was based on the ERAS protocol, guided by the Doppler parameters to maintain an SV greater than 70 mL/beat. Upon emergence from anesthesia, the endotracheal tube and the TED were removed in all subjects. Additionally, post-operative analgesia, intravenous fluid therapy, oral intake and physical activity were based on the ERAS protocol in all subjects.

4.2.2.1 Transesophageal Doppler (TED)

Following endotracheal intubation, the TED (CardioQ-ODM, Deltex Medical, Greenville, SC) was inserted into the esophagus of the subject by an experienced practitioner. Once the appropriate Doppler signal and waveform were obtained, monitoring commenced [136, 152]. Windows (30 second length) of the resulting velocity-time waveform were used to calculate various parameters such as mean acceleration, peak velocity, and stroke distance. SV was calculated from the TED waveform by multiplying the distance a column of blood moves along the aorta at each left ventricular contraction (stroke distance) with a calibration constant (the approximate area of the aorta based on patient age, weight and height).

4.2.2.2 Wearable Patch

The custom wearable patch, which is shown in Fig. 14(a), was applied to the anterior chest of the subject at the mid-sternal level using three Ag/AgCl ECG adhesive-backed gel electrodes (Red Dot 2560, 3M, Maplewood, MN). It weighs ~ 65 grams and has a diameter of ~10 cm, which does not make the subjects uncomfortable during data collection. To capture the ECG, an analog-front-end (AFE) integrated circuit (IC) with on board analog-to-digital converter, ADC, (ADS1292 Texas Instruments, Dallas, TX), which

connects to these three electrodes, was used. Each electrode corresponds to either the positive reference, negative reference, or right-leg drive.

While the ECG waveform measures the *electrical* activity of the heart, the SCG waveform assesses the *mechanical* motions and corresponds to the local thoracic vibrations originating from the contraction of the heart and ejection of blood from the ventricles. An SCG waveform is generated for each contraction, which is characterized by several peaks and valleys. A tri-axial ultralow noise accelerometer (ADXL355, Analog Devices, Norwood, MA) is used to capture the SCG signals in lateral (X-axis), head-to-foot (Y-axis) and dorso-ventral (Z-axis) directions. SCG components each having a specific morphology, can be observed in all three axes as shown in Fig. 14(b), [21, 153].

The patch includes an ATSAM4LS8B (Microchip Technology, Chandler, AZ) to sample all sensors (ECG at 1kHz and the accelerometer at 500 Hz) and to save the data to an on-board microSD card. When interfaced with a computer through a microUSB port, HeartPulse App (Department of Anesthesiology, Northwestern University, Chicago, IL) communicates with the microcontroller to download and subsequently delete the data on the microSD card. Additionally, the inserted microUSB interfaces with the battery charger (BQ24232RGTR, Texas Instruments) to charge the 150 mAh battery. The battery life of the patch is approximately 48 hours, which is more than adequate to record the data for the entirety of most surgeries and the first post-operative day.

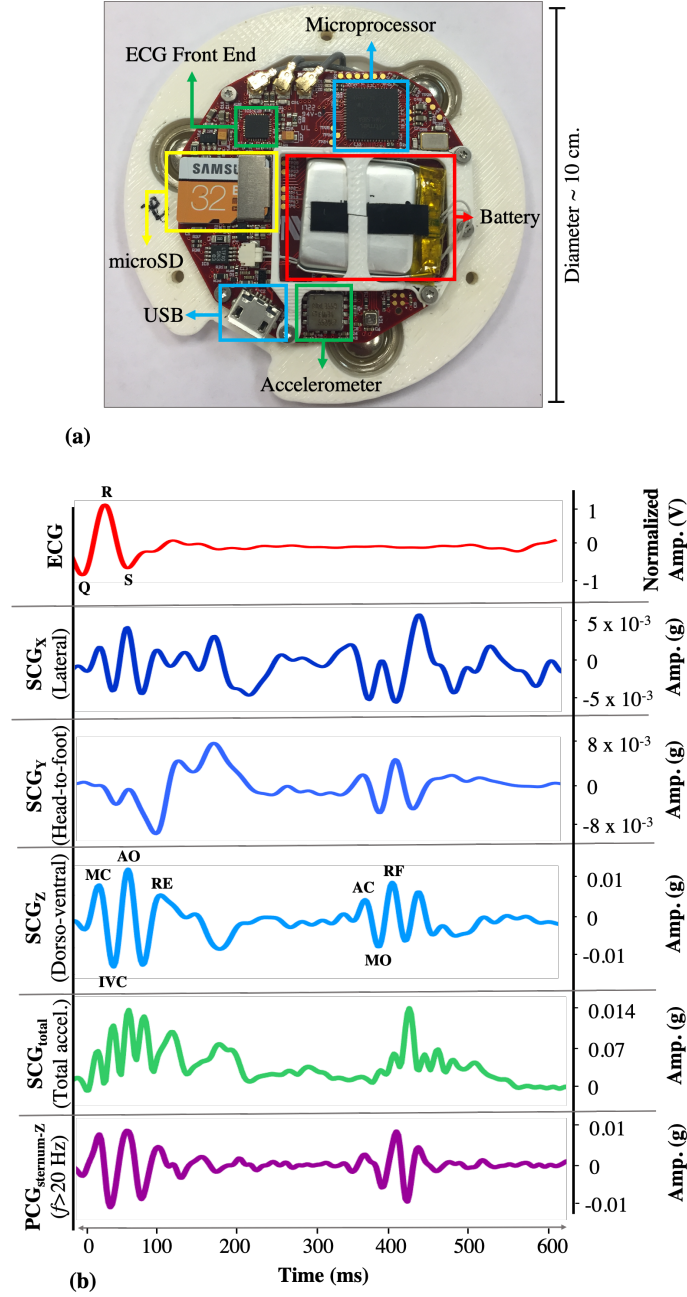


Figure 14: (a) Custom-made patch hardware. (b) A representative beat from the ECG, $SCG_{x,y,z,total}$ and $PCG_{sternum}$ signals (for PCG, only Z-axis is shown as the other axes are less commonly used). The portion within each R-R interval is called a *beat*, which is usually around $\sim 600 - 1000$ ms. The R-peak locations on the ECG signal are used to split the SCG into individual beats. The peaks and valleys on the SCG signals correspond to specific cardiac events.

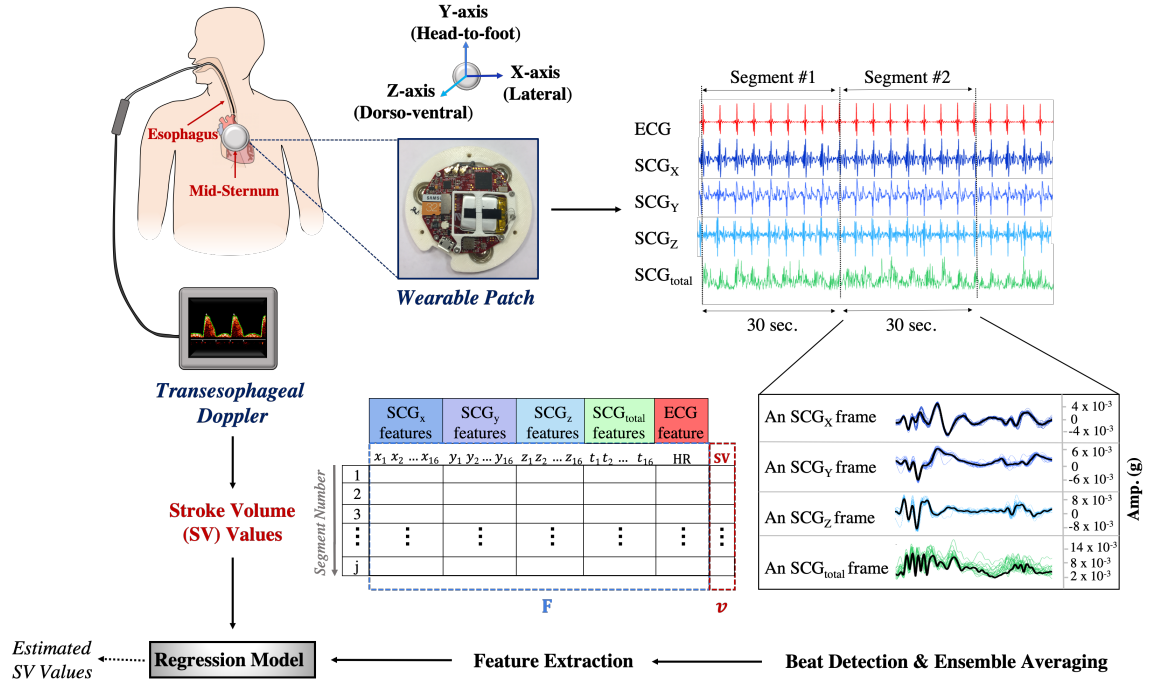


Figure 15: The block diagram for the data acquisition and signal processing. The wearable patch was mounted on the mid-sternum while the patient was undergoing a surgery. The ECG and SCG_{x,y,z} signals were recorded throughout the surgery while the SV values were being acquired with the TED. In addition to the SCG_{x,y,z} signals, the total acceleration magnitude, SCG_{total}, was calculated. As the TED was averaging and storing the SV values every 30 seconds, the ECG and SCG_{x,y,z,total} signals were split into 30 seconds-long segments. Within each segment, the R-peak locations were found on the ECG and used to detect the individual beats on the SCG_{x,y,z,total}. The ensemble average of the detected beats was calculated, so per segment there was one single ensemble averaged frame for each axis. In total 65 features were extracted for the analysis. Then a random forest regression model was trained to predict the corresponding SV values.

4.2.2.3 Pre-processing and Signal Segmentation

The SCG_{x,y,z} and ECG signals were filtered using finite impulse response (FIR) Kaiser-window band-pass filters (1-45 Hz and 0.5-40 Hz, respectively) to reduce the motion artifacts while preserving the shape of the signals. These cut-off frequencies were determined based on the existing literature [154, 155]. With these cut-off values, low

frequency noise caused by respiratory chest wall movements was also attenuated as these movements are usually observed below 0.5 Hz [156, 157]. In addition to the $SCG_{x,y,z}$ signals, the total acceleration magnitude, SCG_{total} , was calculated using the formula $\sqrt{SCG_x(t)^2 + SCG_y(t)^2 + SCG_z(t)^2}$, where $SCG_x(t)$, $SCG_y(t)$ and $SCG_z(t)$ represent the acceleration measured in the X, Y, Z directions at time instant t (Fig. 14(b)). This new signal represents the total physical activity of the heart by assessing the spatial curve of the displacement vector instead of separating the vector space into individual axis components [155, 158]. The TED and patch clocks were aligned using the corresponding timestamps in the stored data. The heart rate (HR) and SV values were obtained from the TED as the reference values. Since the TED averages and updates the HR and SV every 30 seconds, the $SCG_{x,y,z,total}$ and ECG signals obtained from the wearable patch were processed in a similar manner. The ECG and $SCG_{x,y,z,total}$ signals were split into 30 seconds long *segments* to ensure coherence with the TED measurements (Fig. 15).

