# A Comprehensive mHealth System for Chronic Low Back Pain Assessment: Development, Evaluation, and Exploration for Future Works

by

**Zakiy Firdaus Alfikri**

Bachelor of Science, Bandung Institute of Technology, 2012

Master of Science, Bandung Institute of Technology, 2013

Submitted to the Graduate Faculty of the

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

**Zakiy Firdaus Alfikri**

It was defended on

March 13, 2023

and approved by

Kevin M. Bell, PhD, Assistant Professor, Department of Bioengineering

Brad E. Dicianno, MD, Professor, Department of Physical Medicine and Rehabilitation

Andi Saptono, PhD, Assistant Professor, Department of Health Information Management

Leming Zhou, PhD, Associate Professor, Department of Health Information Management

Dissertation Advisor: Bambang Parmanto, PhD, Professor, Department of Health Information Management

**A Comprehensive mHealth System for Chronic Low Back Pain Assessment: Development, Evaluation, and Exploration for Future Works**

Zakiy Firdaus Alfikri, PhD

University of Pittsburgh, 2023

Phenotyping chronic low back pain (CLBP) is essential for developing personalized and adaptive treatments for CLBP. To achieve this, a large amount of CLBP assessment data is needed. In this study, an mHealth system was developed for CLBP assessment. The system consists of an in-clinic app, at-home app, clinician portal, backend, and database, and focuses on collecting biomechanical and behavioral assessment data as part of extensive multifactorial data needed for CLBP phenotyping. The mHealth system was able to collect CLBP assessment data effectively and efficiently from both the structured in-clinic assessment and the assessment in the patients' daily life settings.

Usability evaluations were conducted to assess the usability of the in-clinic and at-home apps. Two usability evaluations were conducted for the in-clinic app, and several updates and revisions were made to address identified usability issues. In the last evaluation, the in-clinic app received a high usability score of 6.00 (SD=1.15). Meanwhile, for the at-home app, five usability evaluations, involving 337 CLBP patients, were conducted. Several updates and revisions were made to address the usability issues identified. In the last usability evaluation, the at-home app received a high usability score of 6.24 (SD=1.37).

Furthermore, two exploratory works for future direction were conducted in this study. The first was an investigation of the correlation between the perceived activity level that patients reported in EMA and the activity level calculated from accelerometer data from the kinematics

sensors. The overall correlation was found to be weak, ranging from 0.095 to 0.260 (mean=0.194, SD=0.054). Even though the overall correlation was weak, the correlation of activity level from sensor data and perceived activity level from 37.5% of CLBP patients was found to be strong. Using the five most accurate activity level representations, the average score for the correlation was 0.716 (SD=0.081), suggesting that some CLBP patients may have a better perception of their activity level. The other exploratory work done in this study was the development of a dataset builder component that was successfully be used to label motion data based on the videos recorded during the in-clinic assessment.

## Table of Contents

## List of Tables

## List of Figures

## Preface

I would like to thank everyone who has supported me during my PhD journey. First, I would like to thank my family for their endless love and support. My beautiful wife, Anindayu Pradetha, who has always been there for me, I love you and thank you for always pushing me to finish my PhD. I don't think I can get to this point without your support and love. To my parents, Sri Haryoko and Tri Bastuti, thank you for always supporting me and letting me study far away from home. I also want to thank my brothers and sisters, Syifa, Fadil, Yasya, and Ridhal. I couldn't wait to share this happiness with you all. Special thanks to my cats, Nautilus and Ocean, for keeping me company while I worked on my dissertation.

I would like to express my appreciation and gratitude to my advisor, Dr. Bambang Parmanto. Thank you for your guidance and advice during my time as a PhD student. To Dr. Kevin Bell, thank you for being a mentor and for always providing guidance and support for my dissertation work. To Dr. Brad Dicianno, Dr. Leming Zhou, and Dr. Andi Saptono, thank you for being my dissertation committee members and for providing your expertise and support that was very valuable for my dissertation work.

I would like to thank my colleagues and friends in HARI Lab. To Gede and Wayan, thank you for being my mentors and for inspiring me to achieve this. To Agus, Leon, and Firdaus, good luck with your work, I will root for you all. I also would like to thank my colleagues and friends in the LB3P research team. I am grateful to have worked with you and I am thankful for your support. To my friends in Pittsburgh and Indonesia, thank you for your support and friendship. It means a lot for me.

# 1.0 Introduction

## 1.1 Problem Statement

Chronic low back pain is one of the most prevalent health problems in all developed countries, including the US (Hoy et al., 2012; Koes, Van Tulder, & Thomas, 2006; Murphy et al., 2017; Waterman, Belmont, & Schoenfeld, 2012). According to the Centers for Disease Control and Prevention (CDC), in 2019, 39% of adults in the US reported back pain (Lucas, Connor, & Bose, 2021). Additionally, low back pain is one of the leading causes of years lived with disability globally (Hartvigsen et al., 2018; Vos et al., 2016; Wu et al., 2020). Despite its prevalence, chronic low back pain is still a challenging condition to effectively treat (Urits et al., 2019)

There are several treatments that can be used to treat low back pain (Delitto et al., 2012; Van Middelkoop et al., 2011). Selecting the right treatment or combination of treatments is essential for effectively and cost-efficiently treating low back pain (Delitto, Erhard, Bowling, DeRosa, & Greathouse, 1995; Murphy et al., 2017). However, deciding which low back pain treatment is best for a patient is still a challenging task (George, Goertz, Hastings, & Fritz, 2020).

Chronic low back pain is a complex condition that is affected by numerous factors (Allegri et al., 2016; Marras, 2012). Characterizing low back pain based on those factors into unique phenotypes is becoming important (Fairbank et al., 2011; Steinmetz, 2022). Establishing these unique phenotypes can help to formulate precise treatment plans that are associated with each phenotype. It can help health care professionals decide and select the right treatment for

patients with low back pain, which can potentially improve the treatment outcome and reduce cost.

Biomechanical (Karayannis, Jull, & Hodges, 2012; Quirk et al., 2022) and behavioral factors (Carmody, 2001; Feuerstein & Beattie, 1995; Truchon, 2001) are some factors that can be used for low back pain phenotyping. For phenotyping purposes, large biomechanical and behavioral data need to be assessed and collected. The biomechanical and behavioral data can be assessed in structured assessment protocols in clinic (Reid, Williams, & Gill, 2005; Rodrigues et al., 2017) or collected from patient-reported assessment in their daily life settings (Lin, Burke, Schlenk, & Yeh, 2019; May, Junghaenel, Ono, Stone, & Schneider, 2018). To be able to get complete data for phenotyping, assessment data from both settings should be collected.

Mobile health (mHealth) technology can be used to facilitate the assessment process (Lobelo et al., 2016; O'Reilly & Spruijt-Metz, 2013). An mHealth system can collect and record a wide range of data types, including, but not limited to, textual data, multimedia contents, and predefined survey responses. Integrating mHealth with wearable sensors can further enhance the capabilities of the mHealth in collecting and monitoring assessment data (Dobkin & Dorsch, 2011; Kumar, Jeuris, Bardram, & Dragoni, 2020; Munos et al., 2016; Thilarajah, Clark, & Williams, 2016).

The use of mHealth technology has been widespread. It has been used for accommodating assessments, delivering treatments, and helping to manage diverse health conditions (Bell et al., 2019; Dicianno et al., 2015; Parmanto et al., 2013; Pramana, Parmanto, Kendall, & Silk, 2014; Setiawan et al., 2019). However, usability remains an issue that affects the adoption of mHealth and its effectiveness (Holthe, Halvorsrud, Karterud, Hoel, & Lund, 2018; Liew, Zhang, See, & Ong, 2019; Welhausen & Bivens, n.d.). Adopting a user-centered

approach when developing the mHealth system can help optimize its usability (Chowdhary et al., 2022; Fairman et al., 2016; Wang, Zhou, Chen, Hill, & Parmanto, 2018). Usability evaluation plays an important part in the user-centered approach and in evaluating the overall usability of the mHealth system. In the user-centered approach, development is done iteratively, with usability evaluation and users' feedback incorporated in each iteration of the development process (Couture et al., 2018; Farao et al., 2020; Schnall et al., 2016).

In the case of chronic low back pain phenotyping, an mHealth system can be used to collect assessment data, including extensive and diverse biomechanical and behavioral data (Dobkin & Dorsch, 2011). Another example of data that can be collected is activity level data, a unique aspect that is part of both the causal factors and outcome metrics of low back pain. Activity level can be measured either by calculating kinematics data, which is part of the biomechanical component, or by collecting patient-reported data, which is part of the behavioral component. Comparing activity levels from objective kinematics data and perceived activity level from subjective patient-reported data would be interesting to investigate, as it would reveal if they are comparable and if one could replace or complement the other to make the mHealth system more effective and efficient.

One of the main purposes of mHealth is to deliver care or treatment to patients (Free et al., 2010a). In addition to accommodating the assessment for low back pain phenotyping, the same mHealth system can be further developed to deliver treatment. One of the ultimate goals of phenotyping low back pain is to effectively treat low back pain patients utilizing the phenotype information. Further design adjustments and development work need to be done to transform the mHealth system to accommodate treatment delivery as well.

## 1.2 Specific Aims

The main goal of this work is to develop and evaluate a comprehensive mHealth system to be used for chronic low back pain assessment for phenotyping purposes. Another goal of this work is to conduct exploratory research that can be used to further improve and transform the mHealth system into a complete assessment and treatment delivery system for chronic low back pain. To achieve those goals, the following specific aims are formulated:

1. **Specific Aim 1**: To design and develop an mHealth system for comprehensive chronic low back pain assessment.

   An mHealth system was designed and developed for chronic low back pain assessment that can collect extensive assessment data from both biomechanical and behavioral aspects and be used in both structured in-clinic assessment settings and in patients' daily life settings.

2. **Specific Aim 2**: To evaluate the usability of the mHealth system.

   The usability of the mHealth system was evaluated for two purposes: to inform iterative user-centered development of the mHealth system itself and to assess the overall usability of the system.

3. **Specific Aim 3**: To investigate and compare activity level from subjective patient-reported EMA and objective kinematics sensor data.

   The kinematics and EMA data were collected using the mHealth system. The activity level, which can be calculated from the kinematics data or collected from the EMA, was analyzed, and the correlation between the activity level from the two different sources of data was calculated and investigated.

4. **Specific Aim 4**: To design and develop dataset collection tools as preliminary work toward a personalized and adaptive intervention component in the mHealth system.

   As part of exploratory work to further improve the mHealth system for care delivery, a personalized and adaptive intervention component was designed to be integrated into the system. The numerous video recordings of the functional performance tests collected by the mHealth system contain visual information that can be used to help build a machine learning model for the personalized and adaptive intervention component. The aim is to design and develop tools for collecting datasets using the video recording data.

## 1.3 Significance

The significance of this study lies in the development of an mHealth system for the comprehensive assessment of chronic low back pain. Traditional assessment methods for chronic low back pain are often time-consuming, resource-intensive, and can be burdensome for both the physical therapist and the patients. This study develops and evaluates an mHealth system for comprehensive chronic low back pain assessment that can streamline the assessment process and reduce the burden on both physical therapists and patients.

The mHealth system developed in this study can collect extensive and rich data from both biomechanical and behavioral aspects of chronic low back pain, making it a valuable tool for chronic low back pain phenotyping. Data collection and integration with the mHealth system require minimal effort compared to traditional assessment methods, allowing the whole research process to be more efficient.

Furthermore, the mHealth system can also accommodate Ecological Momentary Assessment (EMA), providing an opportunity for real-time data collection. This can lead to more accurate assessment data by reducing recall bias.

The comprehensive and extensive data collected by the mHealth system have the potential to inform future works aimed at developing personalized and adaptive treatments for chronic low back pain. In fact, several exploratory works were conducted in this study toward developing a personalized and adaptive treatment for chronic low back pain. Furthermore, the exploratory works conducted in this study represent an important step toward improving the management of chronic low back pain. This study represents a significant contribution to the field of mHealth, providing a novel approach for the assessment and potential treatment of chronic low back pain.

## 1.4 Innovation

This study introduces several innovations in the field of mHealth and chronic low back pain assessment and treatment. First, a novel mHealth system was developed to enable comprehensive and integrated assessment of chronic low back pain. The system can collect extensive and multiform assessment data, including both biomechanical and behavioral aspects, and can be used for assessments in both structured in-clinic settings and in patients' daily life settings. Additionally, the mHealth system was designed to be flexible and configurable, allowing it to be used for other assessments and to be further developed to deliver treatment for chronic low back pain in the future.

Second, the study includes exploratory work on comparing objective and subjective/perceived activity levels. This work provides valuable insights into the differences between these two types of measures, which can inform future assessment and treatment strategies for chronic low back pain.

Finally, the study includes exploratory work on designing a personalized and adaptive treatment component that can be integrated with the mHealth system in the future. This work lays the foundation for developing tailored treatment approaches that can better address the unique needs of individual chronic low back pain patients. Overall, these innovations have the potential to help advance the field of chronic low back pain assessment and treatment, ultimately leading to better outcomes for patients.

## 1.5 Dissertation Outline

The remainder of this dissertation consists of:

- Chapter 2.0 provides a background and literature review related to the works conducted in this dissertation.
- Chapter 3.0 presents the design and development process of the mHealth system, addressing the first specific aim.
- Chapter 4.0 presents the usability evaluation of the mHealth system, addressing the second specific aim.
- Chapter 5.0 investigates the correlation between activity level from objective kinematics data and subjective patient-reported EMA, addressing the third specific aim.

- Chapter 6.0 focuses on the design and development of the dataset collection tools for a personalized and adaptive intervention component in the mHealth system, addressing the fourth specific aim.
- Chapter 7.0 provides a summary and discussion of the overall works conducted in this dissertation, reviewing each specific aim.

# 2.0 Background

## 2.1 mHealth System for Chronic Low Back Pain Assessment

### 2.1.1 mHealth

Mobile health (mHealth) refers to the use of mobile devices, such as smartphones and tablets, and other wireless technologies to support and improve health care services. Istepanian defines mHealth as mobile computing, medical sensor, and communications technologies for healthcare (R. Istepanian, Laxminarayan, & Pattichis, 2007; R. S. H. Istepanian, Jovanov, & Zhang, 2004). mHealth includes a wide range of applications, such as health monitoring, disease management, and health promotion. The primary goal of mHealth is to leverage mobile devices and mobile communication to deliver care to patients (Free et al., 2010b). mHealth has several objectives, including increasing access to care, engaging patients in treatment, improving care after treatment, and monitoring treatment progress.

