Detecting Lumbar Muscle Fatigue Using Nanocomposite Strain Gauges

Darci Ann Billmire
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Engineering Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/10023

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Detecting Lumbar Muscle Fatigue

Using Nanocomposite

Strain Gauges

Darci Ann Billmire

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Anton E. Bowden, Chair
David T. Fullwood
Ulrike H. Mitchell

Department of Mechanical Engineering
Brigham Young University

Copyright © 2023 Darci Ann Billmire
All Rights Reserved
Detecting Lumbar Muscle Fatigue Using Nanocomposite Strain Gauges

Darci Billmire
Department of Mechanical Engineering
Master of Science
Abstract

Introduction: Muscle fatigue can contribute to acute flare-ups of lower back pain with associated consequences such as pain, disability, lost work time, increased healthcare utilization, and increased opioid use and potential abuse. The SPINE Sense system is a wearable device with 16 high deflection nanocomposite strain gauge sensors on kinesiology tape which is adhered to the skin of the lower back. This device is used to correlate lumbar skin strains with the motion of the lumbar vertebrae and to phenotype lumbar spine motion. In this work it was hypothesized that the SPINE Sense device can be used to detect differences in biomechanical movements consequent to muscle fatigue.

A human subject study was completed with 30 subjects who performed 14 functional movements before and after fatiguing their back muscles through the Biering-Sørensen endurance test with the SPINE Sense device on their lower back collecting skin strain data. Various features from the strain gauge sensors were extracted from these data and were used as inputs to a random forest classification machine learning model. The accuracy of the model was assessed under two training/validation conditions, namely a hold-out method and a leave-one-out method. The random forest classification models were able to achieve up to 84.22% and 78.37% accuracies for the hold-out and leave-one-out methods respectively.

Additionally, a system usability study was performed by presenting the device to 32 potential users (clinicians and individuals with lower back pain) of their device. They received a scripted explanation of the use of the device and were then instructed to score it with the validated System Usability Score. In addition they were given the opportunity to voice concerns, questions, and offer any other additional feedback about the design and use of the device.

The average System Usability Score from all participants from the system usability study was 72.03 with suggestions of improving the robustness of electrical connections and smaller profiles of accompanying electronics. Feedback from the potential users of the device was used to make more robust electrical connections and smaller wires and electronics modules. These improvements were achieved by making a two-piece design: one piece contains the sensors on kinesiology tape that is directly attached to the patient and the other one contains the wires sewn into stretch fabric to create stretchable electronic connections to the device.

It is concluded that a machine-learning model of the data from the SPINE Sense device can classify lumbar motion with sufficient accuracy for clinical utility. It is also concluded that the device is usable and intuitive to use.
Keywords: muscle fatigue, low back pain, high deflection strain gauges, nanocomposite sensors, system usability, biomechanics, sensors, machine learning, random forest classification, and cross-validation.
Acknowledgments

First, I would like to thank and acknowledge my graduate advisor and mentor Dr. Anton E. Bowden. He has helped me foster skills necessary for me to achieve my goal of becoming an influential engineer and improve lives through the discipline of biomedical engineering. I would also like to thank my other committee members Drs. David Fullwood and Ulrike Mitchell for mentoring me and supporting me along this journey.

Next, I would like to thank my family for their love and support during these many years of schooling. My parents Tad and Donna Billmire have encouraged and supported me from the very beginning. If it weren’t for them raising me to believe and know that I can achieve anything I put my mind to, I would not be here today.

I would also like to thank my friends for their constant support and amazement in my pursuit of my degree in mechanical engineering. Along with them I would like to thank and acknowledge the many undergraduates that I have worked with on the BACPAC project. Working on the design and manufacturing of the SPINE Sense device has only been possible with their many hours of hard work and dedication. They include: Ashley Quinn, Alaina Stewart, Andrew Clawson, Anna Jensen, Braden Trejo, Caleb Swain, Cera Gowans, Emma Bowden, Cristian Payne, Dallin Jacobs, Gavin Johnson, Hailey Jones, Heidi Hunter, James Anderson, Jonathan Nance, Jordan Ryan, Joseph Peterson, Kade Waite, Kaleigh Renninger, Paul Walker, Thomas Andrews, and any others that I may have missed. I also cannot forget to acknowledge my fellow graduate students and friends that I have been able to collaborate with and enjoy many Fun Fridays with: Alex Wonnacott, Jacob Carter, Daniel Orr, Andrew Gibbons, Austin Johnson, Spencer Baker, Kurt Jensen, Brianna Bischoff, David Wood, Paul McMullin, Genesis Moreno, Adam Bilodeau, and Danni Porter.

Most importantly I would like to thank and acknowledge my Heavenly Father and my Savior Jesus Christ. My pursuit to grow and become like them has been my “why” to gaining an education and constantly trying challenging things that push me out of my comfort zone. I know that I have received much help and power through the Atonement of my Savior Jesus Christ over the course of my education that has helped me become the person I am today. I give them all of the glory and praise and echo the words of Ammon and say “I do not boast in my own strength, nor in my own wisdom; but behold, my joy is full, yea, my heart is brim with joy, and I will rejoice in my God. Yea, I know that I am nothing; as to my strength I am weak; therefore I will not boast of myself, but I will boast of my God, for in his strength I can do all things…” (Alma 26:11-12).
# Table of Contents

Abstract iii  
Acknowledgments v  
Table of Contents vi  
List of Figures vii  
List of Tables ix  
1 Introduction 1  
2 Background 3  
  2.1 Low Back Pain 3  
  2.2 Muscle Fatigue 4  
  2.3 Nanocomposite Strain Gauges 5  
  2.4 Phenotyping Motion 6  
  2.5 SPINE Sense System 7  
3 Wearable Strain Sensor-Based Detection of Lumbar Fatigue Using Machine Learning 9  
  3.1 Introduction 9  
  3.2 Methods 10  
  3.3 Results 17  
  3.4 Discussion 19  
4 SPINE Sense System Usability 24  
  4.1 Introduction 24  
  4.2 Methods 25  
  4.3 Results 26  
  4.4 Discussion 26  
5 Conclusions and Future Research Directions 29  
References 32  
Appendix A Subject Test Forms 40  
Appendix B System Usability Study Test Plan 45
List of Figures

Figure 2.1: The SPINE Sense system with exposed nanocomposite strain gauge sensors. 7

Figure 3.1: a) Photo representation of the SPINE Sense system. b) Photo of the SPINE Sense system attached to the back with the smartphone application screen and electronics attached. c) Demonstration of movement in the 3 planes of motion that are included in the 14 exercises performed by participants in the study. d) The unsupported horizontal position held during the Biering-Sørensen endurance test. 12

Figure 3.2: An example of resistance data from one of the 16 nanocomposite strain gauges for 6 repetitions and an expanded view of one repetition with the significant points that were used as features numbered and explained in Table 3.2. 13

Figure 3.3: Example of how the data was organized for the hold-out cross-validation approach of the random forest classification model and illustrates the way the model was trained with 70% of the data in yellow used to train the model and 30% in the gray being used as test data. 16

Figure 3.4: Example of how the data was organized for the leave-one-out cross-validation approach of the random forest classification model and illustrates the way the model was trained with one subject set aside as the testing data set while the model was trained on all other subjects’ data. 16

Figure 3.5: The SHAP values of the top 20 features with the most impact from the leave-one-out model where the number next to the feature title indicates the sensor number of that feature. 18

Figure 3.6: Mean SHAP Value bar plot with the most impactful features in descending order. 19

Figure 3.7: Shows the approximate location of the sensors in relation to the vertebrae represented by the labeled rectangles. 21

Figure 3.8: An example of the resistance curve for pre and post fatigue from one sensor for one repetition from one exercise. 21

Figure 3.9: Histogram that shows the distribution of the starting trough values from sensor 5 pre and post fatigue from every exercise and every subject. 22

Figure 4.1: The version of the SPINE Sense system that was presented to potential users of the device for the System Usability Study. 25

Figure 4.2: The change of wiring of the SPINE Sense device from embedding crimped wire in silicone to sewing wire into the kinesiology tape to the two-piece reusable wiring design. 27
Figure 4.3: Various SPINE Sense design improvements including a) the compact electronics design, b) the micro-HDMI cable, c) mounted micro-HDMI port on the wire piece, d) the full assembly with storage pockets and identification tags.  28

Figure 5.1: Sensors that had the highest occurrence among the features with the top 20 Gini importance from each leave-one-out model.  30
List of Tables

Table 3.1: Summarizes the mean and standard deviation values of the demographics of the participants in the study. 11

Figure 3.2: An example of resistance data from one of the 16 nanocomposite strain gauges for 6 repetitions and an expanded view of one repetition with the significant points that were used as features numbered and explained in Table 3.2. 11

Table 3.3: Accuracy results from various machine learning methods for each of the approaches of organizing the training and testing data. 17

Table 4.1: Breakdown of the System Usability Scale scores from the system usability study. 26
1 Introduction

Lower back pain is one of the most common health problems and disabilities in the world. Unfortunately, this issue contributes to an even larger more dangerous issue of the opioid epidemic. Preventing lower back pain could help billions of people worldwide from pain and potential drug abuse. Muscle fatigue is a common mechanism that contributes to lumbar spinal pathologies and can even transition an acute case of lower back pain to a chronic case. The ability to detect muscle fatigue of the lumbar spine could aid in the prevention of injury from muscle fatigue. With the help of new wearable technologies that collect biomechanical data of spinal motions, high deflection strain gauges can be used to track motion of the lumbar spine. The hypothesis of this research is that a nanocomposite strain sensor array (the SPINE Sense Array) applied on the skin of the lumbar region can detect changes in spinal motion due to muscle fatigue. The structure of the thesis is described below.