The first 20 minutes (data recorded during pre-incision) was considered as *baseline* or *pre-incision* data. For the intra-operative data, the patch signals were segmented into non-overlapping segments in accordance with the TED clock as previously explained (Fig. 15). The average number of intra-operative data points was 228 ± 128 (min:62, max:418). The differences between the number of data points were due to either the varying length of the surgeries or loss in TED/patch signals. For the pre-incision data, a different signal segmentation method was employed to increase the number of instances. As the TED updates the SV and HR values every 30 seconds, there were 40 data points in each of the pre-incision SV and HR vectors. Cubic spline interpolation was applied on these SV and HR vectors in order to artificially increase the number of data points. Cubic spline

interpolation was chosen as it has been found to be an effective and computationally efficient method in the literature [159, 160]. It fits a series of unique cubic polynomials between adjacent pairs of points, so that a smooth and continuous curve is obtained [161]. From the interpolated SV and HR curves, equally spaced data points were obtained to increase the frequency of the SV and HR readings. The interval was selected as 3 seconds, so that the reference SV and HR values could be updated every 3 seconds, rather than every 30 seconds. The pre-incision $SCG_{x,y,z,total}$ and ECG signals were segmented in accordance with these interpolated SV and HR values. A moving 30 seconds long window with 90% overlap (i.e. step size of 3 seconds for coherence with the interpolated SV and HR values) was applied on the $SCG_{x,y,z,total}$ and ECG signals to obtain the segments (Fig. 16).

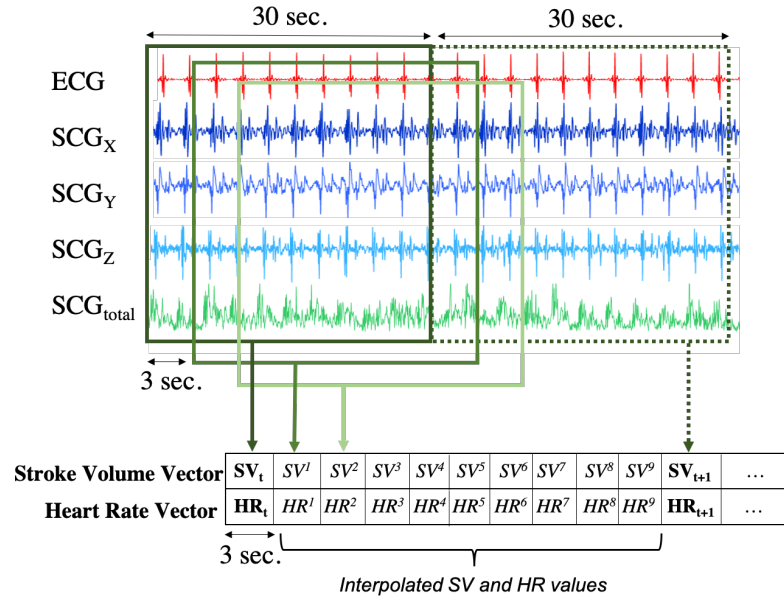


Figure 16: Cubic spline interpolation was performed on the pre-incision SV and HR vectors to increase the number of instances. From the interpolated SV and HR curves, equally spaced (3 seconds) data points were obtained to increase the frequency of the SV and HR readings. Similarly, the patch signals were segmented using a 30 seconds long moving window with 90% overlap (step size = 3 seconds).

4.2.2.4 Feature Extraction

The R-peaks of the ECG were detected using a simple peak detection algorithm within each segment. The peaks were used to extract individual beats of the SCG in each of the three axes. These beats were then truncated to a length of 600 ms and ensemble averaged to produce a single averaged *beat*, or *frame*, per each segment for all three directions [34]. Additionally, for each segment, the average R-R interval duration t_{RR} was calculated using the ratio $60 / t_{RR}$ to obtain the average HR in beats per minute (bpm). These HR values were compared with the reference HR values obtained from the TED, and any 30 seconds-long segment was discarded from the analysis unless the reference and calculated average HR values were in 10% agreement.

Table 6: Features Extracted from the Signals

	Feature Names	Number of Features
SCG_{x,y,z,total} 0-250 ms	First and second largest maxima (amp. and loc.) First and second largest minima (amp. and loc.) RMS power Peak-to-peak amplitude	10 (4 signals=40)
SCG_{x,y,z,total} 250-500 ms	Largest maximum (amp. and loc.) Largest minimum (amp. and loc.) RMS power Peak-to-peak amplitude	6 (4 signals=24)
ECG	Average HR derived from R-R interval length	1
TOTAL		65

SCG signals are characterized by several peaks and valleys, which are labeled in Fig. 14(b) based on their corresponding physiological events: MC (mitral valve closure),

IVC (isovolumetric contraction), AO (aortic valve opening), RE (rapid ejection), AC (aortic valve closure), MO (mitral valve opening), and RF (rapid filling) [21]. In the current analysis, the peaks or valleys were not matched with specific cardiac events as the cardiac events points are subject specific and prone to instability. Such instability may occur because, 1) SCG signals have high inter-subject variability and it is not always possible to detect these points accurately, 2) SCG signal morphology may change depending on the position of the subject or the location of the patch, 3) The current data set consists of subjects undergoing surgery rather than healthy subjects, and thus the subjects' physiological state is not as stable as for healthy subjects. Thus, there was a need to account for instability of the signal morphology in the proposed methodology [41, 153]. Accordingly, rather than putting emphasis on specific physiological points (AO, MC, etc.), the methods in the literature were followed [33, 34] and a new feature set was defined consisting of maxima and minima points of the signal (instead of matching the peaks or valleys with underlying physiological events).

A total of 65 features were extracted for the analysis (Table 6): 64 of these features were from the $SCG_{x,y,z,total}$ frames, i.e. 16 features were extracted from each SCG signal. These features included maxima-minima amplitudes and locations, root-mean-square (RMS) powers, and peak-to-peak amplitudes. As the final feature, calculated average HR values were used. All features were stored in a matrix, and the same feature extraction procedure was repeated for each subject. For n^{th} subject, the feature matrices obtained from the pre-incision and intra-operative data were denoted by \mathbf{B}_n and \mathbf{S}_n , respectively. A sample feature matrix is shown in Fig. 15. In these feature matrices, each row corresponds to a 30 seconds-long segment, and is of the form: (\mathbf{F}_j, v_j) , where $\mathbf{F}_j = (x_{j,1}, x_{j,2}, \dots, y_{j,1}, y_{j,2}, \dots, z_{j,1},$

$z_{j,2}, \dots, t_{j,1}, t_{j,2}, \dots, \text{HR})$ are the 65 features extracted as described in the previous section, and v_j is the SV values we obtained from the TED.

4.2.2.5 Regression Model and Validation

Random forest-based regression was used to estimate SV values as it can achieve high accuracy without overfitting. In this method, rather than relying on a single tree, multiple trees are bootstrapped by using randomized subsets drawn from the original dataset. The trees are trained in parallel independently and their final predictions are averaged to obtain the predicted target value [162, 163]. During the training phase, the model learned the relationship between the reference SV values and extracted features. This trained model was then used to predict the SV values for the incoming unseen segments.

The model was validated using LOSO-CV. The feature matrices obtained from each subject from the pre-incision and intra-operative data as \mathbf{B}_n and \mathbf{S}_n , respectively. In each fold i , the intra-operative data (\mathbf{S}_i) from the i^{th} subject was left out and the model was trained using the remaining \mathbf{S}_n , where $n \neq i$, and all pre-incision data (\mathbf{B}_n). This effectively simulates the clinical scenario of requiring a calibration period with TED \mathbf{B}_i after which time the TED would no longer be required intra- or post-operatively. The combination of these matrices and the corresponding SV values were the training data and training targets, respectively. The model was then tested on \mathbf{S}_n and the predicted SV values were stored. A simplified representation of the proposed LOSO-CV framework is presented in Fig. 17. The same procedure was repeated for each subject.

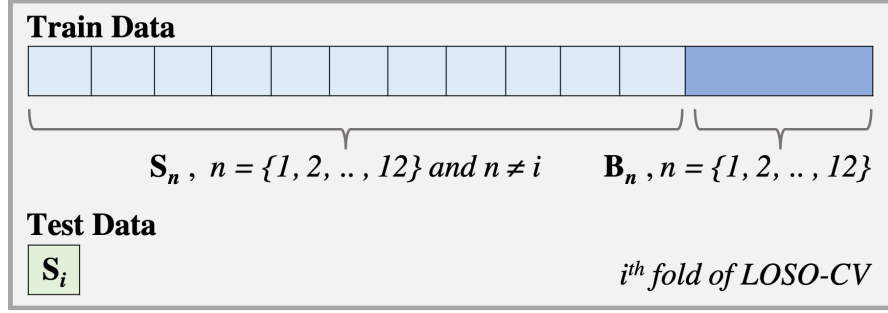


Figure 17: A simplified representation of how LOSO-CV was performed. In each fold, intra-operative data from one subject was left out S_i . The model was trained using the remaining S_n where $n \neq i$ and all B_n . The model was then tested on S_i and the predicted SV values were stored. This step was repeated for each subject.

The correlation coefficient (r-value) and the median absolute error (MedAE) between the reference and predicted SV were computed to assess the performance of the model. Beyond correlation, Bland-Altman methods were used to quantify the agreement between the SCG-based SV values and the reference standard values from TED [164]. Although a Bland-Altman plot presents the bias and limits of agreement, it does not provide direct information on whether the difference between the reference and predicted values is clinically acceptable. The most important reason is inter-subject variability, i.e. one limit of agreement might be acceptable for a subject with high SV, but it may be too high for a subject with low SV [165]. One of the methods to overcome this problem is to calculate the percent error (PE) for each subject, i.e. dividing the limits of agreement by the mean reference SV value [165, 166]. This error should be within $\pm 30\%$ to consider the new method as “interchangeable” with the reference method. It should be noted that in this study, rather than using the limits of agreement across all subjects, each subject’s own limits of agreement and mean reference SV value were used to remove the effect of inter-subject variability. In the MedAE calculation, first the prediction error, i.e. the difference between the reference and predicted SV values of all segments from all subjects, was

calculated. Then the absolute value of the errors was computed, and their median was calculated. This yielded the MedAE of the model. MedAE was selected rather than mean error to reduce sensitivity of the assessment to outliers.