Mobile devices can be utilized as a platform that patients can use anywhere at any time. Nowadays, such devices are ubiquitous, almost always available, and nearby. This ubiquity of personal mobile device usage has created the potential for mHealth to be widely adopted. In addition, the current advancements in mobile technologies have increased the capabilities of mobile devices, such that they can record a wider range of data, including audio, video, location, time, and even device kinematics data.

Utilizing the capabilities of mobile devices, mHealth can be developed as an assessment system for various health conditions. Assessment is an essential component of any management

of health conditions, and mHealth can provide an effective tool for comprehensive and integrated assessment. mHealth can be utilized in the assessment of chronic low back pain. Mobile apps and wearable devices can provide real-time data on physical activity, pain levels, and medication use, which can help health care providers assess the effectiveness of treatment and adjust it accordingly.

### 2.1.2 Chronic Low Back Pain

Chronic low back pain is one of the most prevalent health problems in all developed countries, including the US (Hoy et al., 2012; Koes et al., 2006; Murphy et al., 2017; Waterman et al., 2012). It is a complex condition that is affected by numerous factors (Allegri et al., 2016; Marras, 2012). Low back pain is defined as pain, muscle tension, or stiffness localized below the costal margin and above the inferior gluteal folds, with or without leg pain (Koes et al., 2006). Chronic low back pain is a persistent low back pain that lasts for at least three months. It is estimated that up to 80% of adults will experience low back pain at some point in their lives, and chronic low back pain affects around 20% of those individuals.

The importance of chronic low back pain lies in its significant impact on quality of life, physical function, and psychological well-being. It can limit an individual's ability to perform daily activities, work, and participate in social and recreational activities. Chronic low back pain can also lead to depression, anxiety, and social isolation.

Assessing chronic low back pain involves a comprehensive evaluation of the patient's medical history, physical examination, and diagnostic tests. The medical history should include information about the onset, duration, and nature of the pain, as well as any previous treatments or interventions. During the physical examination, the healthcare provider will assess the

patient's range of motion, strength, reflexes, and sensation in the affected area. Diagnostic tests may include imaging studies, such as X-rays, CT scans, or MRI scans, to evaluate the structure of the spine and detect any abnormalities.

There are several treatments that can be used to treat low back pain (Delitto et al., 2012; Van Middelkoop et al., 2011), including exercise therapy (Hayden, Ellis, Ogilvie, Malmivaara, & van Tulder, 2021; Van Middelkoop et al., 2010), pain medication (Peck et al., 2021), back school (Heymans, van Tulder, Esmail, Bombardier, & Koes, 2004; Parreira et al., 2017), transcutaneous electrical nerve stimulation (Khadilkar et al., 2005), low level laser therapy (Baxter, Bell, Allen, & Ravey, 1991), massage (Furlan, Brosseau, Imamura, & Irvin, 2002), behavioral treatment (Henschke et al., 2010; Ostelo et al., 2005), heat/cold therapy (French, Cameron, Walker, Reggars, & Esterman, 2006), lumbar supports (Van Duijvenbode, Jellema, Van Poppel, & Van Tulder, 2008), and multidisciplinary biopsychosocial treatment (Guzmán et al., 2001).

### 2.1.3 Integrated Assessment System

With advancements in communication protocols, mobile devices can now communicate with other components. For example, there are numerous wearable sensors that can communicate with mobile devices via Bluetooth. mHealth systems can incorporate these other components, like wearable sensors, into their framework because of the availability of data communication between the components.

Nowadays, wearable sensor technologies have advanced and are used for a variety of purposes. The development of early wearable sensors was driven by the need to objectively measure and quantify physiological functions and activities outside of laboratory settings for an

extended period of time (Ghika et al., 1993; Spieker, Jentgens, Boose, & Dichgans, 1995). The design principles of these early wearable sensors were that they be reliable, compact, portable, and simple to use.

The main objectives of wearable sensors are to detect important objective physiological parameters and to be worn by an individual (Shokri, Ward, Anton, Siffredi, & Papetti, 2020), so that they can be used for long-term and continuous monitoring in home or community settings (Bonato, 2010). These objectives are in line with the objectives of mHealth, which include increasing access to care, delivering and engaging patients in treatment, enhancing care after treatment, and monitoring treatment progress.

Integrating wearable sensors into an mHealth system can help improve and enhance the quality of the mHealth system in pursuing its objectives. Wearable sensors can be used to collect objective and real-time data from the patient. With the advancements in technology, wearable sensors can collect a wide range of data, from blood sugar and sleep to mood (Dinh-Le, Chuang, Chokshi, & Mann, 2019). Connecting and receiving data from wearable sensors can help mHealth systems obtain objective and real-time assessment information, monitor health conditions, and keep track of the wearer's compliance with intervention or treatment requirements and the effects of the intervention or treatment as well.

The integration of mHealth and wearable sensors can help transform a hospital-centered system to an individual-centered system. This transformation can reduce the need for hospital visits and decrease health costs (Teng, Poon, Zhang, & Bonato, 2008). Wearable sensors can collect a large amount of data, but mHealth has the potential to enhance this capability by providing more computational power to process the data, enabling the use of the data collected by sensors for early disease detection and timely response to health threats, as well as the ability

to provide personalized and adaptive rehabilitation interventions. mHealth can also be used to present and communicate information from the data collected for personal digital health tracking, personalized feedback, and improved care for caregivers.

The technologies used in wearable sensors have evolved from their origins in the 1990s with the evolution of electronics, sensing technologies, embedded systems, wireless communication technologies, nano technologies, and miniaturization technologies (Kyerematenga, Brousse, & Pech, 2017; Mukhopadhyay, 2015; Tricoli, Nasiri, & De, 2017). Not only their technical performance, but also their convenience of use has made great strides, which improve the adoption of use. Moreover, improvements in biocompatible materials and nano materials (Choi et al., 2018; Lim et al., 2020) have allowed advancements in implantable wearable biosensors (Song, Min, & Gao, 2019), further promoting the integration of wearable sensors with mHealth.

mHealth and wearable sensors can be integrated to support assessment, monitoring, and intervention. Jovanov designed a simple framework to integrate wearable sensors with an mHealth system to be used for assessment and monitoring (Jovanov, 2005). In the framework, the sensors communicate with the mobile device, which acts as and is defined as a personal server, via a Wireless Body Area Network (WBAN). Bluetooth and Zigbee are some examples of WBAN technologies (Georgakakis, Nikolidakis, Vergados, & Douligeris, 2011). This personal server enables the patient to monitor the processed data collected by the sensors. The personal server also acts as a gateway to the internet, transmitting the data to a medical server that a clinician or health provider can use to assess and monitor a patient's data.

This framework consists of several useful components outside of the health information transmission component (Jovanov, Milenkovic, Otto, & De Groen, 2005). For example,

information about the weather forecast can be added to the framework. It can be connected directly through internet to the personal server or through a subsequent connected device. The personal server can use the information from this component to adjust or improve its monitoring presentation for the patient. Emergency and caregiver components can also be connected to the personal server via the internet. With the addition of these components, the personal server can inform emergency personnel or the caregiver with relevant information from the patient's monitoring data as needed.

This framework works for assessment and monitoring if there is seamless connectivity and communication between each component. The patient can assess processed objective health data collected by the sensors, and the continuous and in-time benefits of using integrated mHealth-wearable sensors framework can be used to improve monitoring of the patient's health condition.

Gay et al. implemented a similar framework to develop a heart monitoring mHealth-wearable sensor system (Gay & Leijdekkers, 2007). Their framework sets a focus on emergency alarm/notification. Negative changes in a patient's heart condition cause the application in the smartphone to send out a notification to call for an ambulance or emergency services. Bisio et al. used a similar framework, adjusting the role of the mobile device to that of communication hub; their framework includes additional sensors as well as a computational processor (Bisio, Lavagetto, Marchese, & Sciarrone, 2015). Banos et al. used a similar framework to develop a personal physiological monitoring system (Banos et al., 2014).

Lobelo et al. proposed an mHealth-wearable sensors framework that further incorporates a counseling and intervention component (Lobelo et al., 2016). They implemented the framework to support physical activity assessment, counseling, and intervention for

cardiovascular disease risk reduction. In their framework, the wearable sensors and the app in the mobile device are responsible for collecting physical activity data. The data is transmitted by the app to a digital ecosystem software platform, where it is processed and standardized. The processed data is then transmitted to clinical research center entity for analysis using a clinical outcome prediction algorithm. The meaningful and summarized data is then integrated into an EMR system that is used by the healthcare team to develop a clinical decision. Finally, the clinical decision, in the form of a counseling and clinical intervention program, is sent back to the app used by the patient. Using Lobelo's framework, the patient can review the developed counseling and intervention in their app. The benefit of this framework is that it is useful not only for assessing and monitoring patient's condition but also for developing and adjusting personalized rehabilitation interventions.

## 2.2 System Design, Development, and Evaluation

### 2.2.1 Design Process

Attention to target users and purposes is crucial in guiding the design and development process of a mobile app. It helps the designers and developers to focus on creating a system that satisfies the specific needs of the target audience. By identifying the target users and their purposes, the design team can better understand what capabilities and features the system should include.

The first step in designing and developing a mobile app is to identify the requirements of the system. The requirements will form the basis of the features that will later be developed to

build the entire system. It is important to identify what the target users need, and what their perceptions are in utilizing the system. By gathering this information, the app can be designed and developed to accommodate the needs of the target users.

The design process should focus on creating a user-friendly interface that is easy to navigate and understand. The interface should be visually appealing and should reflect the purpose of the app. Additionally, the features of the app should be designed to be intuitive, efficient, and effective, and should align with the target users' goals and motivations for using the app.

The development process should also focus on ensuring that the app is responsive and compatible with a range of mobile devices, including smartphones and tablets. The app should be tested thoroughly to ensure that it is functional, reliable, and secure. This testing should be done on a range of devices and operating systems to ensure that the app performs optimally across all platforms.

### 2.2.2 User-centered Approach

User-centered design is a broad term to describe design processes in which end-users influence how a design takes shape (Abras, Maloney-Krichmar, & Preece, 2004). Understanding the target users is needed in every phase throughout the design and development life-cycle (“User-Centered Design Basics | Usability.gov,” n.d.). This means that feedback from the target users at each design and development stage is important not only at initial development stages, but also in later stages to revise and improve the design of the app. User-centered design is a critical aspect of developing an effective mHealth system.

In a user-centered design approach, the target users of the mHealth system are involved and have influence on how the system will be designed and developed. The process of user-centered design is iterative, meaning that it involves multiple cycles of feedback and revisions based on user input. This approach helps to ensure that the system is continuously refined and improved to better meet the needs of its users.

Through this iterative development process, the feedback from target users is gathered and analyzed to determine what features and capabilities need to be added, modified, or removed. The design and development team then works to incorporate this feedback into the system, creating a new version that can be evaluated by the target users. This process is repeated until the system meets the needs and expectations of the users.

The benefits of user-centered design are numerous. By involving the target users in the design and development process, the system is more likely to be intuitive and user-friendly, which can increase its adoption and use. Additionally, the system is more likely to be effective in achieving its intended purposes, as it has been tailored to the specific needs and preferences of its users. Ultimately, the goal of user-centered design is to create a system that is usable, efficient, effective, and satisfies the needs of its users.

### 2.2.3 Usability Evaluation

Jakob Nielsen defined usability as a quality attribute that assesses how easy user interfaces are to use (“Usability 101: Introduction to Usability,” n.d.). This metric can be used to evaluate the design solution in a user-centered design process. Designing and developing an app that has good usability is important. Users will not use the app if the app is not usable, meaning

that the app is hard to use, difficult to learn, not efficient in performing tasks, or even is not pleasant to use.

To be able to assess the usability of an app, metrics and measurements of usability components need to be defined. Jakob Nielsen defined usability using five quality components ("Usability 101: Introduction to Usability," n.d.):

- learnability,
- efficiency,
- memorability,
- errors, and
- satisfaction.

Learnability covers whether it is easy to use the app the first time and how easily users learn to use the app. Efficiency explains how quickly they can perform tasks in the app. Memorability involves whether users can easily use the app after a period of not using it. Errors include user perspectives related to errors made, the severity of errors made, and the user's reaction to the error itself. Satisfaction covers whether the user is satisfied with the app.

Representative users from a target user population are needed to test the usability of the app. To assess the usability of the app, the representative users would need to be asked to perform several tasks in the app. Their interaction with the app during the usability test should be observed. Letting them talk and explain their reactions and impressions during the test can enhance the effectiveness of observation as a way to evaluate usability. The five components of usability can be evaluated during the observations in a usability test.

There are several usability questionnaires that can be used to measure the usability of a system, such as System Usability Scale (SUS) (Brooke, 1996) that can be used to evaluate

usability of a general system. Using the most relevant usability questionnaires is important to get the most accurate usability score. Several usability questionnaires are made for a specific system. For example, Telehealth Usability Questionnaire (TUQ) (Parmanto, Lewis, Jr., Graham, & Bertolet, 2016) was developed to evaluate usability of a telehealth system and mHealth App Usability Questionnaire (MAUQ) (Zhou, Bao, Setiawan, Saptono, & Parmanto, 2019) was developed to evaluate usability of an mHealth app.

## 2.3 Assessment

### 2.3.1 Kinematics Sensor

Kinematics sensors, also known as inertial measurement units (IMUs), are devices that measure the acceleration and angular velocity of an object. They are typically composed of accelerometers, gyroscopes, and sometimes magnetometers. The data collected from these sensors can be used to determine the object's position, velocity, and orientation in space, as well as its motion patterns (Benson, Clermont, Bošnjak, & Ferber, 2018).

Kinematics sensors have become increasingly important in various fields, including sports, medicine, and robotics. In sports, kinematics sensors can be used to measure the movements and performance of athletes, such as tracking the velocity and acceleration of a sprinter. In medicine, kinematics sensors can be used to monitor patients' movements and activity levels, which can be useful for rehabilitation or tracking the progression of a disease. In robotics, kinematics sensors are used to help robots maintain balance and orientation, such as in humanoid robots or drones.

### 2.3.2 Ecological Momentary Assessment

Ecological momentary assessment (EMA) is a method that captures real-time data on individuals' experiences, behaviors, and environmental context in their natural setting (Shiffman, Stone, & Hufford, 2008). EMA typically involves the use of mobile devices or wearable sensors to prompt participants to report their current experiences, behaviors, or symptoms, often multiple times per day over an extended period of time.