Chapter 2 provides background with an accompanying survey of the literature related to the thesis topics. In particular, previous efforts to measure and quantify fatigue-related kinematics using different measuring techniques. It provides a deeper understanding of the extent of the problem that lower back pain including its association with the current opioid epidemic. Further information and previous work on the high deflection nanocomposite strain gauges is also discussed along with their use on the SPINE Sense device and its application and use. The use of these sensors in phenotyping motion in comparison to other popular techniques is also explored in this chapter.

Chapter 3 is a draft journal paper that will be submitted to IEEE Transactions on Biomedical Engineering with co-authors being: Dr. Anton Bowden, Dr. David Fullwood, Dr. Ulrike Mitchell, Christophe Giraud-Carrier, Spencer Baker, Joseph Peterson, and Caleb Swain. This work was performed in the BABE L lab at Brigham Young University where a human subject test was conducted on 30 participants to put this hypothesis to the test. Each participant had a SPINE Sense array placed on their lumbar spine where they were then instructed on how to perform basic exercises that included movements in all three kinematic planes of motion and combinations of them. Strain data was collected from the array while the subjects performed these exercises. The subjects then performed an endurance test that fatigued the muscles in the lumbar spine, and immediately repeated the same exercises after the fatigue event. The data pre and post fatigue was processed, and features were extracted to use a random forest classification machine learning model to classify movements as fatigued or not. These models were able to achieve up 84% accuracy, proving that the hypothesis of this research is true.
Chapter 4 presents additional work performed on the SPINE Sense array to assess the usability of the device and attain feedback from potential users of the device to make design changes and improvements on the system. Particular gratitude and thanks is attributed to Eric Lee, Heidi Hunter, Thomas Andrews, Jared Brinkman, and Emma Bowden on their work and contributions to this part of the work. The device was presented to 32 potential users of it, including clinicians and patients with chronic lower back pain, with a scripted explanation of how the system works. The participants then completed a System Usability Score for the system and were then given the opportunity to voice concerns and questions about the design of the device and its potential use. The average score of the System Usability Scores was 72.03, denoting that the system itself is usable and intuitive to use. User feedback about the design of the device was implemented by updating the design of the device to have more robust and reusable electrical connections, wiring, and smaller more compact electronics.

Chapter 5 contains conclusions and suggestions for future work extending the results of this thesis. It is concluded that the SPINE Sense device is capable of changes in motion that denote muscle fatigue, and that the system itself is usable. There are many unexplored capabilities of the SPINE Sense device and the ability to detect muscle fatigue was one of them. The future of this device could help change and improve the lives of many people across the world. Further research could be performed to improve the variability across devices and sensors to improve repeatability and performance of the system itself. The importance of different movements, features, and sensors should also be investigated beyond what was done in this work to better understand what information is needed from the device to phenotype motions of the lumbar spine.
2 Background

2.1 Low Back Pain

Lower back pain is one of the most debilitating health problems in the world.\cite{1} It has been well established that most people experience some form of lower back pain at some point in their lives. \cite{2} Experts estimate that 80% of the population of the world will experience back pain at some point in their life.\cite{3} While it is more common for older individuals and females to experience lower back pain, as with most musculoskeletal pathologies it is prevalent across all ages and both sexes.\cite{4, 5} Of all spinal problems, lumbar spinal pathologies are up to 10 times more prevalent than those of other areas of the spine. \cite{6} Lower back pain is one of the most prevalent health issues in the world and affects people on a daily basis in their work and personal lives. \cite{7}

This problem also costs citizens of the United States of America hundreds of billions of dollars each year.\cite{8, 9} These costs are direct and indirect costs of pathologies sustained in the lower back or the cost of diagnosing and treating non-specific acute and chronic back pain conditions. \cite{10, 11} The bulk of these costs are associated with time and money lost from the inability to work while dealing with a certain ailment of the lumbar spine. A cost that is not generally considered in these studies is the costs associated with the possible addiction to narcotics that are prescribed to cope with the pain.

Lumbar back pain is a large contributing factor to the opioid epidemic in the US as patients become dependent and sometimes addicted to opioids prescribed for the management of their pain.\cite{12-14} While opioids are not the first recommendation for people experiencing acute or chronic lower back pain, these problems have still contributed largely to the prescribing of narcotics and their overuse and addiction.\cite{15, 16} Opioids are the most commonly prescribed class of drugs for lower back pain according to databases of insurance claims.\cite{17} This contributes to the increase in use of opioids that is being tracked regularly where it was found that the increase in sales of opioids quadrupled from 1999 to 2010 whereas the increase of sales for people with spine problems is 600% that is attributed to increased prices and increased use.\cite{14} This increase in use and sales of opioids is matched with the same, quadrupled, increase in abuse and deaths from opioids. Avoiding these associated costs and the pain associated with lower back pain is important to understand for increased health and mental well-being across the globe.
Not only is this problem very prevalent but it is actually getting worse. The prevalence for cases of lower back pain is increasing along with the use of narcotics for the problem.\cite{18, 19} As is the case with most problems, it is always cheaper and better to put effort into prevention of the problem rather than experience and pay for the reaction to the problem after it has already occurred. \cite{20-22} For example, D. Steffens et al showed in a systematic review that exercise and education coupled together have the ability to reduce the risk of a lower back pain episode by 45% and reduce the risk and costs associated with sick-leave by 78%. \cite{23}

This is the case with lower back pain.\cite{23-26} Preventative measures are much easier to implement than the alternative of entering a reoccurring pain cycle. Acute and chronic lower back pain are both very prevalent in the United States and a person with acute lower back pain can be likely to transition to a case of chronic lower back pain when the appropriate treatment and care is not administered or sought out.\cite{27, 28} Studies haves been conducted to determine the best course of treatment to prevent the transition from acute to chronic lower back pain, but results have been inconclusive as to how to avoid long term pain. The studies include randomized clinical trials exploring the effects of early intervention of physical therapy on the subject’s levels of disability and pain, and longitudinal studies that observed the effects of lower back pain on participants ability to work and the extent of participant’s pain and symptoms.\cite{28, 29}

2.2 Muscle Fatigue

Muscle fatigue is part of the common mechanism involved in flare-ups of chronic low back pain. It is common for individuals with chronic low back pain to have structural deficiencies in their lumbar spine, such as degenerated discs\cite{30, 31}, damaged and/or repaired ligament tissue\cite{32, 33}, spondylolisthesis\cite{34-36}, spinal stenosis\cite{37-39}, or other conditions which can trigger consistent, but manageable levels of chronic pain\cite{40-42}. These structural instabilities result in passive instabilities in the control mechanisms for the lumbar spine, which can be partially compensated through core muscle activation.\cite{43-47} Commonly a person will fatigue their back muscles and not notice because the body will recruit additional muscles to perform the same tasks, but the fatigued muscles introduce risk.\cite{48, 49} This risk stems partially from the general increase in injury associated with muscle fatigue. By definition, the muscles are slower to respond and have a decrease in their capacity to generate whatever force is needed to perform a certain task.\cite{50} However, the more substantial risk is the loss of muscle control to provide secondary stabilization in the absence of passive structural stability. This increases the intervertebral instability of the spine, aggravating overuse damage to both hard and soft tissues, as well as triggering nerve pinching, tearing of soft tissue, microtrauma, inflammation, generalized tissue damage, accelerating disc herniation, and provoking other injuries specified under a large umbrella of acute lower back pain, thus thrusting the patient into a vicious pain cycle.\cite{51}

Muscle fatigue is aggravated when people with lower back pain start compensating and move in a pain avoidance pattern that is noticeably different from normal biomechanical movements.\cite{52-57} Due to the changes in movement and other biopsychosocial problems, the recruitment of muscles is different in order of recruitment and in magnitudes of force being recruited.\cite{58} These pain
avoidance patterns have been determined to be effective predictors of the chronicity of patients demonstrating acute lower back pain.[59] This additional recruitment can lead to additional injury of those newly recruited muscles and enervation of the muscles in the back that are in pain and not being recruited. This weakening of the back muscles starts a common pain cycle in patients with lower back pain.[60] They weaken the muscles by not using them when they are in pain and inflamed, which in turn decreases the stability of the lumbar spine resulting in a higher risk of reinjury when the pain is subsided starting the cycle over again.