4.2.2.6 Elucidating Mechanistic Insight Regarding Stroke Volume Correlations to SCG

The analysis pipeline was built using the features from both SCG and ECG signals as explained in the previous section. To show that SV prediction would benefit from an electro-mechanical approach rather than relying only on electrical or mechanical schemes, different models were trained using: (i) features from SCG axes (only mechanical), (ii) features from ECG, i.e. HR (only electrical), and (iii) features from SCG axes and ECG together (the electro-mechanical approach presented in the previous section). Then the performance metrics described above (correlation coefficient and MedAE) were calculated.

In addition, the importance of using multiple directions of acceleration (dorso-ventral, lateral, and head-to-foot), and their effect on the regression result were examined to investigate which acceleration direction(s) provide(s) better mechanistic insight into the physiological underpinnings of the signal. Different models were trained using the features derived from each direction (or their combinations) and the correlation coefficient and MedAE value were calculated.

After investigating the effect of acceleration direction, each feature's individual contribution to SV prediction was calculated by obtaining the feature importance scores from the random forest model. In each tree, every node splits the values of a feature based on impurity (variance) criteria, so that the similar feature values are grouped under the

same set. Therefore, feature importance scores can easily be obtained by calculating each feature's contribution to decreasing the weighted impurity [167]. Leveraging this fact, the feature importance scores were obtained from the model and ranked in descending order. While calculating the feature importance scores, two different approaches were employed to investigate the bias and outlier effects on the model: (i) obtaining the feature scores in *each fold* and averaging them across all folds, (ii) re-training the random forest model on the *whole* dataset and computing the feature importance scores.

Finally, to investigate feature contributions further, the Pearson correlation coefficients between feature pairs were calculated. Pearson correlation analysis measures the strength and direction of the relationship between two variables [168]. If any feature pair has an absolute correlation coefficient close to 1 (high correlation), this means that these two features are linearly dependent and have almost the same effect on the dependent variable, in this case: SV. On the other hand, having lower absolute correlation coefficient means that these two features have distinguishable effect on the dependent variable. Thus, having less correlated features, and therefore diversification within a feature set, is actually better for information gain, thus model training.

4.2.2.7 Comparison with State-of-the-Art

Contribution of the Acoustic and Vibration Components

In the literature, it was shown that the SCG signal is composed of two main components: low frequency (i.e. vibration, $1 < f < 20$ Hz) component emerging from the cardiac output and high frequency (i.e. acoustic, $f > 20$ Hz) component originating from the heart sounds [169, 170]. The latter is also known as the phonocardiogram (PCG) waveform

Fig.(14(b)). In this work, in addition to investigating the SCG signal as a whole, the individual contributions of the PCG and vibration components were also investigated. However, it should be noted that the sensor placement was not optimized for PCG and thus it is not quite a fair head to head comparison.

In accordance with the frequency cut-off presented in the literature, the SCG signals were split into PCG and vibration segments: $f > 20$ Hz (PCG_{sternum}) and $1 < f < 20$ Hz (SCG_{vib}). Again, the lower frequency cut-off was selected as 1 Hz to remove any chest movements originated from respiration (generally $f < 0.5$ Hz) [156]. After segmenting the $SCG_{x,y,z,\text{total}}$ signals into PCG and vibration segments, the whole analysis pipeline was run with the same features (*amplitude and location features*) and same ML model (*random forest including 20-min of calibration*) for PCG_{sternum} , SCG_{vib} , and the combination of PCG_{sternum} and SCG_{vib} cases.

Comparison with the Existing Literature

Previously the relationship between SCG and SV was investigated in 8 individuals by recording their uni-axial SCG signals (using 393C, PCB Piezotronics) in two separate sessions at least a day apart [171]. In that study, the following features were extracted from each beat: the time between the ECG R-wave and the opening of mitral valve (t_{MO}), the maximum of rapid ejection (RE_{max}) and the slope of its increase (m_{RE}), the area under curve during rapid systolic ejection (AUC_{RE}), isovolumic contraction time (t_{MC-AO}) and its slope (m_{MC-AO}), systolic ejection time (t_{AO-AC}) and isovolumic relaxation time (t_{AC-MO}). For each subject, a *linear regression model* was trained on the first recording session and the second session was used for testing.

To provide a fair comparison with [171], the points corresponding to the cardiac events (the features above) were extracted from each of the $SCG_{x,y,z,total}$ signals. Then for each subject, the first half of the SCG signal was used to train a linear regression model and the second half was used for testing the model. It should be noted that there were several differences in the experimental setups of this study and the one being compared. In this study, (i) the recordings were taken at a *single session* where the subjects were *undergoing surgery* (i.e. their hemodynamic state was varying), (ii) instead of a uni-axial accelerometer, a *3-axis accelerometer* was used, (iii) the features were extracted from the ensemble averaged frames derived from 30-sec long segments as the TED was updating the SV every 30 seconds.

4.2.3 Results and Discussion

4.2.3.1 Machine Learning Based Estimation of Stroke Volume

The regression model trained on the combination of the SCG and ECG features was correlated to the TED reference standard ($r = 0.81$) with a relatively low MedAE (7.56 mL) across all subjects. Using the cardio-mechanical information only (i.e., SCG signal features alone) resulted in a correlation value of 0.67 and a MedAE of 9.51 mL, whereas using only the electrical (ECG) information resulted in a correlation value of 0.31 and a MedAE of 14.75 mL. Fig. 18(a) provides the regression plot for the electro-mechanical approach, with each subject represented by a different color for ease of visualization.

The results of this study demonstrated that the combination of the SCG and ECG features contains salient information for enabling estimation of SV with sufficiently high correlation and low MedAE across all subjects. The error is higher in SV values greater

than 100 mL, and in particular quite high for SV values above 150 mL. According to the Bland-Altman results (Fig. 18(b)), the differences between the reference and predicted SV values are closer to zero within the 60-100mL range, which justifies that the model achieves higher prediction performance within this range. Most of the data points are within the 95% confidence interval (limits of agreement: +33.08mL and -32.86mL), and the scatter around the bias line (mean difference: 0.11mL) increases as the average SV increases, demonstrating that the variability is not consistent throughout the graph. One reason for the greater error at higher values of SV is the high inter-subject variability in the dataset, as shown with different colors in Fig. 18(a,b); since most of the datapoints at these higher values of SV were from one subject, the model trained on other subjects with lower SV values did not perform as well for this case. In addition to the Bland-Altman plot, PE for each subject was calculated to investigate whether the proposed algorithm could be considered as “interchangeable” with the reference method. This analysis resulted in average PE of 37.5 ± 22.45 for the lower limit (6 subjects being $< 30\%$), and 25.6 ± 16.7 for the upper limit (9 subjects being $< 30\%$).

4.2.3.2 Interpretation of Acceleration Directions and Features

In contrast to the ECG signal which has been studied for more than a century and is well understood regarding its mechanistic underpinnings, the SCG signal is not as well characterized. The literature reports that the dorso-ventral component of the signal represents a combination of body acceleration responses to both heart and blood movement in the thoracic cavity [21]. Our group has previously reported that the head-to-foot component may better reflect blood movement [32]– due to the anatomical alignment of the aorta and the large component of the body acceleration that is originated from the

change of direction of blood through the aortic arch – and thus a sub-hypothesis of this work aimed at elucidating physiological mechanistic insight was that the head-to-foot SCG components would better reflect hemodynamic forces (e.g., stroke volume). In the current study similarly, a relatively higher correlation was obtained between the predicted and reference SV values when the features from the head-to-foot axis were used to train the model ($r=0.65$) compared to the dorso-ventral ($r=0.50$) or lateral ($r=0.22$) axes. In contrast, the dorso-ventral axis, which represents a combination of body acceleration responses to both heart and blood movement in the thoracic cavity, resulted in smaller MedAE (10.86 mL) compared to other axes. When the total heart activity signal was used, the correlation and MedAE values were calculated as 0.56 and 10.49mL, respectively. The regression results of the models trained using features from each SCG axis and their combinations are presented in Table 7.

Each feature's contribution to SV prediction was investigated by analyzing the feature importance scores obtained from the random forest models. Regardless of the method used (averaging across folds vs. re-training on the whole dataset), the same features were obtained as the top 15 most important features. In Table 8, the importance scores from the first method (averaging across all folds) are presented. The feature which has the highest importance is heart rate obtained from the patch ECG. This is followed by the amplitude of the minimum point within the first 250 ms of the dorso-ventral signal. From a physiological perspective, this point corresponds to the start of isovolumetric contraction, which occurs immediately before systolic ejection. Since stroke volume is defined as the difference between the end diastolic and end systolic volumes, finding isovolumetric

contraction as one of the most prominent features is consisted with the underlying physiological events.

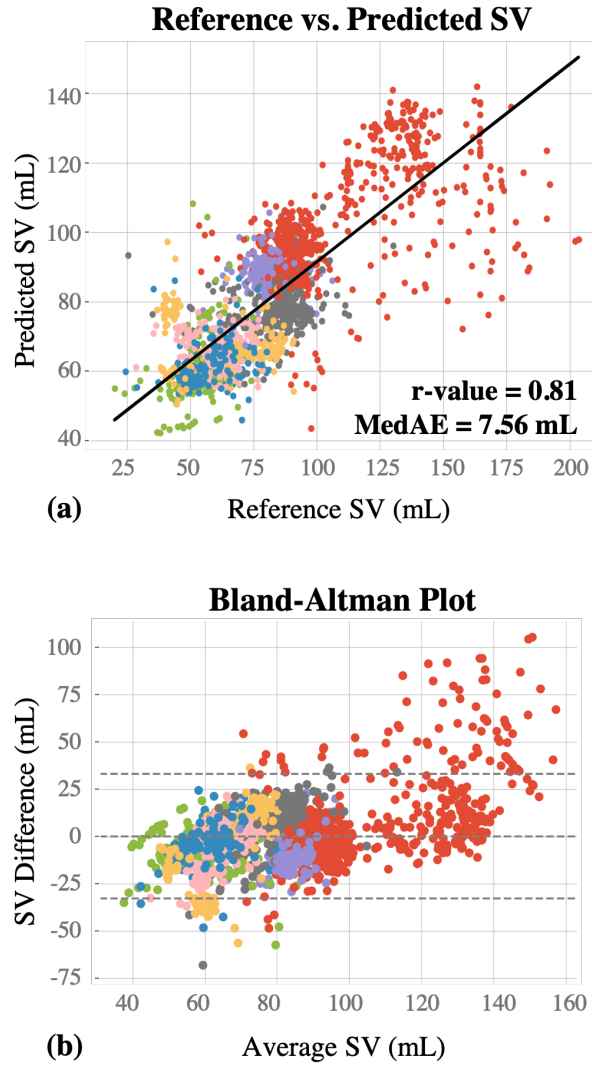


Figure 18: (a) The reference vs. predicted SV values are presented. 65 features were extracted and used to train a random forest regression model. This resulted in a correlation coefficient (r-value) of 0.81 and MedAE of 7.56 mL. (b) The Bland-Altman graph for the reference and predicted SV values. Horizontal line in the middle represents the bias line, whereas the area between the upper and lower horizontal lines represent the 95% confidence interval, (mean \pm 2*standard deviation).