EMA is important because it allows researchers to capture real-world data that can provide a more accurate representation of an individual's experiences compared to traditional methods of data collection, such as retrospective paper-based self-report (Stone, Shiffman, Schwartz, Broderick, & Hufford, 2002). EMA also reduces recall bias and increases the ecological validity of the data collected (Solhan, Trull, Jahng, & Wood, 2009). EMA has been used in various fields such as psychology, medicine, and public health to investigate a wide range of topics, including mood, stress, substance use, physical activity, and medication adherence.

For example, EMA has been used to study substance use and relapse (Shiffman, 2009). This study asked the participants to report their substance use information in time-based assessment. The study reported that the compliance was high when utilizing EMA. Another example is use in depression study (Armey, Schatten, Haradhvala, & Miller, 2015). The EMA was used to get insight of self-harm behavior and mood disorder in real-time. Similar study to observe alcoholics was also utilizing EMA (Litt, Cooney, & Morse, 1998).

### 2.3.3 Physical Activity Level

Physical activity level is a measure of an individual's level of physical activity. Activity level can be used to assess their risk for chronic diseases and to monitor the effectiveness of interventions aimed at increasing physical activity levels.

Accelerometer data from kinematics sensors can be used to calculate activity level using Actigraph algorithm (Neishabouri et al., 2022). The Actigraph algorithm is one of the most common methods to calculate activity level. The Actigraph algorithm for calculating activity levels involves several steps:

- activity counts processing,
- intensity level classification, and
- time proportion calculation.

First, accelerometer data is then processed into activity counts, which are accumulated over specific time intervals, such as 60 seconds. Next, the activity counts are categorized into intensity levels based on cutpoints. Cutpoints are predetermined activity count values that are associated with different levels of physical activity intensity. One of the most used cutpoints for activity level is Freedson's cutpoint (Freedson, Melanson, & Sirard, 1998).

Once the activity counts have been categorized into intensity levels, the total time spent at each level is calculated. This allows for the determination of the proportion of time spent in sedentary behavior, light physical activity, moderate physical activity, and vigorous physical activity.

## 2.4 Adaptive Intervention Framework

Each individual may have different intervention needs for their health issue. What's right for one individual might not be suitable for the another, even if the health issue is the same. Treatment dosage can also vary for one individual over time (Collins, Murphy, & Bierman, 2004). Adaptive intervention has been proposed to address this issue. In adaptive intervention, different individuals can get different dosages of treatment components (Collins et al., 2004).

There are several components in adaptive intervention: treatment, tailoring variables, measurement of tailoring variables, decision rules, and implementation of decision rules (Collins et al., 2004). Those components are interdependent. Tailoring variables are variables that moderate or are expected to moderate the effect of the treatment. Treatment dosage will be determined based on the value of the tailoring variables using the rules in the decision rules component. In time-varying adaptive intervention, tailoring variables need to be assessed periodically over time to determine whether the treatment dosage needs to be adapted or adjusted.

Just-in-time adaptive intervention (JITAI) is a specific form of adaptive intervention. Similar in concept with time-varying adaptive intervention, JITAI aims to provide the right support, at the right time, by adapting to the individual's internal and/or contextual variables changes (Nahum-Shani, Hekler, & Spruijt-Metz, 2015; Nahum-Shani et al., 2018). JITAI frameworks have been used to support behavior change interventions (Goldstein et al., 2017; Nahum-Shani et al., 2015).

Since being able to give the right support at the right time is one of the most important aims of JITAI, monitoring the individual's internal and contextual state continuously and ecologically is important as part of the intervention. The continuous monitoring aspect of JITAI

can be enhanced using the mHealth-wearable sensors system, which is designed to be able to do continuous monitoring.

There are four components of JITAI framework: intervention options, tailoring variables, decision points, and decision rules (Nahum-Shani et al., 2018). Intervention options are possible treatments or interventions that can be given at any decision point. Intervention options are also often called ecological momentary interventions (EMIs) since they are provided during everyday lives in natural settings (Heron & Smyth, 2010). Tailoring variables are types of information about the individual that can be used to decide when and what intervention to provide to the individual (Collins et al., 2004; Nahum-Shani et al., 2018). This information can be retrieved from active assessments (ecological momentary assessments or EMAs), passive assessments, or both. Decision points are the times at which an intervention decision is made. These can occur at a prespecified time interval, specific time of day, or following random prompts. Decision rules are a set of rules that specify which intervention or treatment to offer, for whom, and when. Other principal design components of JITAI are distal outcome and proximal outcomes. Distal outcome is the main objective of the intervention, while proximal outcomes are the short-term goals.

# 3.0 mHealth System for Chronic Low Back Pain Assessment: Design and Development

## 3.1 Introduction

Chronic low back pain is one of the most prevalent conditions in the US. It is a complex condition and can have many different causes. People also experience chronic low back pain differently. Some people may have a higher pain tolerance, while others may be more sensitive to pain. Because of the complex nature of chronic low back pain, it is still challenging to find effective and cost-efficient treatment for different individuals with chronic low back pain. There is a need to find a way to be able to develop precise and personalized treatment for chronic low back pain.

Phenotyping chronic low back pain is essential toward developing precision treatment for chronic low back pain. Constructing unique chronic low back pain phenotypes can help to formulate precise treatment plans that are associated with each phenotype. It can help healthcare professionals decide and select the right treatment for patients with low back pain, which can potentially improve the treatment outcome and reduce cost.

In this study, an mHealth system was developed for chronic low back pain assessment. This mHealth system focuses on collecting biomechanical and behavioral assessment data as part of extensive multifactorial data that will be used for phenotyping chronic low back pain. This chapter presents the design and development process of the mHealth system.

## 3.2 Requirement Analysis

Requirement analysis is an important part of every system development. The results of requirement analysis will inform what needs to be developed for the system to meet its purposes. Even tough revision and refinement will almost always be needed, a comprehensive requirement analysis can significantly reduce the number of changes needed in the development process, which can make the entire development process more effective and efficient.

Requirement analysis involves investigating:

- problems that the system aims to solve,
- stakeholders of the system, and
- objectives of the system.

Understanding these elements is the foundation and the first step toward the design and development of the system. To conduct the requirements analysis for development of the system discussed in this study, three use cases were used:

- structured assessment in clinic setting (in-clinic assessment),
- assessment in patients' home and daily life setting (at-home assessment), and
- data monitoring and access for the researchers.

### 3.2.1 In-clinic Assessment

In the in-clinic assessment use case, the physical therapists assess people with chronic low back pain through a specified set of clinical exams, functional performance tests, and

quantitative sensory testing (QST) in clinic. They need to record the results of the exam and test assessments. Kinematics data and video recording of the patients performing some of the tests also need to be recorded and collected. The problems that need to be solved from this use case were identified:

- extensive and different types of data, such as exams and tests completion information, time needed to complete some tests, video and kinematics data of some tests, physical therapists' notes and comments, assessment deviation information, and case report responses, needs to be collected,
- collected data needs to be further inputted to an integrated database,
- since not all participants are able to do all the exams and tests safely, safety screening, that can be extensive and complex, is needed,
- various tools, such as pen, paper, laptop/computer, stopwatch, and camera, are needed to collect and integrate the data, and
- physical therapists need to maintain and juggle between various resources, such as exam guide, safety screening rules, assessment checklist, and case report form, that can be burdensome and can make the assessment session takes longer than needed.

For the stakeholders, three different individual roles were identified:

- physical therapist, who is responsible to screen the patient for safety, direct the patient to perform the exams and tests in the assessment, and input the outcome of the assessment,

- patient, who is responsible to perform and complete the exams and tests in the assessment, and
- study coordinator, who is responsible for organizing the patients' information and assign study subject ID to the patient.

Based on the identified problems and stakeholders' roles, objectives of the system for this use case were explored. The main objective of the system for this use case was to help and accommodate the physical therapists to conduct the assessment process efficiently with minimal burden for both physical therapists and patients. Some detailed objectives were identified:

- the system should be able to collect and record extensive and different type of data,
- the system should automatically integrate all data into the study database without having the physical therapists to reinput the data into different system,
- the system should provide safety screening tools and automatically generate list of exams and tests that are safe for a patient to perform,
- the system should be able to reduce the tools needed, such as pen, paper, laptop/computer, stopwatch, and camera, and provide the physical therapists with the same capabilities,
- the system should provide the physical therapists the needed resources, such as exam guide, safety screening rules, assessment checklist, and case report form, in one component to streamline the assessment process and to reduce the burden on the physical therapists, and

- the system should provide the physical therapists with a portable component that provide the capabilities as mentioned in preceding points and provide the study coordinator with a component to organize and assign patients' information.

### 3.2.2 At-home Assessment

In at-home assessment use case, the patients report self-assessed pain, behavioral, and activity information three times a day during a 7-day assessment period in their own home or daily life settings. Kinematics data of the patients during the 7-day period also needs to be recorded and collected. The problems that need to be solved from this use case were identified:

- the patients need a way to record their self-reported assessment during their daily life activities,
- the patients might forget to report their assessment,
- the patients and the study coordinators need a way to communicate to each other securely if the patients have any questions or concerns during the 7-day assessment period, and
- all patients' self-reported assessment data and the kinematics data need to be collected and inputted to the study main database.

For the stakeholders, two different individual roles were identified:

- patient, who is responsible to record their self-assessment reports during the 7-day assessment period, and

- study coordinator, who is responsible to prepare and brief the patient for the 7-day at-home assessment and to respond to or help the patients during the 7-day assessment period if needed.

Based on the identified problems and stakeholders' roles, objectives of the system for this use case were explored. The main objective of the system for this use case was to accommodate the patients to record their self-assessment reports during the 7-day self-assessment period in their home or daily life settings. Since the setting is in patients' home and daily life places, the system should be able to be integrated seamlessly into patients' life without adding any unnecessary burden to their life. Some detailed objectives were identified:

- the system should be able to provide the patients with a simple and easy-to-fill assessment form that they can use to record their assessment,
- the system should provide reminder and notification system to remind the patients to record their self-assessment reports,
- the system should provide the patients and the study coordinator with a secured two-way communication component and provide the patients with informational materials regarding the at-home assessment,
- the system should automatically integrate all data into the study database without having the study coordinator to reinput the collected data into different system, and
- the system should provide the patients with a portable component that provide the capabilities as mentioned in preceding points and that can be used in their daily life settings.

### 3.2.3 Data Monitoring and Access

In the data monitoring and access use case, researchers need to monitor and access the assessment data collected from both in-clinic and at-home assessment. The problems that need to be solved from this use case were identified:

- extensive assessment data needs to be integrated and aggregated for review and monitoring purposes,
- the researchers need a way to monitor and access all assessment data that have been collected from both in-clinic and at-home assessment anytime and anywhere, and
- the data should only be accessed by authorized researchers.

For the stakeholders, one individual role was identified:

- researcher, who has access and can monitor the assessment data.

Based on the identified problems and stakeholder' role, objectives of the system for this use case were explored. The main objective of the system for this use case was to accommodate the researchers to access and monitor the assessment data. Some detailed objectives were identified:

- the system should have component that can manage extensive assessment data,
- the system should provide a component that can be used by the researcher to access and monitor the assessment data anytime and anywhere, and

- the access to the component should be secure and should require authentication protocol.

The summary of requirement analysis for the three use cases is presented in Table 1.

**Table 1. Summary of Requirement Analysis**

| | **Problems** | **Stakeholders** | **Objectives** |
|---|---|---|---|
| In-Clinic Assessment | • Extensive data.<br>• Scattered data.<br>• Screening process.<br>• Multi-tools requirement.<br>• Multiform assessment resources. | • Physical therapist.<br>• Patient.<br>• Study coordinator. | • Extensive data collection.<br>• Data management and integration.<br>• Screening automation.<br>• Multi-tools capabilities.<br>• Integrated assessment resources and streamlined workflow.<br>• Assessment component for physical therapist and patient management component for study coordinator. |
| At-Home Assessment | • Assessment reporting.<br>• Reporting | • Patient.<br>• Study coordinator. | • Assessment reporting component.<br>• Reminder system. |

| | | | |
|---|---|---|---|
| | adherence.<br>• Communication.<br>• Scattered data. | | • Two-way secured communication.<br>• Data management and integration.<br>• Portable assessment component for patient. |
| Data Monitoring and Access | • Extensive and scattered data.<br>• Data access.<br>• Data security. | • Researcher. | • Data management and integration.<br>• Component for data access and monitoring.<br>• Access authentication and authorization. |

### 3.3 System Design

Based on the requirement analysis, several components were identified for the mHealth system to incorporate. The identified components are listed in Table 2.

**Table 2. Identified mHealth Components from Requirement Analysis**

| | Identified Components | Target User |
|---|---|---|
| In-Clinic Assessment | 1. In-clinic assessment component.<br>2. Patient management component.<br>3. Data management component.[1] | 1. Physical therapist.<br>2. Study coordinator. |
| At-Home Assessment | 1. Portable at-home assessment reporting component.<br>2. Secured communication component.<br>3. Data management component.[1] | 1. Patient.<br>2. Patient and study coordinator. |
| Data Monitoring and Access | 1. Data access and monitoring component.<br>2. Data management component.[1] | 1. Researcher. |

The components were further organized and grouped based on the functionalities and the target users. Four final main components were identified: in-clinic app, at-home app, clinician portal, and backend and database. The four components cover all identified components and objectives from requirement analysis. This grouping was made to streamline the mHealth system and to create a straightforward and efficient integration between all components.

The detail of the four final components and the requirement analysis items they cover is presented in Table 3.

[1] Data management component automatically handles the management and integration of collected/reported assessment data. It does not have any target user who use this component directly.

**Table 3. Final Components for the mHealth System**

| Final Components | Components from Requirement Analysis | Target User | Objectives |
|---|---|---|---|
| In-clinic app | • In-clinic assessment component. | • Physical therapist | • Extensive data collection.<br>• Screening automation.<br>• Multi-tools capabilities.<br>• Integrated assessment resources and streamlined workflow.<br>• Assessment component for physical therapist. |
| At-home app | • Portable at-home assessment reporting component.<br>• Secured communication component. | • Patient | • Assessment reporting component.<br>• Reminder system.<br>• Two-way secured communication.<br>• Data management and integration.<br>• Portable assessment component for patient. |
| Clinician portal | • Patient management | • Study coordinator | • Patient management component for study |

| | | | |
|---|---|---|---|
| | component.<br>• Secured communication component.<br>• Data access and monitoring component. | • Researcher | coordinator<br>• Two-way secured communication.<br>• Component for data access and monitoring.<br>• Access authentication and authorization. |
| Backend and database | • Data management component.[1] | - | • Data management and integration. |

### 3.3.1 In-Clinic App

The in-clinic app was designed to be used by physical therapists to help them run the assessment and collect the extensive and multiform assessment data. To be able to collect kinematics data, the in-clinic app needs to incorporate kinematics sensors. Using Bluetooth technology, communication and data transfer between the app and the sensors can be established.