Muscle fatigue has been studied for years along with different ways of quantifying it. Maximum voluntary contraction and power output have been used to measure fatigue.[61] Endurance times and electromyography are also considered to be closely related to muscle fatigue. As muscle fatigue increases the endurance times and electromyograph signals both decrease as the muscle activation and force generation decrease. [62] The fatigability of a muscle is the relationship of the amount of muscle force lost over time and is coupled with the idea of recovery which is the rate at which the muscles return to their baseline force output. [63]

2.3 Nanocomposite Strain Gauges

Strain gauges are widely used by billions of people every day and they have been used for decades since their invention 1938 by Edward E. Simmons and Arthur C. Ruge.[64] Typical strain gauges involve metal foils that as their stretch change their cross-sectional area that in turn gives a change in resistance through the metal that is directly correlated to the strain that the sensor underwent. These sensors are typically only accurate for small strains from 0-5%. These sensors are effective and useful for that range of strain but fail when it comes to wanting to measure things like biomechanical strain that typically ranges from 10-40% [65], and can exceed 100% in some instances[66-68].

In recent years there has been a large increase and push for wearable medical devices that can provide patient specific diagnostic information. This increase has facilitated the need for strain gauges that can measure large strains that are typical of biomechanical strain. One biomechanical strain that is of particular interest is that of skin strain. Skin is the body’s largest organ and capturing how it moves and strains can give insight into the biomechanical health of that person. Liquid mercury strain gauges, extensometers, and motion capture have all been used to measure large amounts of strain.[69] Alternative options to measure biomechanical strain have been developed in the last decade.

Work done in the BABEL lab at Brigham Young University under the direction of Drs. Anton Bowden and David Fullwood has developed nanocomposite strain gauges, or they are also referred to as High Displacement Strain Gauges.[65] The nanocomposite strain sensors are composed of Nickel nanostrands (NiNs) and nickel-coated carbon fibers(NCCF) uniformly distributed in a silicone matrix. These sensors achieve very high strains, exceeding 60% strain for thousands of cycles before plastic deformation[69], and can accurately measure even higher strains for shorter number of cycles. The sensors demonstrate a negative piezoresistivity when they are strained. Understanding of this phenomenon was discovered due to work in percolation modeling and quantum
tunneling effect. [69-71] Further work in modeling the sensors and understanding their mechanical and electrical properties is currently being worked on. [72]

These nanocomposites strain gauges have been used on multiple biomedical devices including: a knee sleeve for the therapy and recovery of people who have undergone a total knee replacement, [73, 74] a belly band to be worn by pregnant women to track the movements and kicks of the baby, and the SPINE Sense system that is placed on the lower back to track the segmental motion of the lumbar spine to be used as diagnostic tool for people suffering from non-specific lower back pain. [66, 68, 75]

2.4 Phenotyping Motion

It has been well established for decades that people demonstrate different movements patterns in accordance with different pains and ailments that their bodies are experiencing. [76-78] These movement patterns are called movement phenotypes and have been analyzed in fields of biomechanics and medicine. People consciously and subconsciously identify movements and positions that trigger pain and alter their movements accordingly to avoid these painful movements, producing distinct phenotypes in motion, which are characteristic of underlying pain producing pathologies.

Different characteristics of movement of the lumbar spine have been analyzed to determine what changes in the phenotypes of motion tell us the most information to characterize between asymptomatic and symptomatic subjects. [79-84] These characteristics include features like position, velocity, acceleration, jerk, and angular velocities of different body parts or spinal segments. Original work in this area relied on expensive and time-consuming analysis of marker-based motion capture systems, however recent work has been done to use skin mounted sensors and wearable sensors to classify phenotypes of motion of the lumbar spine. [68, 75, 85]

Current methods for collecting and interpreting biomechanical data include using inertial measurement units (IMUs), accelerometers, motion capture technology, and high deflection strain gauges. IMUs are devices that contain an accelerometer, gyroscope, and most of the time a magnetometer. Accelerometers and IMUs have been used to collect biomechanical data of everything from athletes in sport to asymptomatic and symptomatic cohorts of patients with biomechanical disorders. [86-92] Motion capture has quickly become a gold standard for tracking the position of anatomical landmarks with video cameras, that then employ direct linear transformation formulas to output positions, velocities, accelerations, and angles. That kinematic data can then be used to model biomechanical movements in various software, like OpenSim, to extrapolate other data like forces and torques. [89, 93-95] A more novel form of collecting biomechanical data is the use high deflection strain gauges. Their use and prevalence has increased as the use of wearable sensors and technologies have become more and more popular. [66-68, 74, 75, 96, 97]

Phenotyping motion also applies to changes in patterns of motion due to muscle fatigue. Studies have shown that kinematic patterns of movement change when muscles are in a fatigued state. [98-103] Changes in movement are manifest in the order that muscles are activated (i.e., muscle recruitment pattern), along with alterations in kinematics and timing of movements. [103-105] Alterations like these
can contribute to improper recruitment of the muscles, ligaments, and tendons resulting in injury.

Adding to the difficulty of the challenge, motion phenotypes can be confounded by residual pain avoidance behaviors that can persist after the underlying tissue damage has healed (i.e., learned behaviors), or by psychosocial factors related to perception and treatment of the person with chronic low back pain. Body size and weight are also assumed to play a role in motion patterns. Gender-specific differences in both anatomy and learned motion behaviors may also play a role. To date, these confounding factors have not been well-characterized.

2.5 SPINE Sense System

The SPINE Sense system is an inexpensive, portable system used to capture vertebral motion that interfaces directly with a smartphone application.[75] It is comprised of an array of 16 nanocomposite strain gauge sensors that are mounted on kinesiology tape (e.g., KT Tape) and is shown in Figure 2.1. Electrical connection is made to the sensors through a piece of fabric with a sewn wire path that allows for stretchable electronics that connects to the sensors through metal clothing snaps. The device is adhered to the skin of the lower back and collects skin strain data from the nanocomposite sensors. This skin strain data can be used to phenotype motions from symptomatic subjects suffering from chronic lower back pain and asymptomatic subjects.[68, 97]

![Figure 2.1: The SPINE Sense system with exposed nanocomposite strain gauge sensors.](image-url)

The device’s sensor positioning and alignment on the KT Tape was optimized based on lumbar skin strain data obtained using standard camera-based motion capture of an array of 48 optical markers which were placed on the skin superficial
to 30 healthy patients while performing 17 distinct spinal motions. A lasso regression technique was used to extract information about where the sensors could be placed to capture the most important features of spinal motion during those specific 17 motions or exercises. The area that the sensors cover ranges from the T12 vertebrae down to S2 with the L5 vertebrae being used as a landmark to align the array with a corresponding mark on the device. Importantly, this motion has been shown to be accurate in measuring both overall lumbar motion, as well as segmental-level spinal motion.

The device will eventually be used as a diagnostic tool for people diagnosed with non-specific lower back pain as an alternative to expensive imaging options like X-Ray, magnetic resonance imaging, X-ray fluoroscopy, and other static imaging techniques. This device has been developed in hopes to help alleviate some of the burden that undiagnosed lower back problems have contributed to the opioid epidemic by informing victims of this of the cause of their lower back pain and what can be done to help them.
3 Wearable Strain Sensor-Based Detection of Lumbar Fatigue Using Machine Learning

3.1 Introduction

Low back pain (LBP) is the most disabling condition in the world, resulting in more disability years lived than another other condition. However, for the majority of those with either acute episodic low back pain or chronic low back pain, not every day is the same. Acute flareups of low back pain are distinct from low baseline levels of pain, and demand radical changes in activities of daily living[1, 7], higher doses of pain medications including opioids[12-16], tie up medical practitioner time and resources [9-11], result in lost workdays [1, 7, 10], and can trigger significant psychosocial consequences [107, 108].

Preventative measures for lower back pain are easier, more cost and time efficient, and less painful to implement than reactive ones [23-26]. There is also evidence to suggest that prevention of acute LBP episodes could reduce overall transition rates of patients from acute LBP to chronicity [27-29]. A particularly attractive way to prevent acute flareups of LBP would be to detect when an individual’s back muscles are fatigued and trigger an early warning alert indicating that they are at an increased risk of injury [52, 54-56, 60].

Large general population studies have shown that substantial portions of the population exhibit characteristics on MRI that could be indicative of structural instability in the lumbar spine.[109-112] However, only a relatively small fraction of these individuals present as symptomatic. The reasons for this discrepancy are complex and likely multimodal[113-115], however spinal musculature plays a contributing role. In the absence of structural stability provided by the passive tissues of the spine (i.e., in the absence of healthy and robust spinal discs, facet capsules, and spinal ligaments), the core and paraspinal muscles are known to take an increased role in active stabilization of the spinal segments.[54, 57] When these muscles become fatigued, both passive and active spinal stabilization mechanisms can fail, resulting in higher susceptibility to acute flareups of LBP.