Table 7: Experiments with Axis-Combinations

	Correlation (r-value)	MedAE (mL)
SCG _x	0.22	13.76
SCG _y	0.65	12.46
SCG _z	0.50	10.86
SCG _{total}	0.56	10.49
ECG (HR)	0.31	14.75
SCG _x and ECG	0.58	9.64
SCG _y and ECG	0.74	8.44
SCG _z and ECG	0.67	9.29
SCG _{total} and ECG	0.71	9.65
SCG _{x,y}	0.48	10.81
SCG _{x,z}	0.53	11.80
SCG _{y,z}	0.61	10.94
SCG _{x,y} and ECG	0.77	7.93
SCG _{x,z} and ECG	0.74	8.73
SCG _{y,z} and ECG	0.73	8.86
SCG _{x,y,z}	0.64	10.63
SCG _{x,y,z,total}	0.67	9.51
SCG _{x,y,z} and ECG	0.80	7.61
SCG_{x,y,z,total} and ECG	0.81	7.56

Table 8: Random Forest Feature Importance Ranking

Feature Names	Normalized Feature Scores
Wearable Heart Rate (HR)	0.314
0-250ms, First Min. Amp, Z	0.125
0-250ms, First Max. Amp, X	0.079
0-250ms, Peak-to-Peak, Y	0.050
0-250ms, First Min. Loc, X	0.048
0-250ms, First Min. Loc, Z	0.041
0-250ms, First Min. Loc, Y	0.035
0-250ms, Second Max. Amp, Z	0.026
0-250ms, First Max. Loc, Z	0.024
0-250ms, First Max. Amp, Z	0.020
250-500ms, Peak-to-Peak, Y	0.017
0-250ms, RMS Power, Y	0.014
250-500ms, First Min. Amp, Z	0.013
250-500ms, First Min. Amp, Y	0.013
0-250ms, Second Min. Amp, Y	0.012

Secondly, amplitude values of the maxima and minima points are dominating the location (timing) values of these points. As SV is directly related to the contraction and strength of the heart, it was expected to have mostly the amplitude-related features in the top features. Lastly, most of the top features belong to either the head-to-foot or dorso-

ventral axes. These scores are consisted with the findings of the proposed regression analysis, and explain why dorso-ventral and head-to-foot axes resulted in relatively higher performance compared to the lateral axis. In contrast, none of the total heart activity features appeared in the top 15 list.

Finally, the heatmaps showing the Pearson correlation coefficients between the feature pairs are presented in Fig. 19. Based on these heatmaps, most of the feature pairs have correlation coefficients ranging between -0.25 and 0.50. In the literature, especially in medical research, this range falls under negligible-low correlation group (0-0.3 negligible correlation, 0.3-0.5 low correlation) [123-125]. Therefore, it can be inferred that the features are not linearly dependent, and each feature has a unique contribution to the analysis.

Table 9: Contribution of the Acoustic and Vibration Components ($SCG_{x,y,z,total}$ and ECG)

	Correlation (r-value)	MedAE (mL)
$PCG_{sternum}$	0.69	11.67
$PCG_{sternum} + ECG$	0.75	10.39
SCG_{vib}	0.35	12.67
$SCG_{vib} + ECG$	0.64	10.16
$PCG_{sternum} + SCG_{vib}$	0.76	9.87
$PCG_{sternum} + SCG_{vib} + ECG$	0.81	7.99
THIS WORK	0.81	7.56

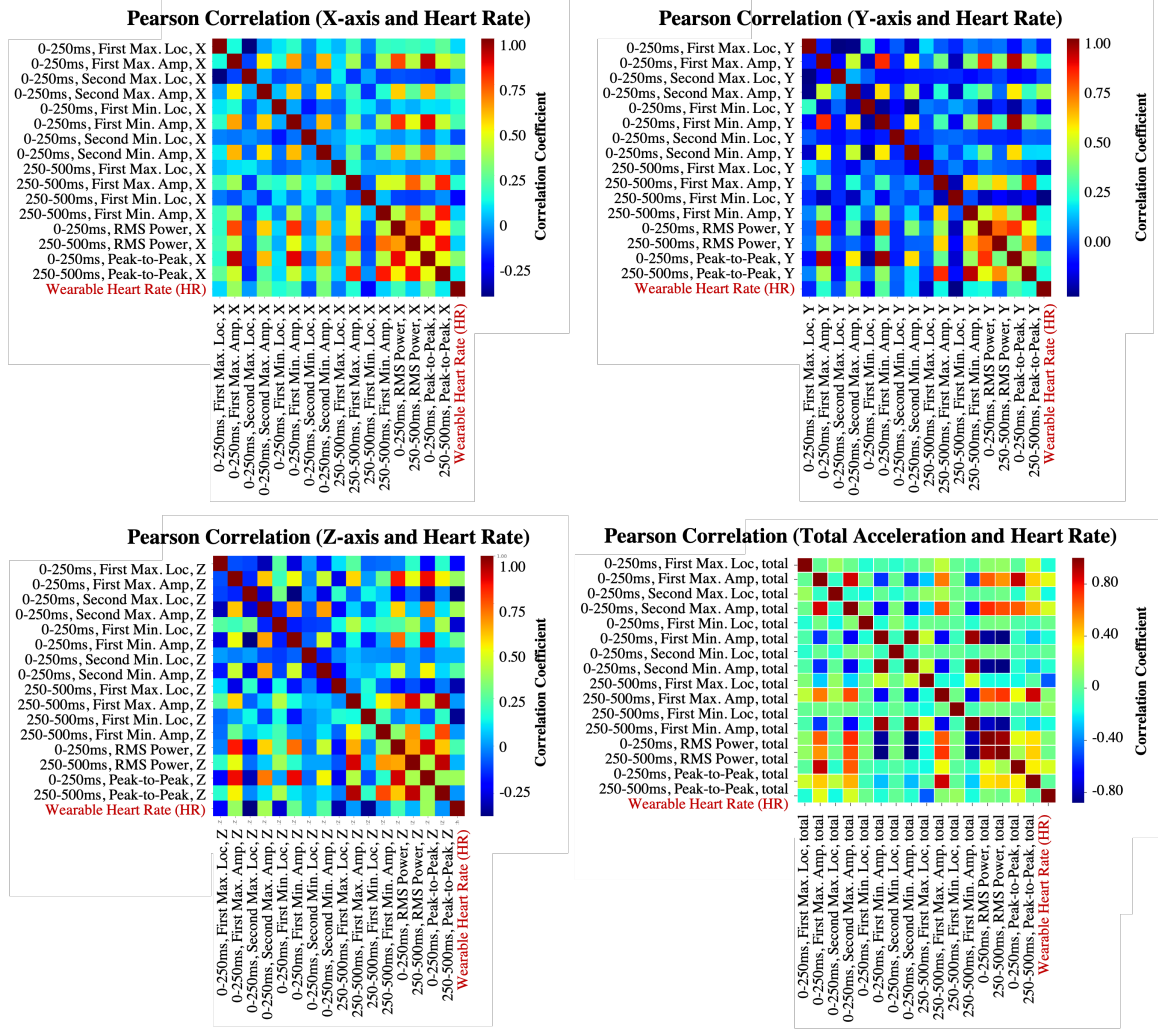


Figure 19: Pearson correlation coefficients between different feature pairs from $SCG_{x,y,z,total}$ signals.

4.2.3.3 Comparison with State-of-the-Art

Contribution of the Acoustic and Vibration Components

After segmenting the $SCG_{x,y,z,total}$ signals into PCG and vibration segments, the whole analysis pipeline was run with the same features (amplitude and location features) and same ML model (random forest including 20-min of calibration) for $PCG_{sternum}$,

SCG_{vib}, and the combination of PCG_{sternum} and SCG_{vib} cases. Results in Table 10 show that the combination of PCG_{sternum} and SCG_{vib} segments has superior performance ($r=0.76$, MedAE=9.87 mL) compared to using only- PCG_{sternum} ($r=0.69$, MedAE=11.67 mL) or only-SCG_{vib} ($r=0.35$, MedAE=12.67 mL) segments. Moreover, using the ECG features, i.e. HR, increases the estimation performance. Indeed, the combination of PCG_{sternum}, SCG_{vib} and ECG resulted in similar correlation ($r=0.81$) and MedAE (7.99 mL) values to the method presented in this work ($r=0.81$, MedAE=7.56 mL). Therefore, SV estimation benefits from both the acoustic and vibrational characteristics of the cardiovascular system. In addition, although the most important feature was heart rate in our proposed algorithm, when PCG_{sternum} and SCG_{vib} components were separated from each other, the first and second minima of the total acceleration magnitude (0-250ms) and the peak-to-peak amplitude of the head-to-foot direction (0-250ms) outperformed heart rate. In a typical PCG, the first oscillation (S1 heart sound) corresponds to the closing of the mitral valve. Thus, having features related to S1 heart sounds as the key features is indeed consistent with the underlying physiological events as stroke volume is defined as the difference between the end diastolic and end systolic volumes [37].

Comparison with the Existing Literature

Finally, the analysis pipeline offered by Tavakolian *et al.* was applied on the current dataset to provide a quantitative comparison. It should be noted that the comparison was not head to head as there were differences in the experimental setups of the two studies (mentioned in Section 4.2.2.7). In addition, the analysis pipeline presented in this study was based on training a global model and validating with LOSO-CV, whereas in [171] subject-specific models were used. Thus, the correlation and MedAE values were

calculated for each subject separately and presented as *mean \pm std* in Table 10. Overall, the MedAE results of the two studies were close to each other, however there was a great difference between the correlation values. In addition, there was a great variance among the correlation and MedAE values of the subjects, e.g. one subject had a correlation value of 0.71, whereas another one had a correlation value of 0.01. One reason might be the insufficiency of linear regression as it assumes a linear relationship between both the dependent and independent variables. Another reason might be the errors during the detection of the cardiac event points (AC, AO, etc.). As the subjects in the current dataset were undergoing surgery, it was challenging to locate and detect these fiducial points accurately. Therefore, it can be deduced that detecting the maxima and minima amplitudes and locations instead of the actual cardiac event points indeed contributes to such frameworks by providing a more representative feature set.

4.2.3.4 Limitations and Future Work

The limitations of this study are the relatively small sample size, and that the study was conducted at a single clinical site. Future studies will be necessary to assess the reproducibility of the methods including larger and more diverse datasets in patients having various diseases, and undergoing a variety of surgical procedures or different anesthetics techniques. For example, more extensive studies with serial measurements under various surgical conditions (such as fluid loading etc.) needs to be done to ensure accurate stroke volume prediction. Similarly, an increase in the number of subjects in the training set will improve the generalizability of the algorithm as well as the determination of key SCG features associated with SV prediction. Once the generalizability is achieved through validation in larger datasets, potential ways to achieve real-time monitoring will also be

studied. Future versions of the wearable patch will attempt to minimize form factor and improve comfort as well.