The in-clinic app was also designed to make it easier for the physical therapist to screen patients for safety and to automatically generate list of exams and tests that are safe for the patient to perform. The app was also designed to reduce the tools needed for the physical therapist to run the assessment, such as pen, paper, laptop/computer, stopwatch, and camera. The app was designed to have the capabilities that can substitute the use of the mentioned tools. The

app was also designed to incorporate several resources that the physical therapist needs, such as exam guide, safety screening rules, assessment checklist, and case report form. Having these resources in one place can streamline the overall assessment process and reduce the burden on the physical therapists.

### 3.3.2 At-home App

The at-home app was designed to be used by the patients to accommodate them to report their assessment in their daily life settings. One of the assessment data collected in the at-home session is kinematics data. To be able to collect kinematics data, the at-home app needs to incorporate kinematics sensors. Using Bluetooth technology, communication and data transfer between the app and the sensors can be established.

The at-home app was also designed to have a reminder system to improve the adherence of the patient in reporting the assessment. Other than that, the app was also designed to have secured communication to enable the patient to communicate with the study coordinator if they have questions or concerns about the assessment process.

### 3.3.3 Clinician Portal

The clinician portal was designed to be used by both the study coordinator and researcher. To enable multiple user roles, the clinician portal was designed to have authorization access capabilities. For the study coordinator, the clinician portal was designed to enable them to input and manage patient data and to provide them with communication module to communicate

with the patient during their at-home assessment period. For the researcher, the clinician portal was designed to be able to gather assessment data and to present them to the researcher.

### 3.3.4 Backend and Database

The database was designed to store the assessment data securely. The backend component was designed to manage all assessment data collected from in-clinic app and at-home app, to manage participant information input in clinician portal, and to integrate and store those data in the database. Ultimately, the backend was designed to be able to communicate with all other components in the mHealth system: in-clinic app, at-home app, clinician portal, and the database itself.

The integration of all components and overall architecture design of the mHealth system is shown in Figure 1.

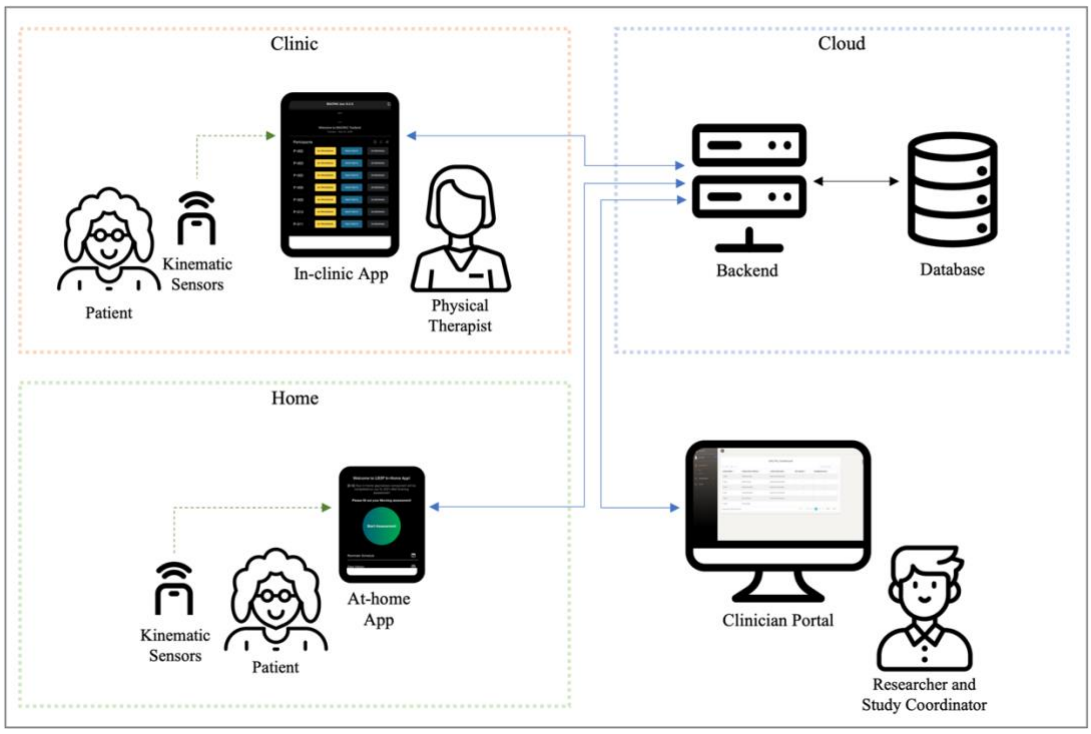


**Figure 1. Architecture design of the mHealth system for chronic low back pain.**

## 3.4 System Development

### 3.4.1 Technology Requirement

Before developing the mHealth system, technology requirement analysis was done to make sure the development is feasible and to purposes of developing the system is achievable. Technology requirement analysis was performed for each component of the mHealth system: in-clinic app, at-home app, clinician portal, and backend and database.

#### 3.4.1.1 In-Clinic App

The in-clinic mobile app was planned to be developed using a cross-platform app development approach to enable installation in different platforms. Using a cross-platform development approach will also make it easier to manage the code base with limited resources. The app project was coded using Ionic Framework (https://ionicframework.com/docs) and Angular Framework (https://angular.io/). The code was written in TypeScript, HTML, and CSS – technologies that are commonly used for web development. Packaging and building of the cross-platform app were done using Capacitor (https://capacitor.ionicframework.com). The development environment was set as listed in Table 4.

**Table 4. Technology Components for In-Clinic App Development**

| Technology Component | Version |
|---|---|
| Ionic | 5.0.0 |
| Angular | 8.2.14 |
| Capacitor | 1.5.0 |
| NodeJS | 13.3.0 |
| npm | 6.13.1 |

The in-clinic app was designed to incorporate the integration with kinematics sensors. For the sensors, kinematics wearable sensors from Lifeware (Lifeware Labs, LLC, Pittsburgh, PA; https://www.lifewarelabs.com/) was planned to be used. The Lifeware wearable sensors can record kinematics data such as accelerometer, magnetometer, gyroscope, and quaternion data. The sensors can communicate and transfer the data through Bluetooth connection.

To integrate the app with the kinematics sensors, both components need to establish communication protocol between the two of them. Unfortunately, at the time of the development, Lifeware could not provide APIs information that was needed for the app to be able to communicate with the sensors through Bluetooth. The in-clinic assessment ended up using different standalone app from Lifeware that collect and manage the kinematics data during the assessment process.

#### 3.4.1.2 At-Home App

Similarly, like the in-clinic app, the at-home mobile app was also planned to be developed using a cross-platform app development approach to enable installation in different platforms. It is important for the app to be able to be used in a large range of mobile device platforms because the app was planned to be installed in the patient's own mobile device. The app project was coded using Ionic Framework and Angular Framework. The code was written in TypeScript, HTML, and CSS. Packaging and building of the cross-platform app were done using Capacitor. The development environment was set as listed in Table 5.

**Table 5. Technology Components for At-Home App Development**

| Technology Component | Version |
|---|---|
| Ionic | 5.0.0 |
| Angular | 8.2.14 |
| Capacitor | 1.5.0 |
| NodeJS | 13.3.0 |
| npm | 6.13.1 |

The in-clinic app was designed to incorporate the integration with kinematics sensors. Similarly, like the in-clinic app, wearable sensors form Lifeware were planned to be used. The Lifeware wearable sensors that were planned to be used for at-home assessment have some differences from the ones for in-clinic assessment. For the at-home assessment, the sensors must be worn during a 7-day period in patient's daily life settings. Because of that, the sensors must be able to collect data for a long period of time and must optimize its battery life. Bluetooth communication is sacrificed to make the battery last longer so the sensors would not need to have bigger battery that will inconvenience the wearer. Since there is no means to communicate with the sensors, the at-home app was not developed to collect kinematics data from the sensors. Kinematics data from the sensors for the at-home session was collected manually by connecting the sensors to a computer via direct universal serial bus (USB) connection.

#### 3.4.1.3 Clinician Portal

The clinician portal was planned to be developed for use in common web browsers, such as Google Chrome, Microsoft Edge, Apple Safari, and Mozilla Firefox. The web portal was coded using Angular Framework. The code was written in TypeScript, HTML, and CSS. The development environment was set as listed in Table 6.

**Table 6. Technology Components for Clinician Portal Development**

| Technology Component | Version |
|---|---|
| Angular | 8.2.0 |
| NodeJS | 13.3.0 |

| | |
|---|---|
| npm | 6.13.1 |

#### 3.4.1.4 Backend and Database

The database was planned to be built on a MySQL database system. The backend system was built using Spring Framework (https://spring.io/). The data source connection to the database used MySQL JDBC driver. The code was written in Java and incorporated MySQL queries. The development environment was set as listed in Table 7.

**Table 7. Technology Components for Backend and Database Development**

| Technology Component | Version |
|---|---|
| MySQL | 5.7.32 |
| Java | 1.8 |
| Spring | 2.0.1 |

### 3.4.2 In-Clinic App Development

Using the information on the requirement analysis and the constructed design of the mHealth system, the initial version of the in-clinic app was developed with several core features and services:

- login module and authentication service,
- home/main menu page,
- safety screening module and processing service,

- exam/test module,
- recording module and service,
- case report form/questionnaire module and processing service,
- local data storage service, and
- remote communication service.

The login module and authentication service were developed to make sure that only approved physical therapist can access the in-clinic app. Home/main menu page was developed to list the patients that the study coordinator has registered in the clinician portal. The physical therapist can select the patient from the list to perform safety screening process and to run the assessment process. The safety screening module and processing service was developed to provide the physical therapist with safety screening questionnaires. The safety screening processing service was designed to automatically generate a list of exams/tests that are safe for the patient to perform.

The exam/test module was developed to contain the exam/test information and guide. The physician therapist can use this module to direct the patient to perform the exam/test assessment. The recording module and service were designed to record the time and video of the patient's exam/test performance. The case report form/questionnaire module and processing service were designed to provide the physical therapist with specific case report form for each exam/test. The case report form was designed to be easy to fill and to help validate the data type for each answer to the question item. The exam/test module, recording module, and case report form/questionnaire module were designed to be integrated with each other to create a seamless process for the physical therapy when they run the in-clinic assessment process.

Local data storage service was designed to securely store assessment data in local device temporarily before sending it to the remote main study database. Remote communication service was designed to communicate with the backend component to retrieve and send data from and to the remote main database. The physical therapist wouldn't have to reinput the assessment data collected using the in-clinic assessment in a different system. Data integration was designed to be handled by the in-clinic app. All of these modules and services were designed to help the physical therapist run the whole assessment process effectively and efficiently. Using the in-clinic app should minimize the burden and effort on the physical therapist in the in-clinic assessment process.

The screenshots of the modules developed in the in-clinic app are shown in Figure 2.

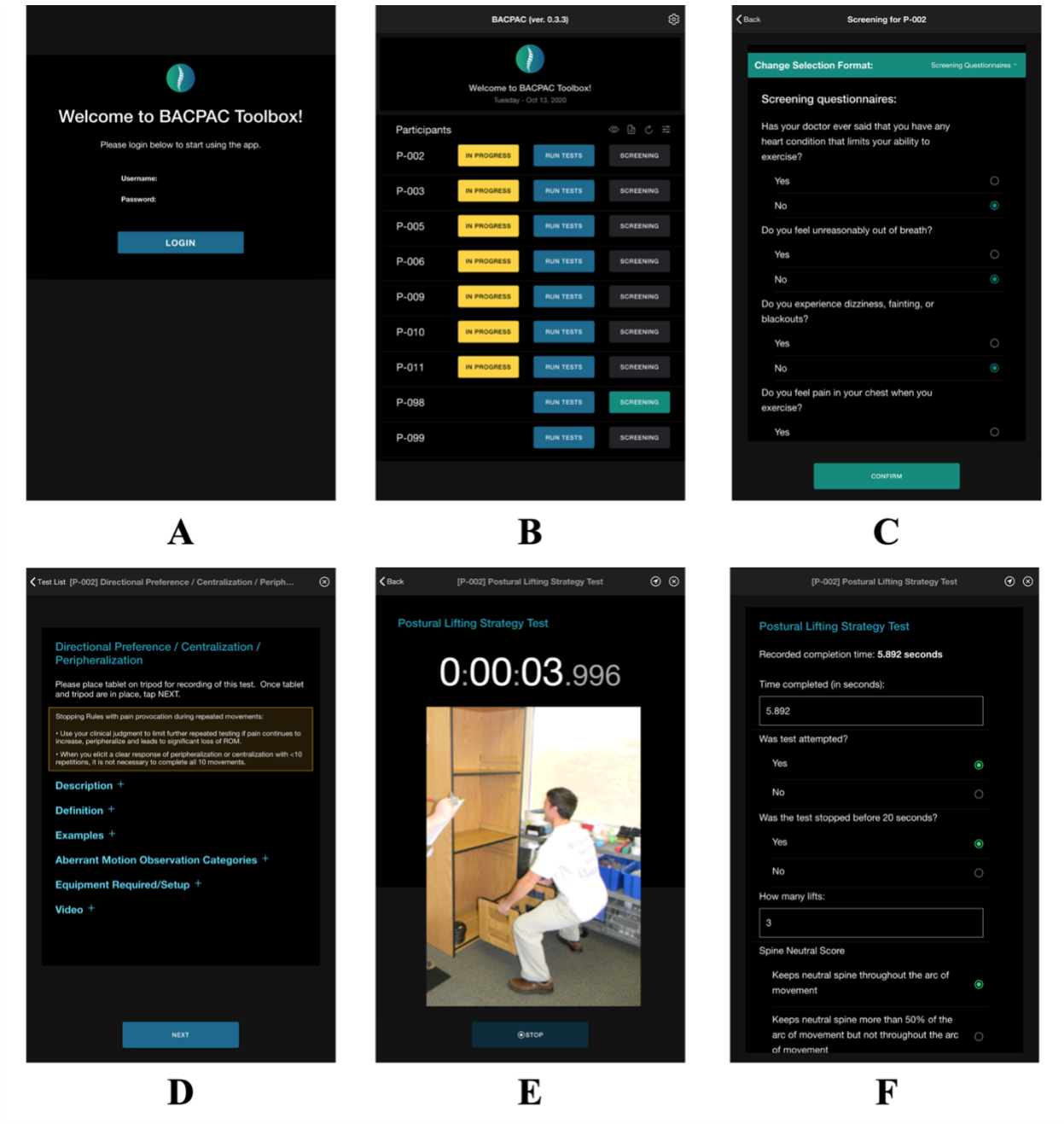


**Figure 2. Screenshots of the modules in the first version of the in-clinic app.**

**Figure A shows the login module. Figure B shows the main/home page. Figure C shows the safety screening module. Figure D shows the exam/test module. Figure E shows the recording module. Figure E shows the case report form module.**

The integration of the modules was developed to follow the assessment process workflow. The whole modules were developed to connect with each other seamlessly. The app was developed to help the physical therapist from the start to the end of the in-clinic assessment process. At first, the physical therapist would need to login in to the in-clinic app using their own credential. After successful login, the app will show the main/home page that will provide the physical therapist with the list of patients. The physical therapist selects the patient from the list and performs safety screening process. After the list of exams/tests for the patient is generated, the physical therapist can start directing and recording the patient through several sets of exams/tests. The exams/tests were listed based on order from top to bottom on the main test list page to make it easier for the physical therapist to run those tests. After each exam/test, the physical therapist inputs the assessment data in the case report form that the app shows after the test is recorded.