Muscle fatigue is the decrease in the muscle’s force output and the time of response. People are at a greater risk of injury when their muscles are fatigued because the fatigued muscles can introduce changes in patterns of movement and mechanics of the movements. [99, 100, 102-104, 116] These alterations in patterns of movement can manifest themselves in altered kinematics and adjustments in muscle recruitment patterns. Bonato, Dingwell, Descarreaux, and others have found that fatigue of the muscles in the lumbar spine result in a reduction in hip flexion and an increase in range of motion and trunk lean associated with lumbar
muscle fatigue. These kinematic changes are likely correlated with the flexion relaxation phenomenon found in muscles due to their length tension characteristics.[99, 100, 104, 117] Decreased stabilization of joints and other biomechanical structures like the spine can also be an effect of muscle fatigue. [50, 51, 118] These changes in movement and kinematics can be measured and classified as fatigued and prefatigued.

Movement phenotyping has been used in the field of biomechanics for decades. Many features of the mechanics and kinematics of movement are measured and tracked to determine the difference between patterns of movement from different groups of subjects like symptomatic and asymptomatic or fatigued and prefatigued.[89-91, 119] These features include joint angles, position, velocity, angular velocity, acceleration, jerk, muscle forces and torques, etc.

Recent work has been done to phenotype motion of the spine with wearable skin mounted strain sensors.[85] To date, skin strain-based motion phenotyping has been reported in the context of distinguishing between specific motion patterns or distinguishing chronic low back pain motion from that of asymptomatic subjects. However inexpensive, wearable techniques for identifying changes in a particular individuals biopsychosocial state holds promise for non-invasive monitoring of health, particularly in the context of conditions such as low back pain, which can be aggravated or moved from baseline to acute based on muscle fatigue.

In the present work, we evaluated the utility of an inexpensive, skin-mounted wearable array of stretch sensors in identifying changes in low back motion consequent to paraspinal muscle fatigue. The SPInal Nanosensor Environment (SPINE Sense system) is comprised of an array of 16 nanocomposite strain gauge sensors that are mounted on a multi-directional kinesiology tape. The device is adhered to the skin of the lower back and collects skin strain data from the nanocomposite sensors. The utility of this system for distinguishing the motion of chronic LBP subjects from that of healthy controls has previously been reported.[67, 68, 75, 97] The purpose of this study was to test the abilities of the SPINE Sense system to detect changes in spinal motion in a cohort of 30 asymptomatic adults due to muscle fatigue.

3.2 Methods

In this study, 30 volunteers (18 males and 12 females) between the ages of 18-35 were recruited. The rationale for this age range was to target an age group that is at lower risk of injury performing exercises and during an endurance test required in the protocol. Sample size was selected based on a power analysis (G*Power 3.1, ANOVA: Repeated measures, within factors, medium effect size, alpha of 0.05, power of 0.8, 10 measurements per subject) as well as typical study sizes for similar previously published work.[118, 120, 121] Exclusion criteria for participants in the study included people with a history of lower back pain or prior surgery on the back, knees, or hips. The demographics of the patients, including age, height, and weight, are summarized in Table 3.1. The study protocol was approved by Brigham Young University’s Institutional Review Board. Participants provided their written and informed consent prior to participation in the study.
Table 3.1: Summarizes the mean and standard deviation values of the demographics of the participants in the study.

<table>
<thead>
<tr>
<th>Number of Participants</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/Female Participants</td>
<td>18/12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patient Information</th>
<th>Average Value (± std dev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>23.6 (± 2.5) years</td>
</tr>
<tr>
<td>Height</td>
<td>69.2 (± 4.8) inches</td>
</tr>
<tr>
<td>Weight</td>
<td>162 (± 33.3) lbs</td>
</tr>
</tbody>
</table>

Participants were informed about the details of the study and what would be required of them, including instruction on how each functional movement and the endurance test would be performed. They were then invited to give their consent and fill out the participant information sheet that collected basic information about their age, weight, height, and gender. The skin of the participant’s lower back was prepared to have the SPINE Sense System adhered on by removing the hair on the back if necessary. The participants had their lumbar spine palpated by an experienced and qualified individual who marked their L5 vertebrae as a landmark to place the SPINE Sense system. The system was then applied to the back with Cramer Tuf-Skin adhesive spray.

The device’s sensor positioning and alignment on the KT Tape was optimized based on lumbar skin strain data obtained using standard camera-based motion capture of an array of 48 optical markers which were placed on the skin superficial to 30 healthy patients while performing 17 distinct spinal motions.[66, 67] A lasso regression technique was used to extract information about where the sensors could be placed to capture the most important features of the spinal motion during those specific 17 motions or exercises. The area that the sensors cover ranges from the T12 vertebrae to S2 with the L5 vertebrae being used as a landmark to align the array with a corresponding mark on the device as shown in Figure 3.1.

Each participant performed the 14 functional movements listed below, 6 repetitions each. These exercises were chosen because they include motion of the spine in all 3 kinematic planes of motion (shown in Figure 3.1) as well as combinations of those movements which includes all kinds of motion the spine can undergo in 14 exercises. Impedance data was collected from the SPINE Sense System to capture the motion of the lower back through strain of 16 nanocomposite strain gauge sensors at a rate of 50 Hz per sensor.

- Knee Touch
- Maximum Extension
- Extension Fast
- Extension to the Right
- Extension to the Left
- Rotation to the Right
- Rotation to the Left
- Lateral Bending to the Right
- Lateral Bending to the Left
- Flexion to the Right
- Flexion to the Left
- Sit to Stand
- Maximum Flexion
After all the exercises were complete the participant completed the Biering-Sørensen endurance test. This test has been proven to fatigue the lower back muscles. The test involves the participant supporting their hips and legs in a horizontal position while holding their trunk and upper body in an unsupported horizontal position shown in Figure 3.1 until they can no longer hold or reach 4 minutes. The time that the participant was able to achieve was then recorded on their information sheet. The participant then immediately began performing the 14 functional exercises again, so that the motion data captured from the 16 nanocomposite strain gauge sensors pre-fatigue could be compared to the data captured post-fatigue (i.e., post Biering-Sørensen endurance test). No rest time was given to ensure that the participant completed the exercises in a fatigued state. Studies have shown that there is a minimum 10-minute window after the exercise that the muscles remain adequately fatigued and the body has not recovered. When the data collection was complete the array was removed with the Mueller Tape and Tuffner remover spray.

Figure 3.1: a) Photo representation of the SPINE Sense system. b) Photo of the SPINE Sense system attached to the back with the smartphone application screen and electronics attached. c) Demonstration of movement in the 3 planes of motion that are included in the 14 exercises performed by participants in the study. d) The unsupported horizontal position held during the Biering-Sørensen endurance test.
Figure 3.2: An example of resistance data from one of the 16 nanocomposite strain gauges for 6 repetitions and an expanded view of one repetition with the significant points that were used as features numbered and explained in Table 3.2.
Table 3.2: Explains the name and significance of the important points or features extracted from the data.

<table>
<thead>
<tr>
<th>Point Number</th>
<th>Name</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Starting Trough</td>
<td>The starting resistance value.</td>
</tr>
<tr>
<td>2</td>
<td>First Peak</td>
<td>The maximum peak resistance value as the subject is moving into the exercise.</td>
</tr>
<tr>
<td>3</td>
<td>Trough</td>
<td>The minimum resistance value at which the subject is fully into the exercise.</td>
</tr>
<tr>
<td>4</td>
<td>Second Peak</td>
<td>The peak resistance value as the subject is moving out of the exercise.</td>
</tr>
<tr>
<td>5</td>
<td>Ending Trough</td>
<td>The resistance value at the end of the repetition.</td>
</tr>
<tr>
<td>6</td>
<td>Time In</td>
<td>The time taken to get into the exercise.</td>
</tr>
<tr>
<td>7</td>
<td>Time Out</td>
<td>The time taken to get out of the exercise.</td>
</tr>
<tr>
<td>8</td>
<td>Delta R</td>
<td>The difference of the trough minus the starting resistance values.</td>
</tr>
<tr>
<td>N/A</td>
<td>Rdot</td>
<td>Delta R divided by time in. Represents the rate of change of the resistance.</td>
</tr>
</tbody>
</table>

The electrical resistance data from the 16 nanocomposite stretch sensors for each of the exercises pre- and post-fatigue was obtained for each of the participants and passed through a Hampel filter to remove any outliers. An example of a resistance curve for six repetitions from one sensor during one movement series is shown in Figure 3.2. The peaks and troughs of the curves indicate changes in the movement of the participant within an exercise and are key points of the exercise. Every peak and trough were manually identified and recorded for each exercise from each participant using MATLAB. From this data the values pointed out in Figure 3.2 and described in Table 3.2 were recorded for every subject for each exercise and each repetition from each sensor. These points were imported as features for a machine learning model used to classify exercises as fatigued or not fatigued in Python.