Table 10: Comparison with the Existing Literature

		Tavakolian <i>et al.</i> [171] (mean \pm std)	THIS WORK
SCG _x	Corr. (r-value)	0.15 \pm 0.21	0.22
	MedAE (mL)	13.29 \pm 10.98	13.76
SCG _y	Corr. (r-value)	0.09 \pm 0.13	0.65
	MedAE (mL)	13.54 \pm 10.68	12.46
SCG _z	Corr. (r-value)	0.11 \pm 0.26	0.50
	MedAE (mL)	13.21 \pm 10.75	10.86
SCG _{total}	Corr. (r-value)	0.13 \pm 0.21	0.56
	MedAE (mL)	12.96 \pm 10.91	10.49

Another drawback of this study is the additional prediction error introduced by inter-subject variability. The current problem was framed as monitoring the changes in SV throughout and after the surgery relative to the reference SV values taken in pre-incision period, therefore some baseline data was included in the training set for each subject. This reduced the added inter-subject variability error in SV prediction, but requires a brief calibration period (pre-incision) during which both the wearable device and TED were being used. To adapt the wearable patch as an alternative to the TED, the ideal scenario would be having a model without any baseline data. Therefore, future work will also focus

on methods to eliminate the effects of inter-subject variability, and the need for a baseline set to enhance the prediction performance of the trained model.

4.2.4 Conclusion

The results of this study demonstrate the feasibility of a wearable patch system to monitor the SV continuously and non-invasively that is applicable to pre-, intra- and post-operative periods. With this proposed technology, the wearable sensor could be applied to conscious, ambulatory patients and used to monitor SV without requiring anesthesia or any invasive intervention. There are obvious benefits of this technology from a patients' perspective. It eliminates the discomfort and risks associated with other techniques for hemodynamic monitoring. Additionally, this technology also has several benefits for healthcare providers. The rapidity and ease with which the device is applied to the patient and the lack of the risks associated with inadvertent needlesticks and exposure to bloodborne infections cannot be underestimated.

In conclusion, the proposed wearable patch system could be a viable alternative for peri-operative hemodynamic monitoring of surgical patients throughout the entire encounter. Combined with appropriate remote capabilities, it could facilitate monitoring the patients remotely after discharge to their homes as well. In addition, as the success of fluid management protocols based on monitoring intra-operative SV has been recently shown, this wearable technology can also be used in the future to potentially guide peri-operative fluid management.

CHAPTER 5: CONCLUSION AND FUTURE WORK

With recent advances in wearable technology and machine learning, engineered devices and algorithms may soon transform healthcare delivery from intermittent clinic visits to feedback controlled continuous health monitoring. This dissertation described the use of wearable acoustic and vibration measurements to derive digital biomarkers, which can be used together with existing medical information to assist in clinical decisions.

The first part of the dissertation mainly focused on the wearable acoustic measurements for biomechanics, and discussed how knee acoustical emissions could be leveraged to derive a *knee health score* to assist in joint health assessment, specifically in JIA. This analysis was followed by a novel click detection and classification algorithm leveraging the Teager Energy Operator, which can detect the clicks in joint sound signals and distinguish between physiologic and pathologic clicks.

The second part of the dissertation focused on the wearable acoustic and vibration measurements for cardiovascular assessment in two different applications. In the first work, it was shown that the operating sounds of LVADs may provide substantial information regarding pump thrombosis when combined with machine learning algorithms. The second project discussed the use of wearable SCG in non-invasive SV estimation and suggested that SV estimation benefits from both the acoustic and vibrational characteristics of the cardiovascular system.

Although the current frameworks showed that the wearable acoustic and vibration measurements could potentially be used to derive digital biomarkers, future studies will be

necessary to assess the reproducibility of the methods. To achieve more robust and accurate detection and analysis frameworks, future work will include (i) ensuring reliability of the measurements, (ii) verifying and validating the proposed frameworks, and (iii) achieving biomarker phenotyping. Below, these research questions and possible solutions are discussed.

5.1 Future Work

5.1.1 Ensuring Reliability of the Measurements

The first step to bring medical sensing out of the clinical settings requires investigating the ways to improve the robustness of the current frameworks. Many systems and algorithms this dissertation presented have mostly relied on controlled in-lab experiments for evaluation. Considering that the real-world environments will not be controlled, the proposed frameworks will be prone to motion artifacts, ambient noise and hardware-related hurdles.

In practical settings, patients will be expected to have these devices mounted on their body all the time for continuous monitoring. However, as patients will engage in vigorous activity throughout the day (walking, working out, etc.), signals may be corrupted by the motion artifacts and various noise sources like clothing interference, speaking, sweating, etc. Another case is when the patient will not be required to wear the device continuously, but he/she will need to mount it to record their physiological signals. Based on these challenges, future work will include investigating (1) how different methods of interfacing the sensors to the body affect the acoustic and vibration measurements, and (2) how much noise is introduced through environmental sources or hardware flaws.

To address these challenges, the first step would be ensuring standardization across the measurements, i.e. maintaining controlled data collection. This could be achieved through proper calibration methods, such as calibration through reference comparison. Another major step would be identifying the noise and error sources through appropriate noise quantification and signal quality index schemes as previously done in [154]. Once these sources are identified, they could then be labeled and detached from the actual signals, and quantified to derive actionable correction factors. Such factors could also be leveraged in the calibration step to achieve more reliable, accurate and error-free measurements.

5.1.2 Validating the Proposed Frameworks

To bring medical sensing out of the clinical settings and to use these digital biomarkers in clinical decisions, several verification and validation steps should be completed. First, the sensing systems' measurement reliability should be confirmed as confidence is the key factor for both the caregivers and patients [1]. Necessary engineering bench tests should be performed and rather than just comparing the current findings with the subjective physical examinations, direct correlations with pre-validated diagnostic tools, such as MRI, should be derived. In addition, larger studies including different patient populations should be designed. Proposed algorithms may show changing performance across different patient groups by outputting varying and inconsistent performance metrics [1]. Thus, future work should include necessary steps to ensure proper verification of the sensing systems and validation of the proposed algorithms.

5.1.3 *Achieving Biomarker Phenotyping*

This dissertation presented novel wearable frameworks leveraging vibration and acoustic measurements for knee joint and cardiovascular health assessment. However, when the goal is instead to predict the underlying physiology or the cause and symptom of the decreased condition, one may need more extensive and interpretable biomarkers. To meet this goal, different biomarkers should be derived for specific disease and injury types.

For example, the presented framework in this dissertation could distinguish between the healthy joints and joints with arthritis. However, the inherent disorder might be arthritis, meniscal tear, ligament tear, etc. Thus, the main question is how to distinguish arthritis and another joint disease or injury from each other. If the range of motion, swelling and structure of the joints could be quantified through different digital biomarkers; a more comprehensive and patient-specific wearable scheme could be developed. Therefore, it is important to determine which features encapsulate which symptom, i.e., to achieve biomarker phenotyping.

Similarly, cardiovascular issues are not only limited to deterioration of assist devices or hemodynamic monitoring. For example, a unique wearable device acquiring cardiogenic and pulmonary data (vibrations, acoustics, pressure, etc.) could be designed to track the clinical status of the heart, lung and blood simultaneously through quantifying pulmonary activity (edema, respiratory rate, etc.), cardiovascular performance (atrial and ventricular performance, etc.), baroreflex sensitivity, and blood pressure and oxygen saturation.

In addition, effect of the spatial differences across the acquired recordings should be studied through appropriate experimental models to understand the underlying clinical relevance better. For instance, the location of the LVAD recordings might potentially explain the location of the thrombosis within the device (inflow cannula, outflow cannula, etc.) or the amount of the thrombus formed within the pump. Similarly, recording the joint sounds from different knee locations could potentially tell more about the underlying cartilage and soft tissue degradations, and the severity of disease/injury-of-interest. As the recording location may affect the information we obtain from the signals, cadaver and animal models should be leveraged to understand how underlying anatomy contributes to the information acquired from the signal, which directly relates to the biomarker phenotyping we want to achieve.

5.2 Final Remarks

Wearable technology is in its early stages today and the projects presented in this dissertation are only some examples of how wearable technology can be used to quantify human health. This dissertation paved the way for leveraging wearable acoustic and vibration measurements for remote health monitoring. Once verified and validated through large studies, such systems can potentially assist in clinical decisions and improve the management of various diseases and injuries outside the physical confines of the clinic.

REFERENCES

- [1] A. Coravos, S. Khozin, and K. D. Mandl, "Developing and adopting safe and effective digital biomarkers to improve patient outcomes," *NPJ digital medicine*, vol. 2, no. 1, pp. 1-5, 2019.
- [2] K. Strimbu and J. A. Tavel, "What are biomarkers?," *Current Opinion in HIV and AIDS*, vol. 5, no. 6, p. 463, 2010.
- [3] S. Meister, W. Deiters, and S. Becker, "Digital health and digital biomarkers—enabling value chains on health data," *Current Directions in Biomedical Engineering*, vol. 2, no. 1, pp. 577-581, 2016.
- [4] R. Mayeux, "Biomarkers: potential uses and limitations," *NeuroRx*, vol. 1, no. 2, pp. 182-188, 2004.
- [5] G. Eysenbach, "What is e-health?," *J Med Internet Res*, vol. 3, no. 2, p. e20, 2001/6/18 2001, doi: 10.2196/jmir.3.2.e20.
- [6] S. D. Dorn, "Digital health: hope, hype, and Amara's law," *Gastroenterology*, vol. 149, no. 3, pp. 516-520, 2015.
- [7] S. R. Steinhubl, E. D. Muse, and E. J. Topol, "The emerging field of mobile health," *Science translational medicine*, vol. 7, no. 283, pp. 283rv3-283rv3, 2015.
- [8] T. Wang, T. Azad, and R. Rajan, "The emerging influence of digital biomarkers on healthcare," *Rock Health*, 2016.
- [9] O. Inan *et al.*, "Digitizing clinical trials," *NPJ digital medicine*, vol. 3, no. 1, pp. 1-7, 2020.
- [10] 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 sensors journal*, vol. 18, no. 22, pp. 9128-9136, 2018.
- [11] D. Whittingslow, B. Semiz, L. Ponder, P. Vega-Fernandez, O. Inan, and S. Prahalad, "Knee Joint Sounds: A Non-Invasive Modality for Classifying Knee Joint Health in Juvenile Idiopathic Arthritis," in *ARTHRITIS & RHEUMATOLOGY*, 2017, vol. 69: WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA.
- [12] D. Whittingslow, B. Semiz, L. Ponders, P. V. Fernandez, O. T. Inan, and S. Prahalad, "Analysis and Implications of Non-Invasive Knee Acoustical Emissions in Juvenile Idiopathic Arthritis," in *ARTHRITIS & RHEUMATOLOGY*, 2017, vol. 69: WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, pp. 90-91.