The flow and connection between each module are presented in the following figures. Figure 3 shows the flow when the physical therapist logs into the app. Figure 4 shows the flow when the physical therapist screens the patient for safety. Figure 5 shows the flow when the physical therapist directs the patient to perform an exam/test. Figure 6 shows the flow when the physical therapist records the time and video of the patient performing the test and fills out the case report form for the test performed by the patient. Finally, Figure 7 shows the overall in-clinic app flow.

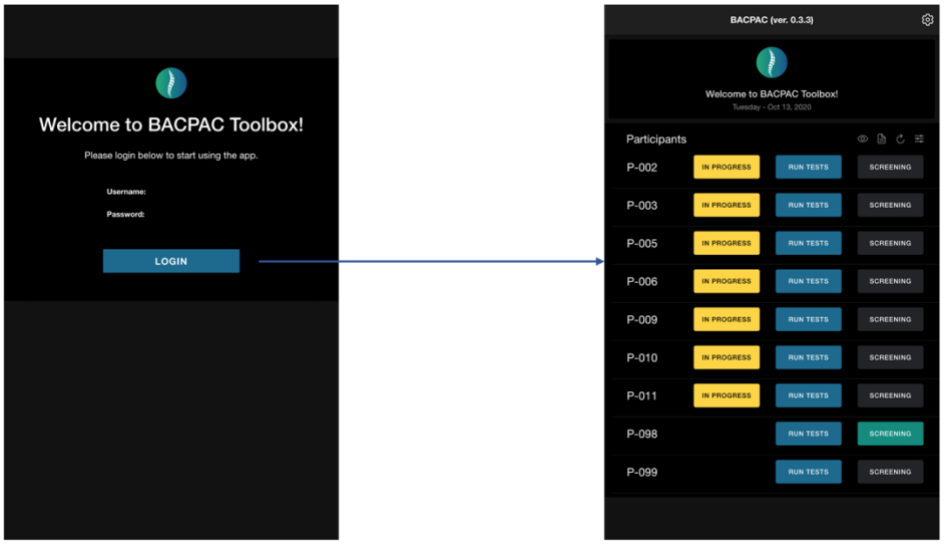


**Figure 3. App flow on logging into the in-clinic app.**

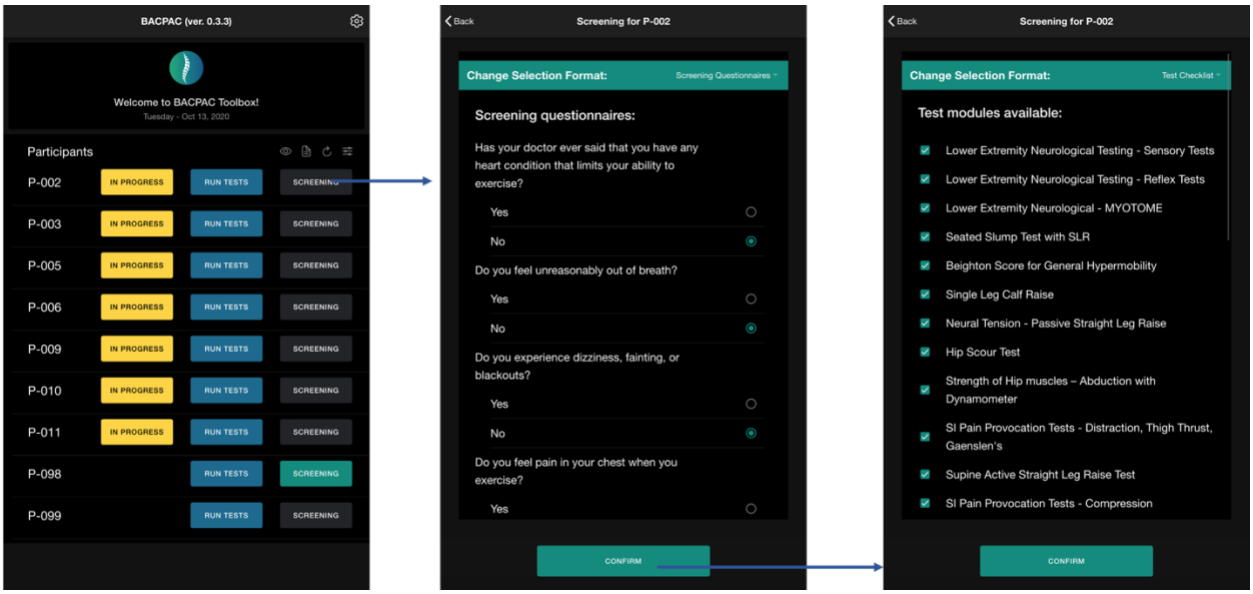


**Figure 4. App flow on screening the patient for safety.**

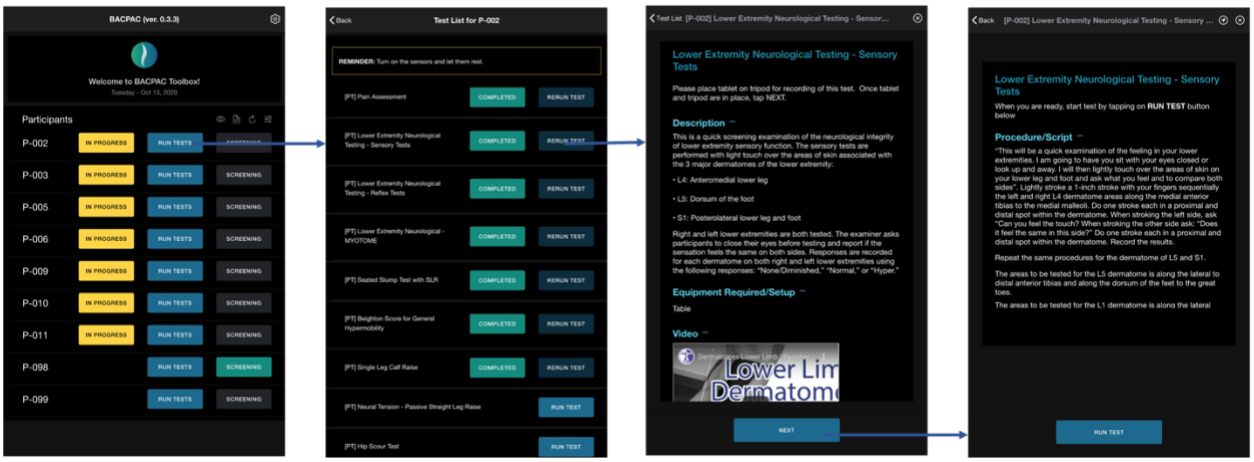


**Figure 5. App flow on running an exam/test.**

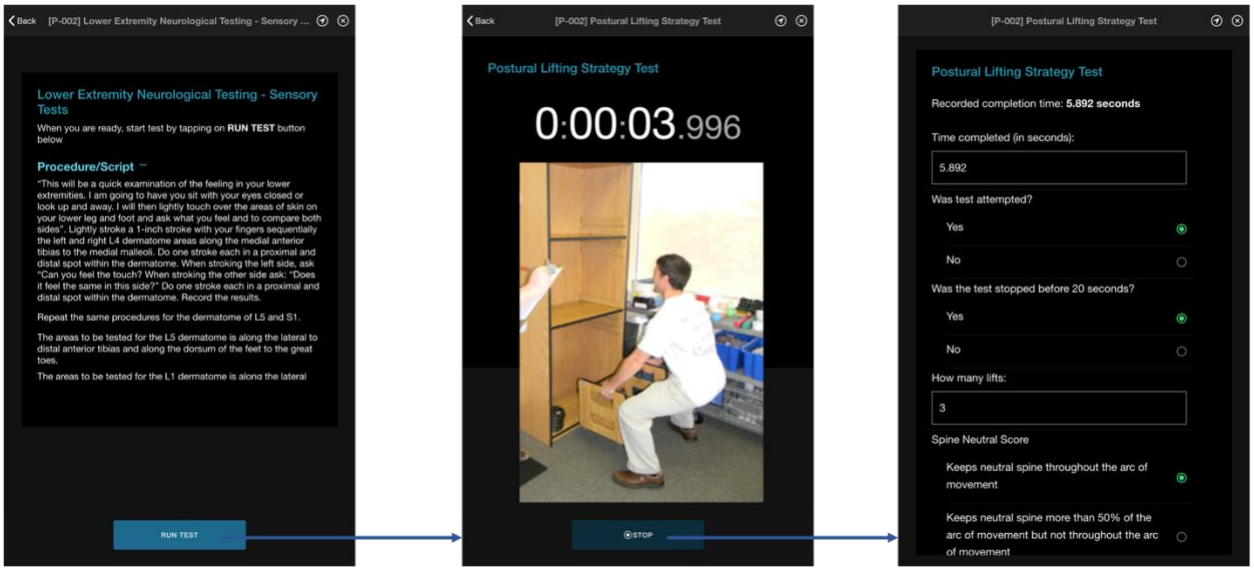


**Figure 6. App flow on recording the test and reporting the assessment data.**

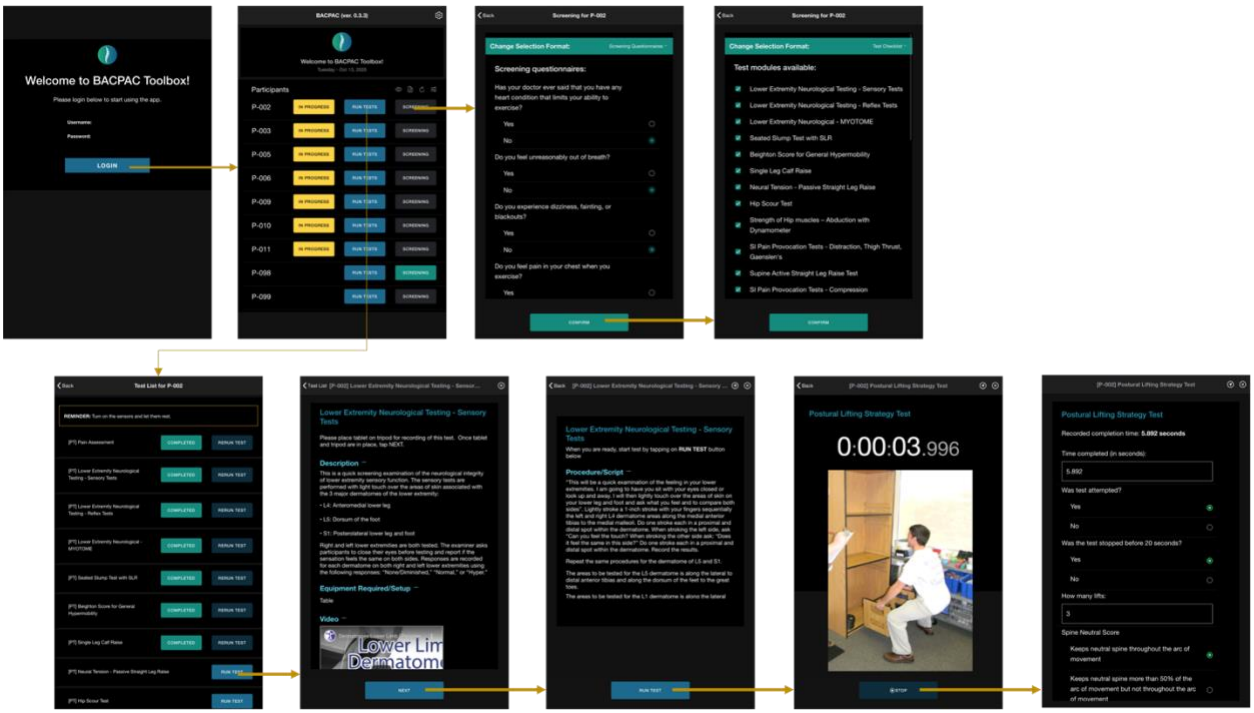


**Figure 7. Overall app flow of the in-clinic app.**

### 3.4.3 At-Home App Development

Using the information on the requirement analysis and the constructed design of the mHealth system, the initial version of the at-home app was developed with several core features and services:

- login module and authentication service,
- home/main menu page,
- Ecological Momentary Assessment (EMA) module,
- reminder module and service,
- messaging module,
- notes, account, and settings modules,

- local data storage service, and
- remote communication service.

The login module and authentication service were developed to ensure that only the patient can access their own data in the app and to ensure the assessment data collected in the at-home app is linked to the correct patient. Home/main menu page was developed as the main landing page in the at-home app. Here, the patient can access other modules through the menu provided on this page. Ecological Momentary Assessment (EMA) module was developed to provide the patient with the assessment form that the patient can fill out with their assessment data. EMA was used because the assessments in the at-home session need to be assessed and reported over time in their daily life settings.