All data was normalized to account for the single subject use of each SPINE Sense system (i.e., since the device is attached to the skin of the subject, it is not transferable to another subsequent subject). This normalization process also helped account for the viscoelastic effects of the sensors on the resistance readings of the sensors.[72] Normalization was accomplished by using the Delta R and Trough values from the Knee Touch exercise from each participant (Equation 3.1).

\[
Normalized\ Data\ Value = \frac{Data\ Value - Trough}{\Delta R}, \tag{3.1}
\]

The data was organized as shown in Figure 3.3 and Figure 3.4 with the first column binarily indicating pre or post fatigue. The features include the values from each sensor for features 1-5, 8, and Rdot from Table 3.2 and the times into and out of the exercises. There are also 14 features, one for each exercise, in which one hot encoding was implemented to indicate which exercise the data for that row was from. One hot encoding is a proven technique used when integer categorical variables are replaced with new binary variables for each category.[123] Different iterations of the models were tested to include patient information like age, height, weight, and gender that was recorded and used as features. Models without the
The random forest classification model was trained and cross-validated using two different approaches to test the robustness of the model. One way that is referred to as hold-out cross-validation is to randomly select 70% of the data as the training set and reserve the remaining 30% to test the accuracy of the model. This is illustrated in Figure 3.4 with the yellow indicating the training data set and the gray indicating the test data set. For this set each repetition excluding the first repetition for each subject and each exercise was included as a row in the data organization found at the bottom of Figure 3.3. The model was run or cross-validated 1000 times to vary the random sampling each time and the accuracy was averaged over those repetitions.

The other method of training the model is illustrated in Figure 3.4 and is known as leave-one-out cross-validation. For this model the feature values from all the repetitions, excluding the first one, were averaged together and input as one row in the data organization illustrated in the bottom of Figure 3.3. The model was run 30 times with one subject being left out each time and used as the testing data set. The accuracies were averaged to get an overall accuracy of the approach. This method was implemented to mimic a new subject receiving the device and their data is thrown into the model without any of their data being included in the training of that model.

For each of these approaches a range of number of estimators was applied to the model and the one that resulted in the highest accuracy was recorded. SHAP Values were also recorded for the leave-one-out approach to determine which features had the highest impact on the classification decisions of the model.

An additional ablation study was performed to determine which exercises had a high impact on the model. This was done by withholding one exercise at a time from the model and then withholding groups of exercises. Those groups of exercises included all extension exercises (i.e., Maximum extension, extension fast, extension to the right, and extension to the left), rotation exercises (i.e., Rotation the left and rotation to the right), lateral bending exercises (i.e., Lateral bending to the left and lateral bending to the right), and flexion exercises (i.e., Maximum flexion, flexion fast, flexion to the right, and flexion to the left). Then a paired t-test was performed to test if there was a significant change in the accuracy due to the elimination of one of the exercises or a group of exercises.
Figure 3.4: Example of how the data was organized for the hold-out cross-validation approach of the random forest classification model and illustrates the way the model was trained with 70% of the data in yellow used to train the model and 30% in the gray being used as test data.

Figure 3.3: Example of how the data was organized for the leave-one-out cross-validation approach of the random forest classification model and illustrates the way the model was trained with one subject set aside as the testing data set while the model was trained on all other subjects’
3.3 Results

The accuracy results of the hold-out and leave-one-out approaches are outlined in Table 3.3. In both cases, the baseline accuracy, corresponding to the most frequent class, is 50%. For comparison, K-Nearest Neighbor, Logistic Regression, and Support Vector Machine were run using the same training and testing approaches as shown in Figure 3.3 and Figure 3.4. While all models were significantly above baseline, Random Forest Classification gave the best results overall, yielding a classification accuracy of 84.22% for the hold-out model, and 78.37% for the leave-one-out model.

Table 3.3: Accuracy results from various machine learning methods for each of the approaches of organizing the training and testing data.

<table>
<thead>
<tr>
<th>Machine Learning Approach</th>
<th>Hold-Out Model (%)</th>
<th>Leave-One-Out Model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest Classification</td>
<td>84.22</td>
<td>78.37</td>
</tr>
<tr>
<td>K Nearest Neighbor</td>
<td>71.02</td>
<td>73.62</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>75.92</td>
<td>64.20</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>78.43</td>
<td>67.85</td>
</tr>
</tbody>
</table>

As expected, accuracy results on the hold-out model are higher than results of the leave-one-out model, since for all test instances of any subject, that model would have had access to some data from each subject during model training. Hence, the accuracy of the hold-out model tends to be somewhat over-stated. Furthermore, the hold-out model is less actionable in a clinical setting, requiring “on the fly” updating of the machine learning model based on measured data from the subject. The leave-one-out model, on the other hand, although it has slightly lower accuracy, is directly applicable since it can make relevant predictions about new subjects who are not included in the training dataset (i.e., it can correctly identify fatigue status based solely upon the pre-existing model without additional training).

A feature analysis was used to determine the most important features and sensors that the model used in the random forest classification decisions. Shapley values were attained for the cross-validated results of the leave-one-out method. Shapley values, or SHAP values (SHapley Additive exPlanations), are a popular way of determining feature importance based on a calculation that incorporates game-theory to determine the impact each feature has on the model.[124] The results are illustrated in Figure 3.5 and Figure 3.6, where the most important features are in descending order and the horizontal location of the dots (i.e. the “tails” of the spread) indicates the impact, positive or negative, that feature has on the model. The colors of the dots shows whether the variable for that observation was a high or low value by the colors red and blue respectively. For example, for the top
feature of the starting trough value from sensor 5, low values for this feature have a positive impact on model’s classification as fatigued.

The ablation study that was performed resulted in two different sets of exercises that had a significant increase in the accuracy when they were withheld from the model. When flexion left and all flexion exercises were omitted from the model, there was an increase in the accuracy of the model. Running a paired t-test confirmed that there was a statistically significant change in the mean of the accuracy when the flexion left exercise and all flexion exercises were withheld.

Figure 3.5: The SHAP values of the top 20 features with the most impact from the leave-one-out model where the number next to the feature title indicates the sensor number of that feature.
There was a 0.917 % increase in the accuracy of the model when the flexion left exercise was withheld and a 2.256 % increase when the group of all flexion exercises was withheld. This resulted in p values of 0.028 and 0.012 respectively for the paired t-test.

3.4 Discussion

The stretch sensors used in the SPINE Sense system were developed and have been extensively validated in previous work. [65-69, 72-75, 97] The strain sensors are composed of nickel nanostrands (NiNs) and nickel-coated carbon fibers (NCCF) uniformly distributed in a silicone matrix. The silicone base of the sensors allows for a strain gauge to repeatedly achieve extremely high strains typical of biological tissues such as the skin.[69] During deformation, the electrical resistance of the strain sensors drops dramatically, yielding accurate and repeatable measurement of a wide range of strains from 0 to over 100%.

Figure 3.6: Mean SHAP Value bar plot with the most impactful features in descending order.
The purpose of this study was to test the ability of the Spine SENSE system to detect the difference between fatigued and not fatigued motion using machine learning models. The strain data from the 16 sensors on the Spine SENSE device from the 14 exercises were enough to train the model accurately classify motions as fatigued or not with an accuracy of up to 84.22%. It is important to note that the manner in which the subjects’ muscles were fatigued was done using an endurance fatigue test. Results may vary if the muscles are fatigued done using a strength fatigue test.[125] This is due to the fact that muscles with higher proportions of Type II (fast twitch) muscle fibers will fatigue more quickly, particularly under strength training conditions, leading to altered recruitment patterns as compared to fatigue training conditions which rely more on the Type I muscle fibers (slow twitch). Consequently, distinct fatigue tests are very likely to elicit overlapping but distinct changes in both muscle recruitment and movement patterns. Similarly, differences in muscle fiber populations between individuals even within the same muscle groups, can elicit distinct changes in fatigue-related muscle recruitment and motion alterations. Given these differences, it is encouraging to note the achieved accuracy of the models in the present work.

The two ways of training the model, hold-out and leave-one-out model, were implemented to simulate different situations of future users of the device receiving the device and what data would be needed from them to be put into the model. The hold-out model represents a scenario in which the new user of the device would need to collect data pre and post fatigue as a calibration step to be added to the model before the device could be used to classify their data. Whereas the leave-one-out model represents a case where the new user of SPINE Sense would just need to start using the device and their data would be classified based on the training the model has already received from other users.

Both models were successful in classifying between fatigued and not fatigued states with accuracies of 84.22% and 78.37%. It is hypothesized that data from more subjects may create a larger training set that could potentially result in higher accuracies as including more subjects could train the model on more biomechanical variation that naturally occurs between humans and their unique movements.