- [13] B. Semiz, S. Hersek, D. C. Whittingslow, L. Ponder, S. Prahalad, and O. T. Inan, "Change Point Detection in Knee Acoustic Emissions using the Teager Operator: A Preliminary Study in Patients with Juvenile Idiopathic Arthritis," in *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 2019: IEEE, pp. 1-4.
- [14] B. Semiz *et al.*, "Use of Ventricular Assist Device Acoustical Signatures to Detect Device Thrombosis," *Journal of Cardiac Failure*, vol. 24, no. 8, p. S19, 2018.
- [15] B. Semiz *et al.*, "Blood-biomarker-based acoustical feature selection improves overall classification accuracy for pump thrombosis in left ventricular assist devices," *Journal of the American College of Cardiology*, vol. 73, no. 9 Supplement 1, p. 897, 2019.
- [16] B. Semiz *et al.*, "Detecting Suspected Pump Thrombosis in Left Ventricular Assist Devices via Acoustic Analysis," *IEEE Journal of Biomedical and Health Informatics*, 2020.
- [17] B. Semiz *et al.*, "Non-Invasive Wearable Patch Utilizing Seismocardiography for Peri-Operative Use in Surgical Patients," *IEEE Journal of Biomedical and Health Informatics*, 2020.
- [18] M. Raghu, "Study to Explore the Effects of Sound Vibrations on Consciousness," *Int. J. Soc. Work Hum. Serv. Pract.*, vol. 6, pp. 75-88, 2018.
- [19] J. P. Cowan, *Handbook of environmental acoustics*. John Wiley & Sons, 1993.
- [20] C. N. Teague *et al.*, "Novel methods for sensing acoustical emissions from the knee for wearable joint health assessment," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1581-1590, 2016.
- [21] O. T. Inan *et al.*, "Ballistocardiography and seismocardiography: A review of recent advances," *IEEE journal of biomedical and health informatics*, vol. 19, no. 4, pp. 1414-1427, 2014.
- [22] R. Mollan, G. McCullagh, and R. Wilson, "A critical appraisal of auscultation of human joints," *Clinical orthopaedics and related research*, no. 170, pp. 231-237, 1982.
- [23] S. C. Abbott and M. D. Cole, "Vibration arthrometry: a critical review," *Critical Reviews™ in Biomedical Engineering*, vol. 41, no. 3, 2013.
- [24] A. P. Yoganathan *et al.*, "Use of the fast Fourier transform for frequency analysis of the first heart sound in normal man," *Medical and biological engineering*, vol. 14, no. 1, pp. 69-73, 1976.

- [25] A. Verikas *et al.*, "Data dependent random forest applied to screening for laryngeal disorders through analysis of sustained phonation: acoustic versus contact microphone," *Medical engineering & physics*, vol. 37, no. 2, pp. 210-218, 2015.
- [26] S. Swarup and A. N. Makaryus, "Digital stethoscope: technology update," *Medical devices (Auckland, NZ)*, vol. 11, p. 29, 2018.
- [27] S. Leng, R. San Tan, K. T. C. Chai, C. Wang, D. Ghista, and L. Zhong, "The electronic stethoscope," *Biomedical engineering online*, vol. 14, no. 1, pp. 1-37, 2015.
- [28] M. Di Rienzo *et al.*, "A wearable system for the seismocardiogram assessment in daily life conditions," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011: IEEE, pp. 4263-4266.
- [29] Y. Chuo *et al.*, "Mechanically flexible wireless multisensor platform for human physical activity and vitals monitoring," *IEEE transactions on biomedical circuits and systems*, vol. 4, no. 5, pp. 281-294, 2010.
- [30] D. Da He, E. S. Winokur, and C. G. Sodini, "A continuous, wearable, and wireless heart monitor using head ballistocardiogram (BCG) and head electrocardiogram (ECG)," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011: IEEE, pp. 4729-4732.
- [31] Y.-J. Hong, I.-J. Kim, S. C. Ahn, and H.-G. Kim, "Mobile health monitoring system based on activity recognition using accelerometer," *Simulation Modelling Practice and Theory*, vol. 18, no. 4, pp. 446-455, 2010.
- [32] O. T. Inan *et al.*, "Novel wearable seismocardiography and machine learning algorithms can assess clinical status of heart failure patients," *Circulation: Heart Failure*, vol. 11, no. 1, p. e004313, 2018.
- [33] H. Ashouri, S. Hersek, and O. T. Inan, "Universal pre-ejection period estimation using seismocardiography: Quantifying the effects of sensor placement and regression algorithms," *IEEE sensors journal*, vol. 18, no. 4, pp. 1665-1674, 2017.
- [34] M. M. H. Shandhi, B. Semiz, S. Hersek, N. Goller, F. Ayazi, and O. T. Inan, "Performance analysis of gyroscope and accelerometer sensors for seismocardiography-based wearable pre-ejection period estimation," *IEEE journal of biomedical and health informatics*, vol. 23, no. 6, pp. 2365-2374, 2019.
- [35] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Aaai*, 2005, vol. 5, no. 2005, pp. 1541-1546.
- [36] N. Amini, M. Sarrafzadeh, A. Vahdatpour, and W. Xu, "Accelerometer-based on-body sensor localization for health and medical monitoring applications," *Pervasive and mobile computing*, vol. 7, no. 6, pp. 746-760, 2011.

- [37] A. Taebi, B. E. Solar, A. J. Bomar, R. H. Sandler, and H. A. Mansy, "Recent advances in seismocardiography," *Vibration*, vol. 2, no. 1, pp. 64-86, 2019.
- [38] H. Ashouri and O. T. Inan, "Automatic detection of seismocardiogram sensor misplacement for robust pre-ejection period estimation in unsupervised settings," *IEEE sensors journal*, vol. 17, no. 12, pp. 3805-3813, 2017.
- [39] P. K. Jain, A. K. Tiwari, and V. S. Chourasia, "Performance analysis of seismocardiography for heart sound signal recording in noisy scenarios," *Journal of medical engineering & technology*, vol. 40, no. 3, pp. 106-118, 2016.
- [40] O. Lahdenoja *et al.*, "Atrial fibrillation detection via accelerometer and gyroscope of a smartphone," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 108-118, 2017.
- [41] S. Hersek, B. Semiz, M. M. H. Shandhi, L. Orlandic, and O. T. Inan, "A Globalized Model for Mapping Wearable Seismocardiogram Signals to Whole-Body Ballistocardiogram Signals Based on Deep Learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1296-1309, 2019.
- [42] A. McAuley, "Digital health interventions: widening access or widening inequalities?," *Public Health*, vol. 128, no. 128, pp. 1118-1120, 2014.
- [43] K. Oen, "Comparative epidemiology of the rheumatic diseases in children," *Current opinion in rheumatology*, vol. 12, no. 5, pp. 410-414, 2000.
- [44] P. J. Manners and C. Bower, "Worldwide prevalence of juvenile arthritis why does it vary so much?," *The Journal of rheumatology*, vol. 29, no. 7, pp. 1520-1530, 2002.
- [45] B. Prakken, S. Albani, and A. Martini, "Juvenile idiopathic arthritis," *The Lancet*, vol. 377, no. 9783, pp. 2138-2149, 2011.
- [46] A. Ravelli, B. Schiappapietra, S. Verazza, and A. Martini, "Juvenile idiopathic arthritis," in *The Heart in Rheumatic, Autoimmune and Inflammatory Diseases*: Elsevier, 2017, pp. 167-187.
- [47] J. Packham and M. Hall, "Long-term follow-up of 246 adults with juvenile idiopathic arthritis: functional outcome," *Rheumatology*, vol. 41, no. 12, pp. 1428-1435, 2002.
- [48] A. D. Woolf and B. Pfleger, "Burden of major musculoskeletal conditions," *Bulletin of the world health organization*, vol. 81, pp. 646-656, 2003.
- [49] B. Manaster *et al.*, *Diagnostic and surgical imaging anatomy: musculoskeletal*. Amirsys, 2007.

- [50] G. F. McCoy, J. D. McCrea, D. E. Beverland, W. G. Kernohan, and R. Mollan, "Vibration arthrography as a diagnostic aid in diseases of the knee. A preliminary report," *The Journal of bone and joint surgery. British volume*, vol. 69, no. 2, pp. 288-293, 1987.
- [51] S. Cai, S. Yang, F. Zheng, M. Lu, Y. Wu, and S. Krishnan, "Knee joint vibration signal analysis with matching pursuit decomposition and dynamic weighted classifier fusion," *Computational and mathematical methods in medicine*, vol. 2013, 2013.
- [52] Y. Wu, *Knee joint vibroarthrographic signal processing and analysis*. Springer, 2015.
- [53] W. E. BLODGETT, "Auscultation of the knee joint," *The Boston Medical and Surgical Journal*, vol. 146, no. 3, pp. 63-66, 1902.
- [54] A. Steindler, "Auscultation of joints," *JBJS*, vol. 19, no. 1, pp. 121-136, 1937.
- [55] M. L. Chu, I. Gradisar, M. Railey, and G. Bowling, "An electro-acoustical technique for the detection of knee joint noise," *Medical research engineering*, vol. 12, no. 1, pp. 18-20, 1976.
- [56] R. A. B. Mollan, "Vibration emission in bone and joints," 1989.
- [57] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.
- [58] K. Umapathy and S. Krishnan, "Modified local discriminant bases algorithm and its application in analysis of human knee joint vibration signals," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 517-523, 2006.
- [59] 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 Transactions on Biomedical Engineering*, vol. 47, no. 6, pp. 773-783, 2000.
- [60] 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: IEEE, pp. 223-228.
- [61] J.-H. Lee, C.-C. Jiang, and T.-T. Yuan, "Vibration arthrometry in patients with knee joint disorders," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 8, pp. 1131-1133, 2000.
- [62] R. M. Rangayyan and Y. Wu, "Screening of knee-joint vibroarthrographic signals using statistical parameters and radial basis functions," *Medical & biological engineering & computing*, vol. 46, no. 3, pp. 223-232, 2008.