Reminder module and service were developed to help remind the patient to report their assessment data at specific times of the day. The patient can set their preferred time as long as the time is within the time range to report the assessment. Messaging module was developed to provide the patient with a means to communicate with the study coordinator if they have any questions or concerns regarding the at-home assessment. Notes, account, and settings modules were not part of the requirement analysis result, but, during the development process, the need to enable the patient to put in notes related to the assessment, to review their account information, and to be able to logout from the app was identified.

Local data storage service was designed to securely store assessment data locally. Remote communication service was designed to communicate with the backend component to retrieve and send data from and to the remote main database. The patient nor the study coordinator wouldn't have to reinput the assessment data collected using the at-home assessment

in a different system. Using the at-home app should enable the patient to report their assessment over time in their daily life settings.

The screenshots of the modules developed in the at-home app are shown in Figure 8.

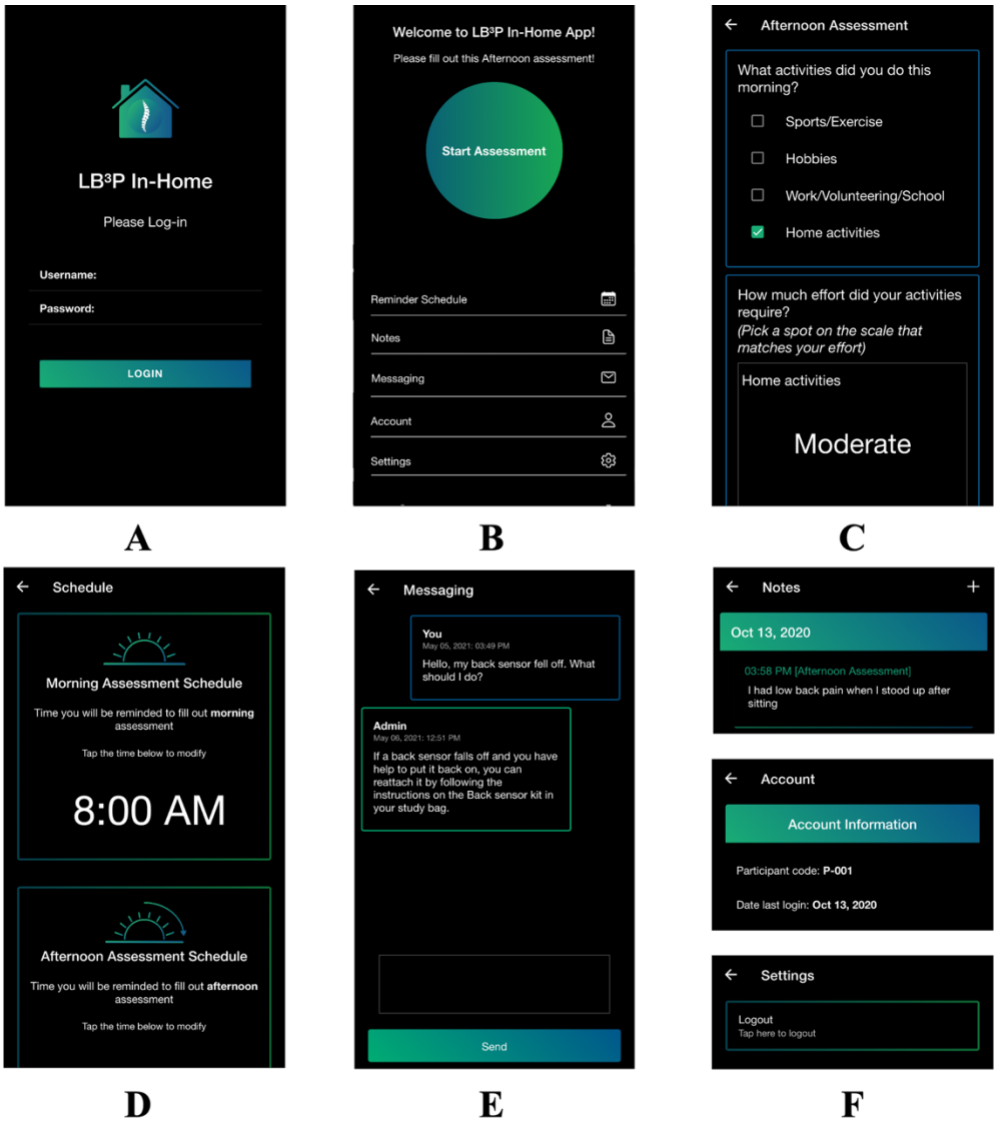


**Figure 8. Screenshots of the modules in at-home app.**

**Figure A shows the login module. Figure B shows the main/home page. Figure C shows the EMA module. Figure D shows the reminder module. Figure E shows the messaging module. Figure E shows the account and setting modules.**

Since the assessments in the at-home session were designed to be reported over time, the app was developed to provide the patient easy access to the EMA module when they open the app. Access to other modules were also made to be accessible by putting all access on the main/home page. From the home page, the patient can access the EMA module, reminder module, notes module, messaging module, account module, and settings module. The EMA module was highlighted to give more importance. The app flow on filling out EMA module is illustrated in Figure 9. The flow for accessing the other modules is quite straightforward. Overall, the app flow is illustrated in Figure 10.

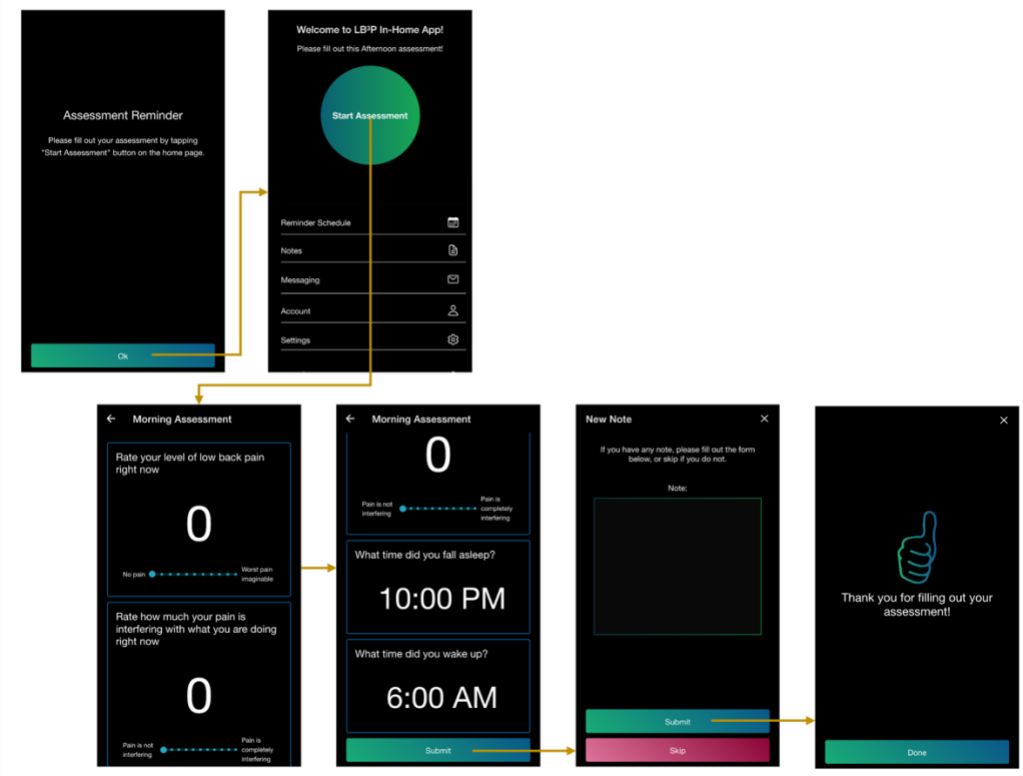


**Figure 9. App flow on reporting morning EMA.**

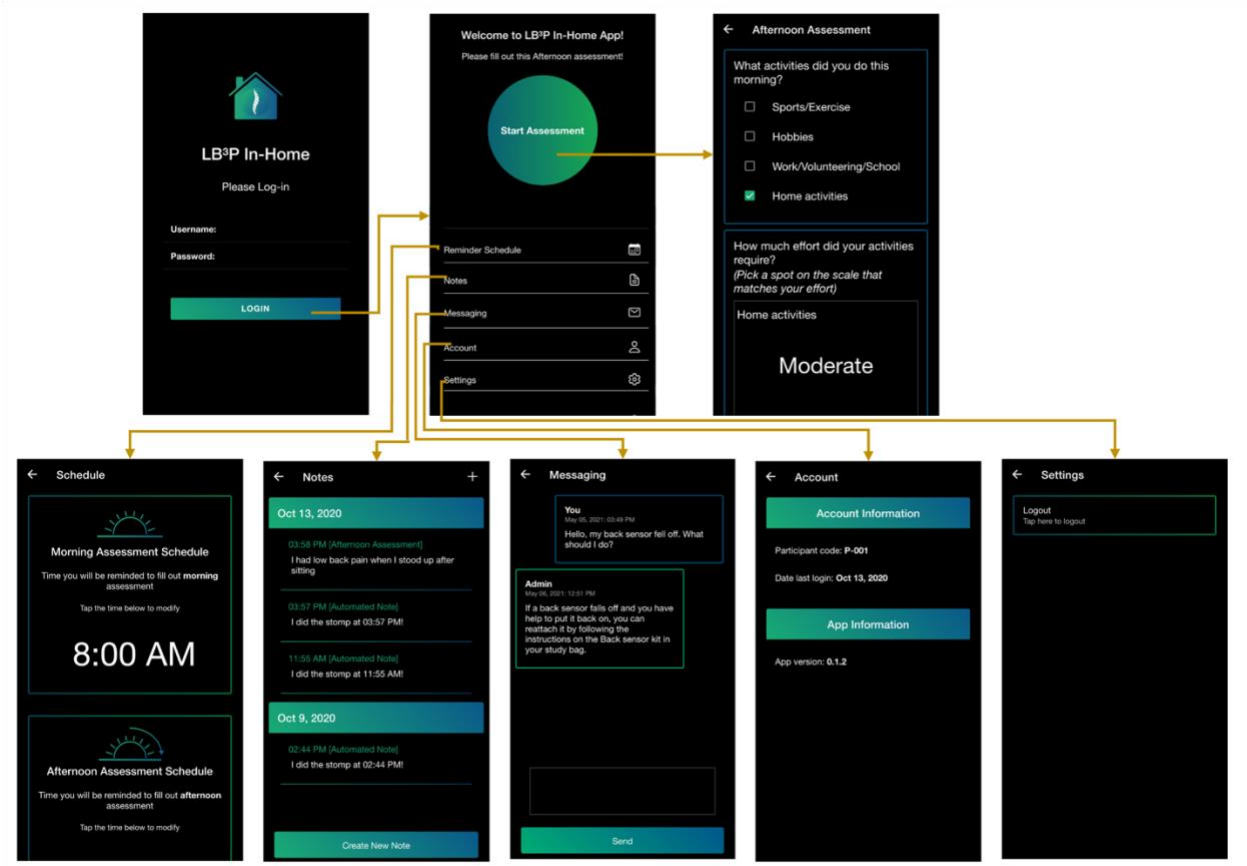


**Figure 10. Overall app flow of the at-home app.**

### 3.4.4 Clinician Portal Development

Using the information on the requirement analysis and the constructed design of the mHealth system, the clinician portal was developed with several core features and services:

- login page and authentication service,
- home/main menu page,
- assessment data page,
- patient account management module,

- messaging module, and
- remote communication service.

The login module and authentication service were developed to ensure only approved study coordinator and researcher can access the clinician portal. The home/main menu page was developed to provide a landing page once the user login to the portal and to provide menu to access other modules. Assessment data page was developed to present the assessment data collected from both in-clinic and at-home assessment. Patient account management module was developed to enable the study coordinator to register a new account or update account information of the patient. Registering an account for a patient is an important step because the in-clinic app gets the patient account information from here. Messaging module was designed to provide a means for study coordinator to communicate with the patient during their at-home assessment period. Remote communication service was designed to communicate with the backend component to retrieve and send data from and to the remote main database.

The screenshots of the modules developed in the at-home app are shown in Figure 11.

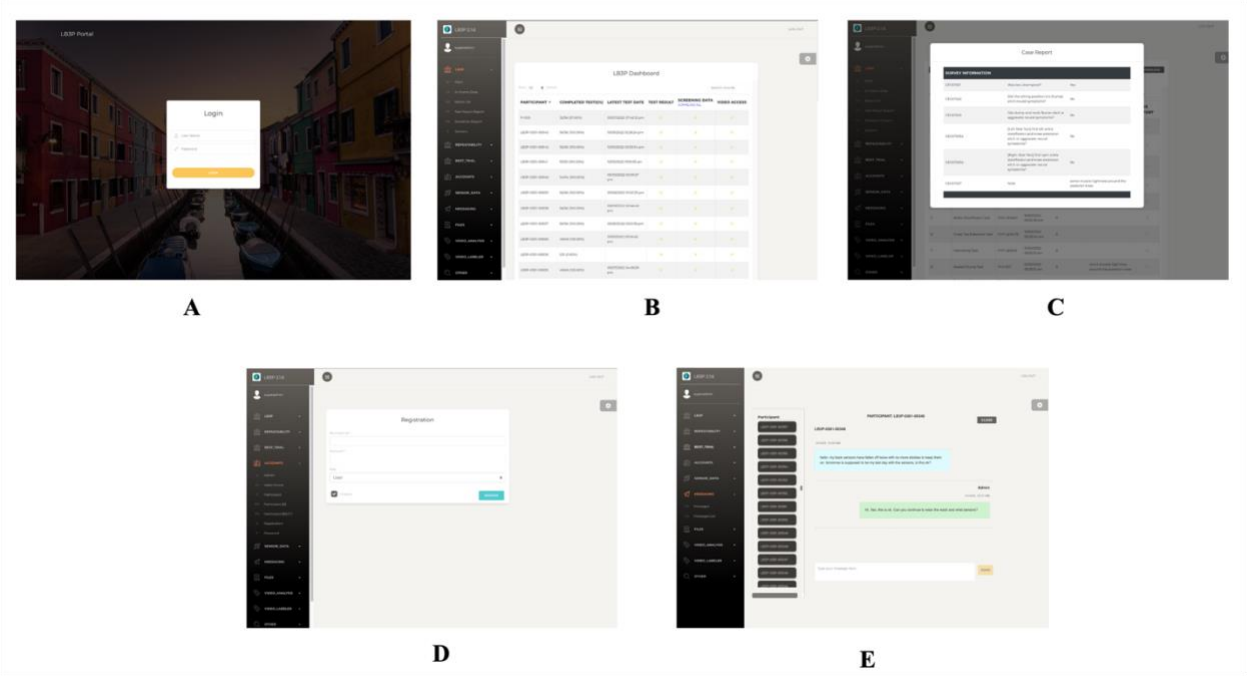


**Figure 11. Screenshots of the modules in the clinician portal.**

**Figure A shows the login module. Figure B shows the home/main menu page. Figure C shows the assessment data module. Figure D shows the patient management module. Figure E shows the messaging module.**

### 3.4.5 Backend and Database Development

Using the information on the requirement analysis and the constructed design of the mHealth system, the backend and database were developed. The main feature of the database is to provide secure storage for patients' information and assessment data from both in-clinic and at-home assessment sessions. The backend was developed with several core features and services:

- access to the database,

- data processing services,
- APIs for in-clinic app,
- APIs for at-home app, and
- APIs for clinician portal.