According to the SHAP values, the features with the highest impact include the Starting Trough, First Peak, and Second Peak values all from sensor 5 and the Starting trough values from sensor 0. All of these values have low values with a positive impact on the classification decision of the model according to the SHAP value summary plot provided in Figure 3.5. An example of a pre fatigued repetition
Figure 3.8: Shows the approximate location of the sensors in relation to the vertebrae represented by the labeled rectangles.

Figure 3.7: An example of the resistance curve for pre and post fatigue from one sensor for one repetition from one exercise.
compared to a post fatigued repetition is shown in Figure 3.8. It was suspected that the times into and out of the exercise would have a high impact on the model. Both of those features appear in the top 20 most impactful features where their high values have a negative impact on the classification decision. As for the other features that appear in the top twenty, all features derived from a sensor resistance value have low values with a positive impact. This is true for all except for the starting values from sensor 9 where high values have a positive impact on the model.

Biomechanically, the lower values for sensor 5 in the starting trough and peak values is indicative that the skin is strained more during the post-fatigue movements in the region where sensor 5. A distribution of the starting trough values from sensor 5 pre and post fatigue is shown in Figure 3.9 where a visible shift in the resistance values of the peaks can be seen. The location of the sensors on the back in relation to the vertebrae is shown in Figure 3.7. As show in the figure, sensor 5 is in the lumbar region of the back near the L4 and L5 vertebrae. Additional skin strain is indicative of the spine being in additional flexion than it was in its pre-fatigue state. This finding reaffirms conclusions from other research studies that found an increase in trunk lean, a decrease in hip flexion, and increased range of motion in the lumbar spine due to muscle fatigue.[99, 100, 104] It is correlated with the length tension characteristics of muscles to generate more tension or force when they are in a lengthened or stretched state, but also unfortunately indicative of instabilities in the spine that could lead to LBP problems. [51, 104, 117, 118]

These conclusions and logic apply to all features in the top 20 that are derived from sensor resistance values except for the high values from sensor 9 that have a positive impact on the model’s classification decision. It is noted that the starting value of sensor 9 is the only feature that comes from the horizontal set of sensors.

![Histogram of Pre vs Post Fatigue Data of Starting Trough Value from Sensor 5](image)

**Figure 3.9:** Histogram that shows the distribution of the starting trough values from sensor 5 pre and post fatigue from every exercise and every subject.
closer to the thoracic region of the spine and may be indicative of the same findings Bonato found where subjects resorted to an alternative lifting method that required less trunk bending post-fatigue.[99] With less trunk bending, the skin of the thoracic region of the spine would have less strain and consequently higher strain values like those of sensor 9.

It was assumed that trough values would be impactful to the model because those features are strain values from when the subjects are fully extended into movement and often times at their end range of motion, but only the trough values from one sensor was in the top 20 most impactful features. It was also assumed that Rdot, and DeltaR values would be a greater indicator of fatigue because there is an element of time in their calculations, but none of these features appeared in the top 20. This may indicate that the rate at which the skin is strained has a less significant impact than we thought, although previous work on the sensors has proven that the magnitude of the peak resistance values is indicative of the rate at which the sensors are being strained.[72] The rate at which the exercises are being performed is also captured in the time into and out of the movements for which high values have a negative impact on the model’s classification. In other words, the more time that was taken the more likely that the model will classify it as pre-fatigued. This presumably could be due to the fact that the subjects were more comfortable with the movements the second time performing them post-fatigue because of their increased familiarity with the protocol.

The results of the ablation study to determine which exercises impacted the accuracy of the model showed that the flexion left and all flexion exercises being withheld from the model increased the accuracy of the model. This indicates that the flexion exercises do not contribute as much as expected to the model’s ability to accurately classify fatigue. It is possible that the variation between the individuals is higher in magnitude than the actual effects of the fatigue for those particular exercises, so the effects are not as detectable within the noise between individuals for those particular exercises.

It is concluded that the SPINE Sense device has a refined enough ability to phenotype motion as fatigued or not. Future work should be done to expand the data set to a larger population and include people of a larger age range. Additionally, investigating different ways in which the muscles are fatigued (i.e. applying a strength fatigue test rather than an endurance fatigue test as done in this study) will expand the abilities of the device to detect fatigue of the muscles in different forms. Extracting information about which sensors have the highest impact on phenotyping of motion may also be informative to the design team of the SPINE Sense system. This research has started the exploration of the capabilities of this device in being used as a tool to combat lower back pain. This can be done by using the device as a tool to detect alterations in the kinematics and biomechanics of a person’s movements due to muscle fatigue and alert users to exercise caution when they are at this state of elevated risk of lower back injury.
4 SPINE Sense System Usability

As a part of the development of the SPINE Sense system, the usability of the system was tested in effort to gauge the patient and clinician perception of the device. The results were informative to the team in making educated and needed design decisions catered to potential users in the further development of the device. The widely used System Usability Scale was employed to quickly assess the usability of the system. Potential users of the device, like patients with chronic lower back pain and clinicians, were the individuals chosen to assess the usability of the system.

4.1 Introduction

The SPINE Sense system is an inexpensive, portable system used to capture vertebral motion that interfaces directly with a smartphone application. It is comprised of an array of 16 nanocomposite strain gauge sensors that are mounted on Kinesiology Tape (KT Tape®). The device is adhered to the skin of the lower back and collects skin strain data from the nanocomposite sensors.

The high deflection strain gauge sensors used on the SPINE Sense device are a product and development of the BABEL lab at Brigham Young University and have been tested extensively. They are composed of Nickel Nanostrands (NiNs) and Nickel Coated Carbon Fibers (NCCF) distributed in a silicone matrix. This provides strain gauges that can capture high deflection strains that are often seen in biomechanical applications like that of the strain of the skin. As the sensors stretch the resistance decreases and allows for repeatable strains over 100%.

The device’s sensor positioning and alignment on the KT Tape was optimized based on lumbar skin strain data. They were obtained using standard camera-based motion capture of an array of 48 optical markers which were placed on the skin of 30 healthy patients while performing 17 distinct spinal motions. A lasso regression technique was used to extract information about where the sensors could be placed to capture the most important features of the spinal motion during those specific 17 motions or exercises. The area that the sensors cover ranges from the T12 vertebrae to S2 with the L5 vertebrae being used as an anatomical landmark to align the array with a corresponding mark on the device.

A study was conducted to test the usability of the SPINE Sense device and receive feedback on its design to make design changes in accordance with user...
defined requirements. The System Usability Scale (SUS) was designed as a “quick and dirty” way to quantify and test the usability of the system and has been reviewed and compared against other usability scales. [127-129] It has been well established and widely applied to many products, devices, services, and websites since its creation in 1996.[126, 128] Modifications and other slight changes have been validated and tested to change the word choice and phrasing to make it more applicable for different classes of systems and devices. [126, 128]

4.2 Methods

A system usability study was conducted employing SUS to quantitatively assess clinician and patient perceptions of the usability of the design in the context of clinical practice. As a part of the study the physical prototype shown in Figure 0.1 of the SPINE Sense system was presented to 32 potential users of the system namely clinicians (n=19) and chronic low back pain patients (n=13). The prototype and a scripted explanation (Appendix B) of the system including its application to the lower back and its integration with the associated smartphone app was presented to the participants. They each completed an SUS survey afterwards, which is an industry standard set of scaled statements used to assess the usability of a system that employs a Likert scale to calculate the overall score.[126, 128, 129]

Figure 4.1: The version of the SPINE Sense system that was presented to potential users of the device for the System Usability Study.

The participants were then asked open-ended questions regarding the design of the device and had the opportunity to voice ideas, concerns, and questions. These comments were audio or video recorded and later reviewed and tabulated.
The comments and feedback were then taken back to the design team of the SPINE Sense System and used in upgrades and design improvements made to the device.

4.3 Results

The possible scores of the SUS are 0-100. A score above 68 is considered above average and represents a device that is intuitive to use. [127] The mean SUS score for this system usability study was 72.03 (“above average”) and a more specific breakdown based on respondent is presented in Table 0.1.

Table 4.1: Breakdown of the System Usability Scale scores from the system usability study.

<table>
<thead>
<tr>
<th>Item</th>
<th>Clinicians</th>
<th>Patients</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>19</td>
<td>13</td>
<td>32</td>
</tr>
<tr>
<td>Mean Score</td>
<td>70.79</td>
<td>73.85</td>
<td>72.03</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>13.74</td>
<td>6.69</td>
<td>11.51</td>
</tr>
</tbody>
</table>

Common positive feedback from the participants included how they liked the simplicity of the system along with the flat compact design. The ease of application of the device with the Kinesiology Tape was also seen as a big positive point of the device as it is a material and product with which a majority of clinicians and patients are familiar. Participants also mentioned how they liked the ease of use of a smartphone app to receive data and feedback from the device with a wireless connection through Bluetooth.