- [63] K.-R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz, "Machine learning for real-time single-trial EEG-analysis: from brain-computer interfacing to mental state monitoring," *Journal of neuroscience methods*, vol. 167, no. 1, pp. 82-90, 2008.
- [64] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *Journal of Neuroscience methods*, vol. 167, no. 1, pp. 115-125, 2008.
- [65] B. Hamadicharef *et al.*, "Learning EEG-based spectral-spatial patterns for attention level measurement," in *2009 IEEE International Symposium on Circuits and Systems*, 2009: IEEE, pp. 1465-1468.
- [66] W.-C. Lin *et al.*, "Non-invasive knee osteoarthritis diagnosis via vibroarthrographic signal analysis," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 5, no. 3, pp. 497-507, 2014.
- [67] L.-K. Shark, H. Chen, and J. Goodacre, "Knee acoustic emission: a potential biomarker for quantitative assessment of joint ageing and degeneration," *Medical engineering & physics*, vol. 33, no. 5, pp. 534-545, 2011.
- [68] S. Hersek *et al.*, "Acoustical emission analysis by unsupervised graph mining: A novel biomarker of knee health status," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 6, pp. 1291-1300, 2017.
- [69] T. Giannakopoulos and A. Pikrakis, *Introduction to Audio Analysis: a MATLAB® approach*. Academic Press, 2014.
- [70] L. v. d. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579-2605, 2008.
- [71] G. J. Berman, D. M. Choi, W. Bialek, and J. W. Shaevitz, "Mapping the stereotyped behaviour of freely moving fruit flies," *Journal of The Royal Society Interface*, vol. 11, no. 99, p. 20140672, 2014.
- [72] M. West *et al.*, "Predicting the clinical status of human breast cancer by using gene expression profiles," *Proceedings of the National Academy of Sciences*, vol. 98, no. 20, pp. 11462-11467, 2001.
- [73] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785-794.
- [74] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189-1232, 2001.
- [75] T. Dietterich, "Ensemble learning. The handbook of brain theory and neural networks," *Arbib MA*, 2002.

- [76] S. Hersek *et al.*, "Wearable vector electrical bioimpedance system to assess knee joint health," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 10, pp. 2353-2360, 2016.
- [77] J. Friedman, T. Hastie, and R. Tibshirani, "The elements of statistical learning. vol. 1 Springer series in statistics," *New York*, 2001.
- [78] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.
- [79] S. Nalband, A. Sundar, A. A. Prince, and A. Agarwal, "Feature selection and classification methodology for the detection of knee-joint disorders," *Computer methods and programs in biomedicine*, vol. 127, pp. 94-104, 2016.
- [80] I. Nakamura, K. Michishita, M. Tanno, and K. Ito, "Synovial impingement after posterior cruciate-retaining total knee arthroplasty for rheumatoid arthritis," *Journal of Orthopaedic Science*, vol. 11, no. 3, pp. 303-307, 2006.
- [81] A. Ravelli, "Toward an understanding of the long-term outcome of juvenile idiopathic arthritis," *Clinical and experimental rheumatology*, vol. 22, no. 3, pp. 271-275, 2004.
- [82] M. Sang Jun Song, "Noise around the Knee," *Clinics in Orthopedic Surgery*, pp. 1-8, 2018.
- [83] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 32, no. 6, pp. 1109-1121, 1984.
- [84] J. F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," in *International conference on acoustics, speech, and signal processing*, 1990: IEEE, pp. 381-384.
- [85] A. Erdamar, F. Duman, and S. Yetkin, "A wavelet and teager energy operator based method for automatic detection of K-Complex in sleep EEG," *Expert Systems with Applications*, vol. 39, no. 1, pp. 1284-1290, 2012.
- [86] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5-32, 2001.
- [87] J. K. Kirklin *et al.*, "Second annual report from the ISHLT mechanically assisted circulatory support registry," *The Journal of Heart and Lung Transplantation*, vol. 37, no. 6, pp. 685-691, 2018.
- [88] C. A. Milano *et al.*, "HVAD: the ENDURANCE supplemental trial," *JACC: Heart Failure*, vol. 6, no. 9, pp. 792-802, 2018.

- [89] S. S. Najjar *et al.*, "An analysis of pump thrombus events in patients in the HeartWare ADVANCE bridge to transplant and continued access protocol trial," *The Journal of Heart and Lung Transplantation*, vol. 33, no. 1, pp. 23-34, 2014.
- [90] R. C. Starling *et al.*, "Unexpected abrupt increase in left ventricular assist device thrombosis," *New England Journal of Medicine*, vol. 370, no. 1, pp. 33-40, 2014.
- [91] J. G. Rogers *et al.*, "Intrapericardial left ventricular assist device for advanced heart failure," *New England Journal of Medicine*, vol. 376, no. 5, pp. 451-460, 2017.
- [92] M. Oz *et al.*, "Bridge experience with long-term implantable left ventricular assist devices: are they an alternative to transplantation?," *Circulation*, vol. 95, no. 7, pp. 1844-1852, 1997.
- [93] N. Uriel *et al.*, "Device thrombosis in HeartMate II continuous-flow left ventricular assist devices: a multifactorial phenomenon," *The Journal of Heart and Lung Transplantation*, vol. 33, no. 1, pp. 51-59, 2014.
- [94] J. M. Stulak, J. Cowger, J. W. Haft, M. A. Romano, K. D. Aaronson, and F. D. Pagani, "Device exchange after primary left ventricular assist device implantation: indications and outcomes," *The Annals of thoracic surgery*, vol. 95, no. 4, pp. 1262-1268, 2013.
- [95] N. Uriel *et al.*, "Development of a novel echocardiography ramp test for speed optimization and diagnosis of device thrombosis in continuous-flow left ventricular assist devices: the Columbia ramp study," *Journal of the American College of Cardiology*, vol. 60, no. 18, pp. 1764-1775, 2012.
- [96] A. S. Flores *et al.*, "Echocardiographic assessment for ventricular assist device placement," *Journal of thoracic disease*, vol. 7, no. 12, p. 2139, 2015.
- [97] J. M. Schaffer, S. K. Singh, B. A. Reitz, P. E. Oyer, R. C. Robbins, and H. R. Mallidi, "Heart transplant graft survival is improved after a reduction in panel reactive antibody activity," *The Journal of thoracic and cardiovascular surgery*, vol. 145, no. 2, pp. 555-565, 2013.
- [98] N. G. Smedira *et al.*, "Unplanned hospital readmissions after HeartMate II implantation: frequency, risk factors, and impact on resource use and survival," *JACC: Heart Failure*, vol. 1, no. 1, pp. 31-39, 2013.
- [99] T. Hasin *et al.*, "Readmissions after implantation of axial flow left ventricular assist device," *Journal of the American College of Cardiology*, vol. 61, no. 2, pp. 153-163, 2013.
- [100] A. Kilic, "The future of left ventricular assist devices," *Journal of thoracic disease*, vol. 7, no. 12, p. 2188, 2015.

- [101] F. Kaufmann *et al.*, "Acoustic spectral analysis for determining pump thrombosis in rotary blood pumps," *ASAIO journal*, vol. 60, no. 5, pp. 502-507, 2014.
- [102] T. Makino *et al.*, "Estimation of early-stage malfunction using implantable artificial heart sound in animal experiments," *Artificial organs*, vol. 30, no. 5, pp. 360-364, 2006.
- [103] L. Hubbert, P. Sundbom, M. Loebe, B. Peterzén, H. Granfeldt, and H. Ahn, "Acoustic analysis of a mechanical circulatory support," *Artificial organs*, vol. 38, no. 7, pp. 593-598, 2014.
- [104] G. L. Yost, T. J. Royston, G. Bhat, and A. J. Tautoles, "Acoustic characterization of axial flow left ventricular assist device operation in vitro and in vivo," *ASAIO journal*, vol. 62, no. 1, pp. 46-55, 2016.
- [105] P. Sundbom *et al.*, "Sound analysis of a left ventricular assist device: A technical evaluation of iOS devices," *The International journal of artificial organs*, vol. 41, no. 5, pp. 254-260, 2018.
- [106] P. Markey *et al.*, "The value of pump audiosignals in patients with left ventricular assist devices," *The Journal of Heart and Lung Transplantation*, vol. 32, no. 4, p. S30, 2013.
- [107] C. Feldmann *et al.*, "An acoustic method for systematic ventricular assist device thrombus evaluation with a novel artificial thrombus model," *Journal of thoracic disease*, vol. 10, no. Suppl 15, p. S1711, 2018.
- [108] M. S. Slaughter *et al.*, "Utilization of acoustic signatures to identify HeartMate XVE device end-of-life," *The Journal of heart and lung transplantation*, vol. 26, no. 6, pp. 579-583, 2007.
- [109] C. Bowles *et al.*, "278 Successful Treatment of Implantable Rotary Left Ventricular Assist Device (LVAD) Thrombus Is Associated with Normalisation of Acoustic Properties," *The Journal of Heart and Lung Transplantation*, vol. 30, no. 4, p. S98, 2011.
- [110] F. Castagna *et al.*, "Acoustic Analysis of a Continuous-Flow Left Ventricular Assist Device before and after Suspected Pump Thrombosis," *The Journal of Heart and Lung Transplantation*, vol. 35, no. 4, p. S326, 2016.
- [111] M. Noor, N. Hitchins, E. Doyle, N. Banner, K. Parker, and C. Bowles, "007 analysis of the left ventricular assist device (lvad) acoustic signal: A novel clinical diagnostic method for suspected pump thrombosis," *Heart*, vol. 99, no. suppl 2, p. A11, 2013.
- [112] R. Rana, K. Q. Schwarz, and J. R. Kolodziej, "Non-Invasive Fault Detection in an Axial Flow Blood Pump Used as a Ventricle Assistive Device," in *Dynamic*

Systems and Control Conference, 2014, vol. 46186: American Society of Mechanical Engineers, p. V001T06A002.

- [113] C.-C. Lin, S.-H. Chen, T.-K. Truong, and Y. Chang, "Audio classification and categorization based on wavelets and support vector machine," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 644-651, 2005.
- [114] J. Fan and J. Lv, "A selective overview of variable selection in high dimensional feature space," *Statistica Sinica*, vol. 20, no. 1, p. 101, 2010.
- [115] L. Chen, "Curse of dimensionality," in *Encyclopedia of Database Systems*, 2009.
- [116] X. Geng, T.-Y. Liu, T. Qin, and H. Li, "Feature selection for ranking," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, 2007, pp. 407-414.
- [117] V. Sood *et al.*, "Correlation of pre-explant lactate dehydrogenase concentrations and findings during post-explant pump analysis of the HeartMate II left ventricular assist device," *The Annals of thoracic surgery*, vol. 103, no. 4, pp. 1207-1213, 2017.
- [118] A. B. Nguyen, N. Uriel, and S. Adatya, "New challenges in the treatment of patients with left ventricular support: LVAD thrombosis," *Current heart failure reports*, vol. 13, no. 6, pp. 302-309, 2016.
- [119] D. L. Jennings and P. A. Weeks, "Thrombosis in continuous-flow left ventricular assist devices: pathophysiology, prevention, and pharmacologic management," *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, vol. 35, no. 1, pp. 79-98, 2015.
- [120] P. Shah *et al.*, "Lactate dehydrogenase is superior to serum free hemoglobin as a marker of pump thrombosis in left ventricular assist devices," *The Journal of Heart and Lung Transplantation*, vol. 32, no. 4, p. S37, 2013.
- [121] J. A. Cowger *et al.*, "Hemolysis: a harbinger of adverse outcome after left ventricular assist device implant," *The Journal of heart and lung transplantation*, vol. 33, no. 1, pp. 35-43, 2014.
- [122] C. Spearman, "The proof and measurement of association between two things," *The American journal of psychology*, vol. 100, no. 3/4, pp. 441-471, 1987.
- [123] D. E. Hinkle, W. Wiersma, and S. G. Jurs, *Applied statistics for the behavioral sciences*. Houghton Mifflin College Division, 2003.
- [124] A. Ramlaul and P. Hogg, *Medical imaging and radiotherapy research: skills and strategies*. Springer, 2010.
- [125] M. M. Mukaka, "A guide to appropriate use of correlation coefficient in medical research," *Malawi medical journal*, vol. 24, no. 3, pp. 69-71, 2012.