The backend component was developed to manage all assessment data collected from in-clinic app and at-home app, to manage participant information input in clinician portal, and to integrate, process, and store those data in the database. The backend has direct access to the database and was developed to provide APIs for the in-clinic app, at-home app, and clinician portal to enable them to send and retrieve necessary data.

The APIs developed in the backend were tailored to the needs of the other components. For example, APIs to send and retrieve safety screening result data, test result data, and videos were developed to be used by the in-clinic app. For the at-home app, several APIs such as APIs to send and retrieve patients' EMA reports, notes, and messages were developed. Extensive data communication APIs were also developed to be used by the clinician portal. The backend plays an important role in providing communication for the other components in the system.

## 3.5 Discussion

This chapter presents the discussion of the development process of the mHealth system for chronic low back pain assessment. Four main components were identified and developed: in-clinic app, at-home app, clinician portal, and backend and database. The in-clinic app that has been developed addresses the limitations of in-clinic paper-based assessments and

questionnaires. The system provides the necessary assessment resources in one place. Additionally, the use of the in-clinic app streamlines the assessment process and reduces the tools needed for the assessment, such as pens, papers, cameras, and stopwatches.

The at-home app was developed to support ecological momentary assessment (EMA) by providing real-time reporting of pain information and physical activities. This feature enhances the accuracy and reliability of the data while reducing the potential for recall bias. The at-home app also supports compliance in reporting using the reminder system, which ensures that patients report their pain information and physical activities regularly and accurately.

Overall, the mHealth system streamlines the assessment process and make the assessment more effective and efficient. The data collected during the assessment using the in-clinic app and the at-home app can be automatically integrated by the system without requiring additional manual data input. Using this mHealth system saves time and effort for the physical therapists and the chronic low back pain patients.

The result of the development process in this chapter was the initial version of the mHealth system. This version was used to perform assessments for the first patients in this study. However, the development process of the mHealth system did not stop at the first version. The system underwent several iterations of revisions and improvements based on the usability evaluation and feedback from the stakeholders, such as physical therapists and chronic low back pain patients. These revisions and improvements were essential in ensuring that the system is user-friendly, effective, and aligned with the needs of the users of the system.

The usability evaluation of the mHealth system was a critical aspect of its development. The evaluation aimed to identify the usability issues of the system and to obtain feedback from users. Because of that, the iterations of the development after the first version of the mHealth

system will be discussed in Chapter 4.0, along with its usability evaluation. In-clinic app and at-home app will be the focus in Chapter 4.0 when discussing the further development and usability evaluation of the mHealth system.

# 4.0 mHealth System for Chronic Low Back Pain Assessment: Iterative Development and Usability Evaluation

## 4.1 Introduction

Chapter 3.0 discusses the development of the initial version of the mHealth system. The development process did not stop at that point and continued iteratively. Since the development adopted a user-centered approach, evaluation and feedback from the target users were needed to inform the further iteration of the system's development. This chapter discusses the usability evaluation of the in-clinic app and at-home app of the mHealth system and the iterative development of those components. The development discussed in this chapter is the continuation of the development discussed in Chapter 3.0.

Usability evaluation is important to inform if the system is usable and can deliver its intended purposes to the target users. The evaluation is also useful to identify changes and improvements needed in development purposes. Evaluating usability is in line with the user-centered approach used in this development process.

In a user-centered design approach, the target users of the mHealth system are involved and have influence on how the system will be designed and developed. The development process in user-centered approach is iterative. Feedback from the users is analyzed to determine how to proceed with the next cycle of the iterative development process. The need to do the iterative development of the system is determined based on the evaluation and feedback from the target users. This process can be repeated several times as needed. The final goal of this process is to

get a system that is usable and can function as intended to achieve all the defined purposes of the system effectively and efficiently.

## 4.2 Methods

### 4.2.1 Study Description

This chapter aims to discuss the iterative development and usability evaluation of an mHealth system for chronic low back pain assessment. As described in Chapter 3.0, this mHealth system was designed to collect biomechanics and behavioral assessment data during structured assessment session in clinic and during participants' daily life at their home settings. The study was conducted in Pittsburgh, PA. The in-clinic assessment sessions were done at the Physical Therapy – Clinical and Translational Research Center (PT-CTRC).

Two mHealth apps were developed:

1. In-clinic app: an mHealth app for physical therapists to help them conduct the structured assessment session in clinic.
2. At-home app: an mHealth app for chronic low back pain patients to help them record assessment data using their own mobile phones in their daily life settings.

Three physical therapists, who work at PT-CTRC, were tasked to do the in-clinic assessment. They used the in-clinic app to perform the assessment in the clinic and were involved in providing feedback and in usability assessment of the in-clinic app.

In this study, 522 people with chronic low back pain who used the at-home app were asked to evaluate its usability. These 522 participants were part of the larger pool of participants

enrolled in the LB3P research project, which aims to recruit 1000 individuals with chronic low back pain. To be eligible to participate in the study, the participants were required to have chronic low back pain, were 18 years old or older, and can read and speak English. Here, chronic low back pain is defined as back pain that has persisted at least three months and has resulted in pain on at least half the days in the past 6 months.

### 4.2.2 Development of the In-Clinic and At-Home Apps

The main purpose of this mHealth system is to collect biomechanics and behavioral assessment data of people with chronic low back pain in both clinic and home settings. To facilitate this, an in-clinic app was designed and developed for use by physical therapists during clinical exams, functional performance tests, and quantitative sensory testing (QST) of patients. The app enables the therapist to record in-clinic assessment data efficiently.

In addition to the in-clinic app, an at-home app was developed to allow patients to fill out and submit an ecological momentary assessment (EMA) over time, with specified frequency throughout a 7-day assessment period. The EMA included questions about the patient's pain information, activities, and some behavioral routines, which can help to capture the patient's pain experience and daily life activities in naturalistic settings.

The overall mHealth system also consisted of a clinician portal, a backend, and a secure remote database. In this development, the clinician portal was designed for internal use, allowing researchers to access, review, and manage assessment data from both the in-clinic app and at-home app. The backend and secure remote database support data storage and management, data processing, and data analysis for the entire system.

The design framework for the two apps used user-centered design principle. The first iteration of the app development, covered in Chapter 3.0, was done based on the requirement analysis of each app target users' needs. The feedback of the users and usability assessment was used to make changes and improvements to the app during further iterations of development. The development process was iterative, meaning that changes and improvements were implemented as needed, without a set timeline.

During development, feedback from users was assessed as early as possible to support rapid development and get more feedback in an agile development cycle. This approach allowed the development team to incorporate user feedback at various stages of the development process, leading to the creation of a system that is usable, effective, and well-suited to the needs of the target users.

### 4.2.3 Usability and Feedback Evaluation

The mHealth apps were designed using a user-centered approach and the usability was evaluated in iteration. The usability evaluation of the mHealth apps involved the use of the 7-Likert-scale mHealth App Usability Questionnaire (MAUQ) as well as qualitative open-ended questionnaires. For the in-clinic app, the MAUQ for standalone mHealth app used by healthcare providers (Appendix A.1) was used, while the MAUQ for standalone mHealth app used by patients (Appendix B.1) was used for the at-home app. The average scores of the MAUQ components were calculated and examined to assess the overall usability of each app.

Additionally, answers to the open-ended questionnaires were also evaluated to identify any specific usability issues or general feedback that users provided. These open-ended questions were designed to gather feedback on aspects such as the ease of task performance, learnability, and overall user experience. The use of open-ended questionnaires (Appendix A.2 and Appendix B.2) aimed to gather more in-depth information about the user's experience with the app. The questionnaires consist of five questions that covered different aspects of usability:

1. How easy do you think it is to perform tasks using this app?
2. How quickly do you think you can perform tasks using this app after the training?
3. How pleasant is it to use the app to perform the tasks?
4. What do you think can be improved from the app?
5. What is your overall impression in using the app?

The first question asked users to rate how easy they found it to perform tasks using the app. This question aimed to gauge the app's ease of use and the level of difficulty users may have faced when navigating the app's features. The second question focused on the speed of task performance, asking users to rate how quickly they could perform tasks after receiving training on how to use the app. This question aimed to evaluate the app's learnability and the effectiveness of the training provided. The third question inquired about the user's subjective experience of using the app. This question aimed to assess the app's overall user experience and the user's satisfaction with the app's design and interface. The fourth question was an open-ended question that asked users to provide feedback on what they thought could be improved in the app. This question aimed to gather specific feedback on areas of the app that users found problematic or confusing, and to identify areas that needed improvement. Finally, the fifth

question asked users to provide an overall impression of the app. This question aimed to gather a general impression of the user's experience with the app and to identify areas where the app excelled or fell short.

For the in-clinic app, feedback from physical therapists was also collected and examined to gain insights into their experiences using the app. Feedback was communicated directly or through study coordinators. Similarly, feedback from patients was collected and examined for the at-home app. This feedback was obtained directly through the messaging feature and/or the note-taking feature in the app or indirectly through study coordinators. The aim of this feedback was to determine any issues or challenges that patients may have faced while using the app, as well as to identify potential areas of improvement.

Overall, the evaluation and feedback from users played an essential role in the iterative development process of the mHealth system. The insights gained from these evaluations helped to refine the design and functionality of the system, ensuring that it met the needs of both the physical therapists and patients while also remaining usable and efficient. In this development process, there is no specific time allocated to make the revision. Revisions can be made and published anytime, especially if the changes are needed immediately.

## 4.3 Usability Evaluation and Development Iteration

### 4.3.1 In-Clinic App

An initial functional version of the in-clinic app was developed prior to the first usability assessment. This version included several core features and services, such as a login module and

authentication service, a home/main menu page, an exam/test module, a recording module and service, a case report form/questionnaire module and processing service, a local data storage service, a remote communication service, and a safety screening module and processing service. The development of the initial version of the in-clinic app is described and discussed in Chapter 3.4.2.

Several usability evaluations were performed, and changes and revisions were developed based on the results. For the in-clinic app usability evaluation, five questions from the MAUQ were removed for various reasons, which are listed in Table 8.

**Table 8. Excluded MAUQ Questions for In-clinic App Usability Assessment.**

| Removed question | Reason |
|---|---|
| I feel comfortable using this app in social settings. | The physical therapists don't use the app in social settings. |
| The amount of time involved in using this app has been fitting for me. | For each session, the physical therapists use the app for a specific pre-defined time range. |
| The app helped me manage my patients' health effectively. | There isn't any feature to manage patient's health in the app. |
| I could use the app even when the Internet connection was poor or not available. | The physical therapists haven't had any trouble with internet connection so far. |
| This mHealth app provides an acceptable way to deliver healthcare services, such as accessing educational materials, tracking | The physical therapists don't use the app to deliver mentioned services. |

| | |
|---|---|
| my own activities, and performing self-assessment. | |

#### 4.3.1.1 First Usability Evaluation and Development Updates

The first usability evaluation of the in-clinic app was conducted after the initial functional version was developed. Prior to the evaluation, the physical therapists were briefed and presented with the app. During the evaluation, the physical therapists used the app for a mock in-clinic assessment session and were asked to fill out the MAUQ and a custom open-ended questionnaire afterwards.

The first usability assessment of the in-clinic app yielded an average MAUQ usability score of 5.55 (SD = 1.14) as shown in Table 9. This indicates that the physical therapists found the app to be easy to use with a good overall usability score. However, the navigation consistency component received the lowest score. The average score for the question "The navigation was consistent when moving between screens" was 4.50 (SD = 2.12). Even though this doesn't indicate a low usability score, it suggests that some physical therapists may have found it difficult to navigate between screens consistently while using the app. These results suggest that the initial version of the in-clinic app was well-received by physical therapists, but some improvements could be made to the navigation features to make it more user-friendly.

**Table 9. Usability Scores for the First Usability Evaluation of the In-Clinic App**

| **MAUQ Item (see Appendix A.1)** | **Usability Score** |
|---|---|

| | |
|---|---|
| 1 | 5.67 (SD = 0.58) |
| 2 | 6.00 (SD = 1.41) |
| 3 | 4.50 (SD = 2.12) |
| 4 | 5.00 (SD = 0.00) |
| 5 | 5.00 (SD = 1.41) |
| 6 | 4.67 (SD = 1.53) |
| 7 | 6.00 (SD = 0.00) |
| 8 | 5.00 (SD = 0.00) |
| 11 | 7.00 (SD = 0.00) |
| 12 | 6.00 (SD = 0.00) |
| 13 | 6.00 (SD = 0.00) |
| 14 | 6.00 (SD = 0.00) |
| 16 | 5.00 (SD = 0.00) |
| Overall | 5.55 (SD = 1.14) |

The open-ended questionnaires revealed that the overall impression of the physical therapists was positive. They expressed confidence that they would be able to use the app easily if given more time to try it out, and they believed that the app would be helpful for them. However, some physical therapists mentioned issues with inconsistency in navigating within a test module and unresponsiveness in some processes. One physical therapist specifically suggested that "*consistent back access to the previous screen*" could be improved in the app.

The feedback from one of the physical therapists also indicated an issue with the app's responsiveness. They reported experiencing a delay in button responsiveness, which could potentially cause frustration and hinder the overall usability of the app. They mentioned that they experienced "*small delay in responsiveness of buttons today*". To investigate the issue, the troubleshooting was conducted, and it was discovered that the delay was caused by a process that can take more than one second to finish.