Constructive feedback received was the concern about the system only coming as a “one size fits all”, when individual’s body shapes are vastly different. As a consequence the sensors could potentially be misaligned on the vertebrae of the spine and yielding inaccurate data. There were also numerous concerns about wires and their durability for applications on collegiate or professional athletes during in-sport movements and exercises. Other common comments included recommendations for specific alternative choices for the athletic tape substrate based on previous experiences with tackiness, duration of adherence, and ease of removal.

4.4 Discussion

The SPINE Sense System received an above average score from the SUS. Considering that this feedback is from a combination of clinicians and patients who will potentially use the system, it is presumed to have validity. It is concluded that the design of the wearable vertebral motion tracking system is attractive and usable for patients and clinicians.

Taking into consideration the comments and feedback from the system usability study, design changes were made to improve the user experience and aesthetics of the device. These changes included improving the electrical connections to provide more robust and reusable electrical connections shown in Figure 4.2. These changes evolved into a design that went from embedding crimped wire into silicone to sewing wire into the kinesiology tape to a design that incorporated two pieces. One piece has the 16 nanocomposite sensors mounted or glued onto the KT Tape®, and the other is a wire piece that has sewn enamel coated
copper wire (36 AWG) into a stretchy fabric to form stretchable and reusable wiring and electronics. These two pieces are integrated with the use of small metal clothing snaps that are mounted on the ends of each sensor with corresponding and opposite ends found on the wire piece to act as the electrical connection between the two pieces.

Other design improvements and changes include the use of one small micro-HDMI cable that connects to a custom micro-HDMI port mounted on the wire piece shown in Figure 4.3 as opposed to multiple wires coming from the device. Small pockets were also integrated into the array for storage of the battery and circuit board. The team also started adding identification tags, so each device had a number associated with it for tracking of the manufacturing pieces. These improvements are shown in Figure 4.3.

With the design changes implemented, the device was ready for data collection for various studies including the fatigue study. The future of this device is to be used as a phenotyping tool for people diagnosed with non-specific lower back pain to evaluate the best course of treatment for their back pain. It is a more economical alternative to expensive imaging like MRIs and X-Rays that only tell part of the problem because they capture data about the spine in a static position. Whereas the SPINE Sense captures dynamic skin strain data that will give information about the segmental motion of the spine that can help paint a better picture and give more information about the patient is moving and what course of treatment will best help them.

Future research should be performed to repeat the system usability study now that the design changes have been applied to the system. This will help assess and quantify the value the design changes have added to the system in terms of usability.

Figure 4.2: The change of wiring of the SPINE Sense device from embedding crimped wire in silicone to sewing wire into the kinesiology tape to the two-piece reusable wiring design.
Figure 4.3: Various SPINE Sense design improvements including a) the compact electronics design, b) the micro-HDMI cable, c) mounted micro-HDMI port on the wire piece, d) the full assembly with storage pockets and identification tags.
5 Conclusions and Future Research Directions

It is concluded that the SPINE Sense device and the wearable high deflection nanocomposite strain gauges are capable of collecting data that is refined enough to be classified as fatigued or not fatigued using machine learning models. Accuracies of 78.37 and 84.22 were achieved by random forest classification models with data from 30 subjects performing 14 exercises pre and post fatiguing of their back muscles.

SHAP values from the leave-one-out model revealed that features from sensor 5 like the starting trough values and peak values have a high impact on the model and its classification decisions. Low values of these feature have a positive impact on the fatigue decision. This means that the sensors are stretched more, and the resistances are at lower values indicating that the subject is in more flexion. This reaffirms the previous research that has found that fatigue causes an increase in trunk tilt, overall range of motion, and hip flexion.

Further work should be done to increase the number of participants and include older populations to test the system’s capabilities on populations that were not included in this study. Including more people in the study could capture and train the model on more biomechanical diversity that these populations have. Including patient identifiers may become a more important factor when a more diverse population is explored.

The manner in which the back muscles are fatigued should also be investigated. In the study performed in this work, the muscles of the back were fatigued using an endurance test. Other ways of fatiguing muscles would be to perform a strength fatigue test. The results would vary due to a difference in the proportion of Type I and II muscle fibers found in the muscles that are being fatigued. An endurance fatigue test, like the one performed in this work, will fatigue the Type I muscle fibers (slow twitch) more whereas a strength fatigue test will rely more on the Type II muscle fibers (fast twitch). This will cause a change in the muscles that are fatigued and the way in which people alter their biomechanics to compensate for the fatigue felt in these muscles. Investigating other ways in which the muscles are fatigued, like a strength fatigue test, would expand the devices ability to detect fatigue of different forms and improve its ability to be used as a tool to prevent injury and flare-ups of lower back pain that occur as a result of fatigued back muscles.

A more comprehensive look at the features that the machine learning models use the most could also aid in future decisions about the use of the device. Knowing this information could alter the exercises, sensors, and features that need to be captured in data collection. Capitalizing on the most important movements and features while omitting the ones that don’t reveal as much information could allow for a quicker easier data capture experience.
This dive into the features that impact the model the most should also include an investigation as to which sensors have the most impact as well. Preliminary work was done to quantify the occurrence of the sensors that occurred in the features with the top 20 Gini importance from each leave-one-out model. This work is shown in Figure 4.1 where sensor 12 was found to be the most reoccurring sensor in the features with the highest Gini importance’s. This preliminary work is intriguing but should be expounded upon with other importance or impact values. Once this work is done a correlation could be performed with the results from the sensors that impact the diagnosis of chronic lower back pain the most. Knowing which sensors impact the models the most could inform the design team as to how to reduce the number of sensors on the device.

**Top Occurrences of Sensors in Most Important Features in 30 Models**

![Image: Sensors that had the highest occurrence among the features with the top 20 Gini importance from each leave-one-out model.](image)

**Figure 5.1:** Sensors that had the highest occurrence among the features with the top 20 Gini importance from each leave-one-out model.

Work was also done to assess the usability of the device and design changes that were desired from potential users of the device. According to feedback and system usability scores from 32 potential users of the device, the system has received a score that concludes it is usable and intuitive to use. User feedback from these participants also aided the design team in making design improvements to the electrical connections, wiring, and electronics modules of the system. A two-piece design was implemented to allow for electrical wire connections to be reusable. Other design changes included the use of a micro-HDMI cable and a compact electronics module with a small battery pack for power.

As the design changes have improved and altered the device significantly, future work should include a new system usability study to capture a score on the new system and attain user feedback about further improvements. Future research should also include a look at the sensors themselves to achieve more consistent and repeatable performance across sensors and devices to allow for easier and more reliable data processing. Future work might also include design changes that would allow for the device to be removable and re-usable as the current device is a disposable device and cannot be used across subjects.

The capabilities of the SPINE Sense system and affiliated high deflection strain gauge technologies are effective in detecting changes in motion due to
altered recruitment patterns. The ability to detect fatigue has now been explored and demonstrated. Future work should be focused on refining the abilities and repeatability of the device and system itself to focus on a more user-friendly device that has the potential to improve and impact the lives of many people who suffer from lower back pain or have the potential to acquire lower back pain.
References


119. !!! INVALID CITATION !!! [76-84].


Appendix A  Subject Test Forms

Fatigue Spine Kinematics Study
Consent to be a Research Subject

Introduction
This research study is being conducted by Darci Billmire at Brigham Young University (BYU) to measure the spinal motion before and after the muscles of the back have been fatigued. You were invited to participate because you are a healthy adult with no current biomechanical disorders or injuries that might produce abnormal spinal motion.

Procedures
If you volunteer to participate in the study, we will ask you to do the following things:

- Participants should wear comfortable loose clothes they can move around in and if possible, their pants or shorts should have an elastic waistband.
- Provide your name, age, weight, height, and gender, and confirm that you do not suffer from any biomechanical disorders or injuries that might produce abnormal spinal motion, including any previous spinal injury or spinal surgery.
- Your lower back will be prepared for the array that has an adhesive backing to be placed (ie. shaving and skin preparation). The array will be placed on your lower back in the region between the bra line and belt line.
- An experienced and qualified individual will palpate for a lumbar vertebral landmark then align and apply the array with respect to that landmark. The researcher applying the sensors will be of your gender upon request.
- You will then perform the 13 functional exercises listed below in that order while we collect data from the SPINESense array. Common names or descriptions for these movements are as follows: flexion - bend forward, extension - bend backwards, lateral bending
- bend to the side, and rotation - rotating shoulders to the left or right while hips stay in place.
  o Knee Touch
  o Maximum Extension
  o Extension Fast
  o Extension to the Right
  o Extension to the Left
  o Rotation to the Right
  o Rotation to the Left
  o Lateral Bending to the Right
  o Lateral Bending to the Left
  o Flexion to the Right
  o Flexion to the Left
  o Sit to Stand
  o Maximum Flexion
  o Flexion Fast

- Perform what is called the Biering-Sørensen endurance test. For this you will have your legs and hips supported by specifically designed table or chair and you will hold your upper body in a horizontal unsupported position until you are fatigued or reach 4 minutes. See the picture below for an example.