- [126] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning* (no. 10). Springer series in statistics New York, 2001.
- [127] D. J. Goldstein *et al.*, "Algorithm for the diagnosis and management of suspected pump thrombus," ed: Elsevier, 2013.
- [128] S. M. Abbas and A. Hill, "Systematic review of the literature for the use of oesophageal Doppler monitor for fluid replacement in major abdominal surgery," *Anaesthesia*, vol. 63, no. 1, pp. 44-51, 2008.
- [129] T. E. Miller, A. M. Roche, and M. Mythen, "Fluid management and goal-directed therapy as an adjunct to Enhanced Recovery After Surgery (ERAS)," *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, vol. 62, no. 2, pp. 158-168, 2015.
- [130] T. J. Gan *et al.*, "Goal-directed intraoperative fluid administration reduces length of hospital stay after major surgery," *Anesthesiology: The Journal of the American Society of Anesthesiologists*, vol. 97, no. 4, pp. 820-826, 2002.
- [131] M. G. Mythen and A. R. Webb, "Perioperative plasma volume expansion reduces the incidence of gut mucosal hypoperfusion during cardiac surgery," *Archives of Surgery*, vol. 130, no. 4, pp. 423-429, 1995.
- [132] S. Walsh, T. Tang, S. Bass, and M. Gaunt, "Doppler-guided intra-operative fluid management during major abdominal surgery: systematic review and meta-analysis," *International journal of clinical practice*, vol. 62, no. 3, pp. 466-470, 2008.
- [133] S. Noblett, C. Snowden, B. Shenton, and A. Horgan, "Randomized clinical trial assessing the effect of Doppler-optimized fluid management on outcome after elective colorectal resection," *British Journal of Surgery*, vol. 93, no. 9, pp. 1069-1076, 2006.
- [134] H. Wakeling *et al.*, "Intraoperative oesophageal Doppler guided fluid management shortens postoperative hospital stay after major bowel surgery," *British journal of anaesthesia*, vol. 95, no. 5, pp. 634-642, 2005.
- [135] A. M. Roche, T. E. Miller, and T. J. Gan, "Goal-directed fluid management with trans-oesophageal Doppler," *Best Practice & Research Clinical Anaesthesiology*, vol. 23, no. 3, pp. 327-334, 2009.
- [136] B. M. Chamberlain and R. J. Willshire, "Oesophageal Doppler Monitor (ODM) guided individualised goal directed fluid management (iGDFM) in surgery-a technical review," *surgery*, vol. 19, p. 21, 2010.
- [137] M. Singer, "Continuous haemodynamic monitoring by oesophageal Doppler (MD thesis)," *London: University of London*, 1990.

- [138] "CardioQ-ODM Product Brochure." Deltex Medical. www.deltexmedical.com (accessed).
- [139] J.-L. Teboul *et al.*, "Less invasive hemodynamic monitoring in critically ill patients," *Intensive care medicine*, vol. 42, no. 9, pp. 1350-1359, 2016.
- [140] R. Makaryus, T. Miller, and T. Gan, "Current concepts of fluid management in enhanced recovery pathways," *British journal of anaesthesia*, vol. 120, no. 2, pp. 376-383, 2018.
- [141] A. G. Kieback, A. C. Borges, T. Schink, G. Baumann, and M. Laule, "Impedance cardiography versus invasive measurements of stroke volume index in patients with chronic heart failure," *International journal of cardiology*, vol. 143, no. 2, pp. 211-213, 2010.
- [142] E. Lorne *et al.*, "Accuracy of impedance cardiography for evaluating trends in cardiac output: a comparison with oesophageal Doppler," *British journal of anaesthesia*, vol. 113, no. 4, pp. 596-602, 2014.
- [143] G. Cybulski, E. Michalak, E. Koźluk, A. Piątkowska, and W. Niewiadomski, "Stroke volume and systolic time intervals: beat-to-beat comparison between echocardiography and ambulatory impedance cardiography in supine and tilted positions," *Medical and Biological Engineering and Computing*, vol. 42, no. 5, pp. 707-711, 2004.
- [144] M. Etemadi, O. T. Inan, L. Giovannardi, and G. T. Kovacs, "Rapid assessment of cardiac contractility on a home bathroom scale," *IEEE transactions on information technology in biomedicine*, vol. 15, no. 6, pp. 864-869, 2011.
- [145] S. Kamo *et al.*, "Improved pressure contour analysis for estimating cardiac stroke volume using pulse wave velocity measurement," *Biomedical engineering online*, vol. 16, no. 1, pp. 1-19, 2017.
- [146] M. M. Hirschl *et al.*, "Noninvasive assessment of cardiac output in critically ill patients by analysis of the finger blood pressure waveform," *Critical care medicine*, vol. 25, no. 11, pp. 1909-1914, 1997.
- [147] S. Rang, B. de Pablo Lapiedra, G. A. van Montfrans, B. J. Bouma, K. H. Wesseling, and H. Wolf, "Modelflow: a new method for noninvasive assessment of cardiac output in pregnant women," *American journal of obstetrics and gynecology*, vol. 196, no. 3, pp. 235. e1-235. e8, 2007.
- [148] Y. Obata, M. Mizogami, D. Nyhan, D. E. Berkowitz, J. Steppan, and V. Barodka, "Pilot study: Estimation of stroke volume and cardiac output from pulse wave velocity," *PloS one*, vol. 12, no. 1, p. e0169853, 2017.

- [149] H. Ashouri, L. Orlandic, and O. T. Inan, "Unobtrusive estimation of cardiac contractility and stroke volume changes using ballistocardiogram measurements on a high bandwidth force plate," *Sensors*, vol. 16, no. 6, p. 787, 2016.
- [150] M. Etemadi, O. T. Inan, J. A. Heller, S. Hersek, L. Klein, and S. Roy, "A wearable patch to enable long-term monitoring of environmental, activity and hemodynamics variables," *IEEE transactions on biomedical circuits and systems*, vol. 10, no. 2, pp. 280-288, 2015.
- [151] M. Etemadi and O. T. Inan, "Wearable ballistocardiogram and seismocardiogram systems for health and performance," *Journal of Applied Physiology*, vol. 124, no. 2, pp. 452-461, 2018.
- [152] C. Oates, *Cardiovascular haemodynamics and Doppler waveforms explained*. Cambridge University Press, 2008.
- [153] F. Khosrow-Khavar, K. Tavakolian, A. P. Blaber, J. M. Zanetti, R. Fazel-Rezai, and C. Menon, "Automatic annotation of seismocardiogram with high-frequency precordial accelerations," *IEEE journal of biomedical and health informatics*, vol. 19, no. 4, pp. 1428-1434, 2014.
- [154] J. Zia, J. Kimball, S. Hersek, M. M. H. Shandhi, B. Semiz, and O. T. Inan, "A Unified Framework for Quality Indexing and Classification of Seismocardiogram Signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 1080-1092, 2019.
- [155] M. J. Tadi *et al.*, "A real-time approach for heart rate monitoring using a Hilbert transform in seismocardiograms," *Physiological measurement*, vol. 37, no. 11, p. 1885, 2016.
- [156] K. Pandia, O. T. Inan, and G. T. Kovacs, "A frequency domain analysis of respiratory variations in the seismocardiogram signal," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013: IEEE, pp. 6881-6884.
- [157] A. Aguirre, G. Wodicka, C. Maayan, and D. Shannon, "Interaction between respiratory and RR interval oscillations at low frequencies," *Journal of the autonomic nervous system*, vol. 29, no. 3, pp. 241-246, 1990.
- [158] P.-F. Migeotte *et al.*, "Three dimensional ballistocardiography: methodology and results from microgravity and dry immersion," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011: IEEE, pp. 4271-4274.
- [159] S. A. Dyer and J. S. Dyer, "Cubic-spline interpolation. 1," *IEEE Instrumentation & Measurement Magazine*, vol. 4, no. 1, pp. 44-46, 2001.

- [160] S. D. Conte and C. De Boor, *Elementary numerical analysis: an algorithmic approach*. SIAM, 2017.
- [161] S. McKinley and M. Levine, "Cubic spline interpolation," *College of the Redwoods*, vol. 45, no. 1, pp. 1049-1060, 1998.
- [162] L. Breiman, "Random Forests Machine Learning, vol. 45," 2001.
- [163] A. M. Prasad, L. R. Iverson, and A. Liaw, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction," *Ecosystems*, vol. 9, no. 2, pp. 181-199, 2006.
- [164] J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The lancet*, vol. 327, no. 8476, pp. 307-310, 1986.
- [165] P. M. Odor, S. Bampoe, and M. Cecconi, "Cardiac Output Monitoring: Validation Studies—how Results Should be Presented," *Current Anesthesiology Reports*, vol. 7, no. 4, pp. 410-415, 2017.
- [166] L. A. Critchley and J. A. Critchley, "A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques," *Journal of clinical monitoring and computing*, vol. 15, no. 2, pp. 85-91, 1999.
- [167] A. Moosavi, A. Attia, and A. Sandu, "Tuning Covariance Localization Using Machine Learning," in *International Conference on Computational Science*, 2019: Springer, pp. 199-212.
- [168] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise reduction in speech processing*: Springer, 2009, pp. 1-4.
- [169] P. Castiglioni, A. Faini, G. Parati, and M. Di Rienzo, "Wearable seismocardiography," in *2007 29th annual international conference of the IEEE engineering in medicine and biology society*, 2007: IEEE, pp. 3954-3957.
- [170] K. Pandia, O. T. Inan, G. T. Kovacs, and L. Giovangrandi, "Extracting respiratory information from seismocardiogram signals acquired on the chest using a miniature accelerometer," *Physiological measurement*, vol. 33, no. 10, p. 1643, 2012.
- [171] K. Tavakolian, A. P. Blaber, B. Ngai, and B. Kaminska, "Estimation of hemodynamic parameters from seismocardiogram," in *2010 Computing in Cardiology*, 2010: IEEE, pp. 1055-1058.