- The same 13 functional exercises will then be performed again.
- The array of motion sensors will then be removed.
- This whole procedure will take about 45 minutes.

**Risks/Discomforts**
There are minimal risks for participation in this study. However:
- You may feel minor muscle or soft tissue pain as a result of bending and twisting your back and performing the endurance test.
• You may feel discomfort as we remove the adhesive used to attach the sensors. (similar to removing a Band-Aid® or medical tape) We will attempt to reduce the discomfort by using a cotton ball or disposable towel soaked in baby oil to release the adhesive from your skin as we remove the array of sensors.

• You may have minor skin irritation due to the adhesive used to attach the sensors if you have sensitive skin.

• There is a risk of injury while performing the 13 functional exercises and the endurance test. This risk is minimal and not frequent as adequate instruction will be given to each participant on how to perform the basic exercises and endurance test.

If at any time, or for any reason, you wish to discontinue, you may withdraw yourself from this study.

Benefits
There will be no direct benefits to you for your participation in the research study. However, it is hoped that through your participation researchers will gain a deeper understanding of the sensors’ capabilities of differentiating between motion when the muscles are fatigued and not fatigued.

Confidentiality
All the data from this project will be digitally de-identified and stored on password protected computer files. When the data analysis has been completed and ten years have passed (whichever happens last), the data will be deleted. Only de-identified movement-related data (no personal identifiers) will be published or shared with other research institutes.

Data Sharing
We will keep the information we collect about you during this research study for analysis [and for potential use in future research projects]. If the study data contain information that directly identifies subjects: Your name and other information that can directly identify you will be stored securely and separately from the rest of the research information we collect from you. De-identified data from this study may be shared with the research community, with journals in which study results are published, and with databases and data repositories used for research. We will remove or code any personal information that could directly identify you before the study data are shared. Despite these measures, we cannot guarantee anonymity of your personal data.

Participation
Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate entirely without jeopardy to your standing with the university.
**Research Locations**
The research visits will take place at one of the following locations:
- The Engineering Building (BYU Campus), Provo, UT 84602
- The Smith Fieldhouse (BYU Campus), Provo, UT 84602

**Questions about the Research**
If you have questions regarding this study, you may contact Anton Bowden at (801)-422-4760 or abowden@byu.edu.

**Questions about your Rights as Research Participants**
If you have questions you do not feel comfortable asking the researcher, you may contact: Human Research Protection Program; Email: BYU.HRPP@byu.edu

**In Case of Research Related Injury**
BYU makes no commitment to provide financial compensation or free medical care should you be injured as a result of your participation in this research. Nonetheless, in the event of such an injury, after seeking appropriate medical attention, please contact Anton Bowden at (801)-422-4760 or abowden@byu.edu.

**Statement of Consent**
I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Name (Printed): ________________________________

Signature: ________________________________

Date: ________________
Fatigue Spine Kinematics Research Subject Information

Date: __________

Subject Name: ________________________

Age: _______years

Weight: ___________pounds

Height: ______feet and _______inches

Gender:  M  F  Other  Prefer not to say

Do you suffer from any biomechanical disorders or injuries that might produce abnormal spinal motion, including previous spinal injury or surgery? (circle one)

Yes  No

If yes, please explain:

Everything below is filled out by research staff:

Biering-Sørensen Endurance Test Time:

_________Minutes ___________seconds

General Notes:
Appendix B  System Usability Study Test Plan

System Usability Scale Test Plan
BYU Applied Biomechanics Engineering Laboratory – BACPAC Group
By: Darci Billmire

**Scope:** The scope of this project is to test the usability of the BACPAC group’s version 4 prototype. The usability will be tested by conducting interviews with clinicians and patients who will likely be using the product, and having them complete the System Usability Scale with respect to the usability of the prototype.

**Purpose:** The purpose of this project is to identify usability, concerns, and questions from potential users of the product. This will be achieved through an interview and questionnaire conducted with potential users of the prototype.

**Time and Location:** TBA

**Sessions:** The sessions are designed to last between 10 – 20 minutes. The session will start with an expression of gratitude for participation and disclaimers on recorded sessions and the participants ability to drop out of the study at any time. The session will then be a guided discussion explaining the purpose of the prototype and how it is used while asking scripted questions to receive comments, opinions, and feedback about the usability of the prototype. The session will end with having the participant complete the System Usability Scale, which includes 10 questions responded on a 5-point Likert scale of Strongly Disagree – Strongly Agree.

**Equipment:** The equipment used will include prototypes for each of the participants to hold and handle. Other equipment needs include: a laptop, video camera, extra batteries or chargers, and disinfectant wipes to clean between sessions.

**Participants:** The goal is to interview 30 potential users of the prototype. These potential users include clinicians and patients with chronic lower back pain.

**Metrics:** The results of the System Usability scale will be scored according to the SUS scoring standard. This will result in scores within the range of 0 – 100. The lower acceptability limit for this project is the score of 68. The system will be considered usable if the average of the scores is between 68 – 100.

**Procedure:** The interviewer should follow the script below to conduct the interviewing sessions while taking notes and recording the session on video if given permission by the participant. The video will not capture the participant’s face, only their handling of the prototype and the audio of their responses. If possible, perform the interview at a table with the prototype placed in front of the participant, and the video set up with the prototype in frame. The interviewer is free to answer any questions or concerns that the participants have as well as
asking any additional probing questions to receive more feedback or clarify a response.

**Interviewer:** First, we would like to thank you for participating the System Usability Scale Test for the physical prototype of our SPINE Sense System. To better capture your comments and feedback, with your permission, we would like to video record this session. Do we have your permission to record this discussion?

*Await response: start recording now, if given permission. Otherwise, take comprehensive notes during the discussion.

**Interviewer:** Current methods for measuring the biomechanical motion of the vertebrae of the spine are complex, time-consuming, and can be invasive. For example, the most commonly used method involves a time-consuming process of applying an array of motion capture markers to the back and requires the patient to travel to a motion capture laboratory where an array of expensive cameras measure the motion of those markers, which can be processed using specialized software. A visit like this requires 30-60 minutes of patient time, 1-2 hours of technician time, and requires hundreds of thousands of dollars in equipment. Another method of capturing this level of biomechanical data is to collect biplanar x-ray fluoroscopy data, exposing the patient to a significant radiation dose and requiring them to move within a very small space within the field of the x-ray sources. Our goal was to develop an inexpensive, portable system for capturing vertebral motion that could be used in a clinician’s office or in a patient’s home, and would interface directly with a free smartphone app, giving virtually instantaneous feedback on the motion of the patient which could be used as part of their diagnostic process. The entire data collection process only requires a few minutes. We would like to show you a prototype of our device, which we call the “SPINE Sense System” and ask you some questions about how easy you feel it will be to use. We would ask that you use the other systems we have discussed (motion capture and biplanar x-ray fluoroscopy) as reference systems when you consider your answers.

Do you have any questions so far?

We will start with some basic information about you.

**Interviewer:**

**IF PATIENT:** Interviewer: Could please state your name, age, and how long you have suffered with chronic lower back pain.

Name:

Age:

Duration of cLBP:

**IF PHYSICIAN:** Interviewer: Could please state your name, occupation and how many years you have been in this occupation.
Interviewer: This is a prototype of the SPINE Sense system. Feel free to pick up the device and handle it as you wish during our discussion. As you can see, the device is attached to the skin of the lower back using commonly used athletic tape (kinesiology tape). A doctor or therapist would apply the tape backing based on a few anatomical landmarks, and then turn the system on, and it would pair automatically with a nearby smartphone. The patient would then be asked to perform a few simple bending movements (for example, twisting right to left, bending over to pick up a box, rising from sitting to standing position, etc), that would requires about 5-minutes. Then the doctor would push a button on the app, which would then return an analysis of the motion data and a potential diagnosis for underlying sources of mechanical low back pain.

At this point, we would like you to complete this short, 10-question survey about the SPINE Sense.

Interviewer: Thank you! In addition to the survey data, we would like to ask some follow-up questions that may aid us in improving the design for future iterations. We would like open, honest, and candid feedback.

*For Clinicians* How might a system like this aid you in your work with patients of lower back pain?

*For Patients* If wearing something like this aided your diagnosis and possibly your treatment of chronic lower back pain, does it seem to be something you would likely try out?

Design:

Do you have any questions about the design of the device (its overall form and function)?

What do you like about the design of the device?

What don’t you like about the design of the device?

Use:

Do you have any questions about how the device is used?

What do you like about how the device is used?

What don’t you like about how the device is used?

Other Feedback:

What additional feedback would you give our design team?

*Stop Recording