Based on the results of the usability assessment and the feedback from the physical therapists, several changes were implemented to improve the app's usability. These changes included:

- Adding the ability to skip some tests and move to the following test.
- Updating the back navigation to be consistent and similar to the upward flow.
- Adding navigation to go back and forth between sub-tests in a test group, making it easier for users to navigate within a test group.
- Adding a loading spinner to processes that take more than 1 second, indicating that the app is still processing information and is not frozen or unresponsive.
- Adding confirmation dialogs to inform users that they are about to navigate or perform an action in the app, providing additional clarity and preventing accidental actions.

Several modules and pages in the in-clinic app were updated and revised to address the identified usability issue. The main changes were made in the test module, which had been identified as a problem area for the navigation issue. The changes made were aimed at improving the navigation and flow inside the test module. Back button functionality was revised to provide consistent and intuitive back navigation. Back buttons that navigate to recording page and case report form page were removed to prevent overwriting the data that has been submitted. Skip button was also developed to enable the physical therapist to skip a test and navigate to the following test. Skip button was developed to be configurable to enable unskippable test. The changes to address the navigation issues are illustrated in Figure 12.

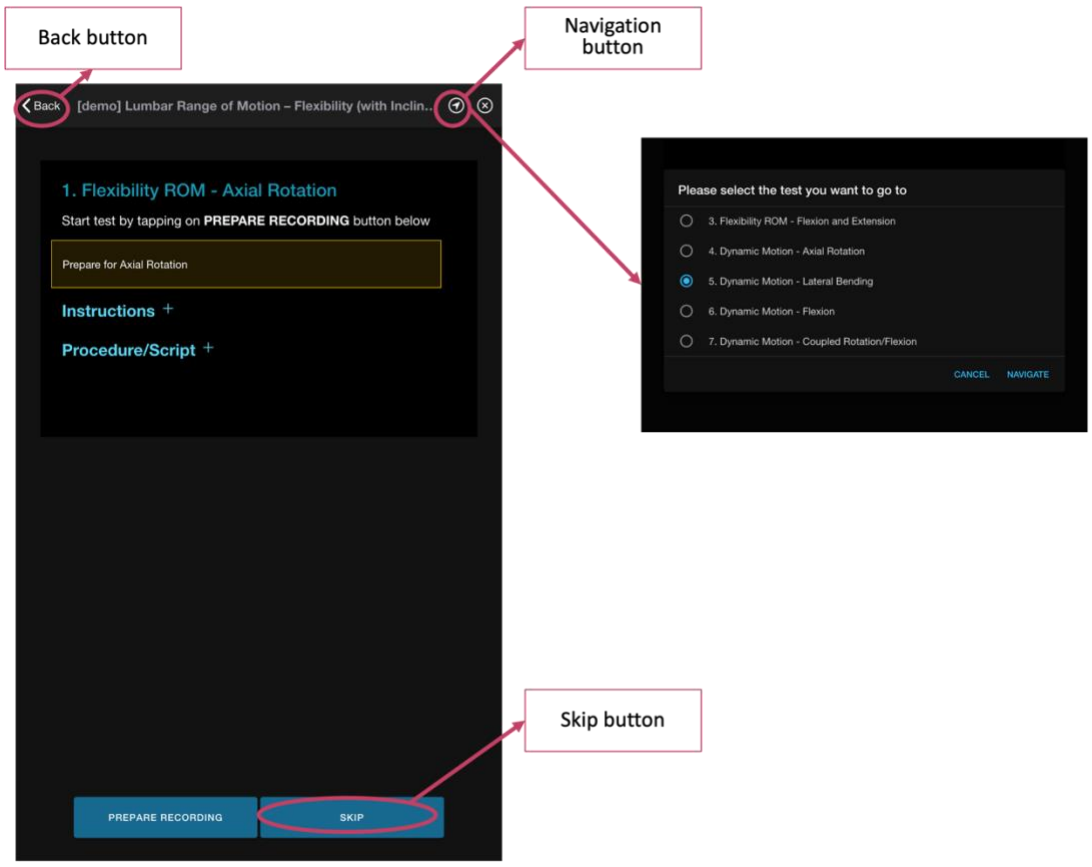


**Figure 12. Navigation changes in the in-clinic app.**

Additionally, the app was updated to address the issue of unresponsiveness that had been reported by one of the physical therapists. To mitigate this issue, loading spinners were incorporated in pages that require data retrieval or processing and take more than one second to complete. The loading spinners were added to inform the user that the app is processing the requested information and is not frozen or unresponsive. This new feature was implemented in various sections of the app, including the home/main page. Loading spinners were added in the home page to indicate that the app is retrieving study participants' information. The implementation of this loading spinners is shown in Figure 13.

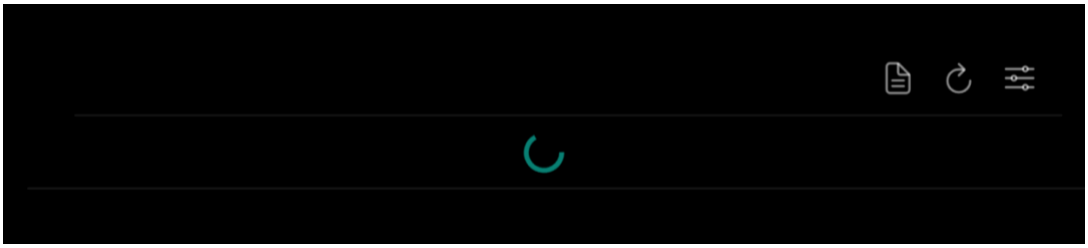

**Figure 13. Example of spinner implementation when the app is retrieving and processing data.**

Another new feature that was developed to address the usability issues identified from the first usability assessment was confirmation dialogs. Confirmation dialogs were developed to inform users that they are about to navigate or perform an action in the app, providing additional clarity and preventing accidental actions. Confirmation dialogs have been implemented in several modules and pages to address usability issues. For instance, in the test module, a confirmation dialog will appear when the exit button is clicked to ensure the user's intention and prevent data loss. In the safety screening module, a confirmation dialog will appear when the submit button is clicked to notify the user that the data has been saved. The examples of the implementation of this confirmation dialogue are shown in Figure 14.

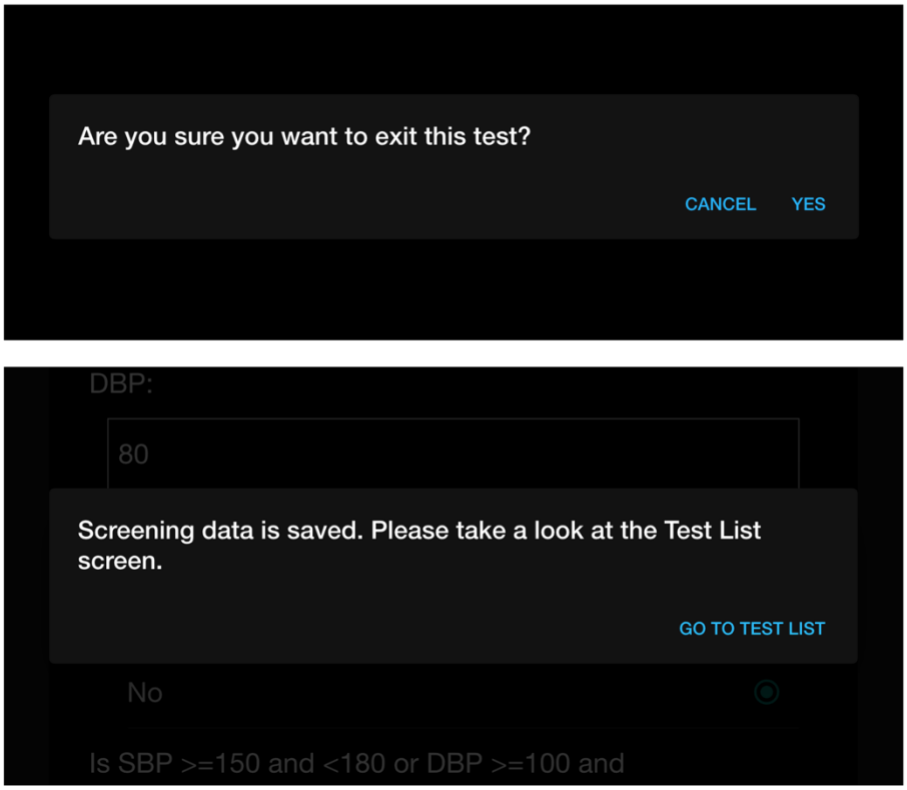


**Figure 14. Example of confirmation dialog that shows up after the user submit safety screening questionnaire responses.**

#### 4.3.1.2 Second Usability Evaluation and Development Updates

The second usability evaluation was conducted three weeks after the first assessment, giving the physical therapists ample time to explore and interact with the app independently. This approach enabled the therapists to develop a deeper understanding of the app's features and functionalities, thereby allowing them to evaluate the usability of the app more accurately. Only one physical therapist participated in the second usability evaluation. The average MAUQ usability score for the second evaluation was 6.00 (SD=1.15), which showed a marked improvement from the first assessment. These results suggest that the changes made to the app

based on the feedback from the physical therapists have resulted in a more user-friendly app. The usability scores can be seen in Table 10.

**Table 10. Usability Scores for the Second Usability Evaluation of the In-Clinic App**

| MAUQ Item (see Appendix A.1) | Usability Score |
|---|---|
| 1 | 5 |
| 2 | 7 |
| 3 | 6 |
| 4 | 7 |
| 5 | 6 |
| 6 | 7 |
| 7 | 4 |
| 8 | 7 |
| 11 | 7 |
| 12 | 6 |
| 13 | 7 |
| 14 | 5 |
| 16 | 4 |
| Overall | 6.00 (SD = 1.15) |

Based on the open-ended questionnaires, the physical therapists had an overall positive impression of the app and appreciated the changes and improvements made to it. However, they did mention a couple of issues that they would like to see addressed:

- They felt that sometimes the processes in the app were a bit slow.

- They suggested adjustments to be made in the case report form module. They found it exhausting and time-consuming to answer all the questions, especially when some of the questions were not relevant based on the test result. For example, if a participant decided not to attempt a test, the physical therapist still needed to go through all the questions and fill them out.
- They mentioned that they needed a bigger interface in general because they often had to stand next to the participant while also monitoring the app from a distance.

Some changes were made after evaluating the usability assessment and the feedback from the physical therapists. A general review of the app's performance was conducted, and several improvements were made to enhance its performance. The changes made included reducing the time needed to load participant data from the remote backend by only retrieving necessary data at a time. A change in the backend was also made to revise the APIs to send only necessary data per request made by the app.

Additionally, the case report form module was improved by restructuring the questions to show up conditionally based on several parameters, such as time recorded and answers to other related questions. The questions were also organized in a way that would reduce the burden on the physical therapists when they fill out the case report form. Several UI changes and improvements were also made, with a focus on making the fonts bigger, especially on the recording page because the physical therapist needs to be standing near the participant, away from the app, when they perform the test that is being recorded for safety purposes. Figure 15 shows examples of the changes in the app interface.

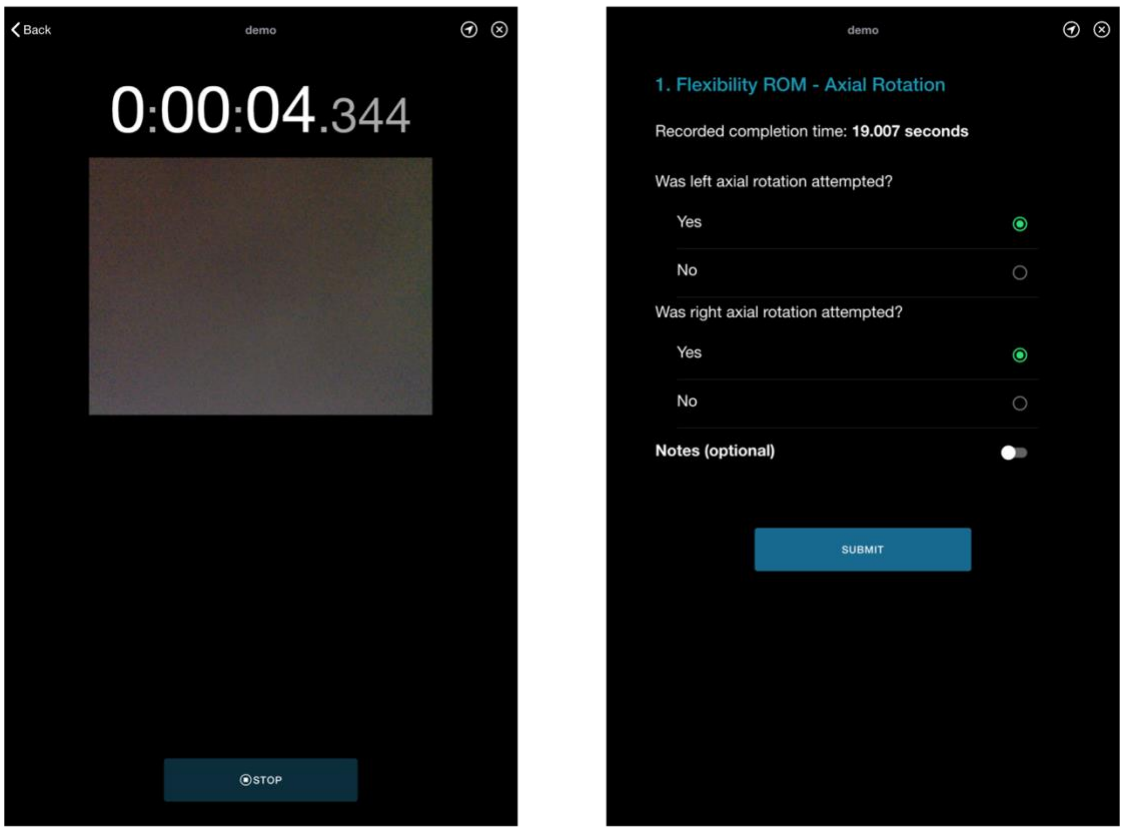


**Figure 15. Updated interface and bigger font to accommodate viewing from distance.**

#### 4.3.1.3 Additional Developments

Since the usability scores from the first and second usability assessments showed good usability of the app, no further usability assessments were conducted. However, feedback from physical therapists was still collected and assessed, and changes and improvements were continuously made based on their feedback. Physical therapists provided their feedback either directly or through the study coordinator. Based on this feedback, several updates and improvements were implemented, including updates to the test structure, development of offline data collection capability, creation of test deviation report processing, and updating the protocol for handling not attempted tests. Additionally, several bug fixes and UI improvements were

made based on feedback from the physical therapists. Some screenshots of the latest version of the in-clinic app can be seen in Figure 16.

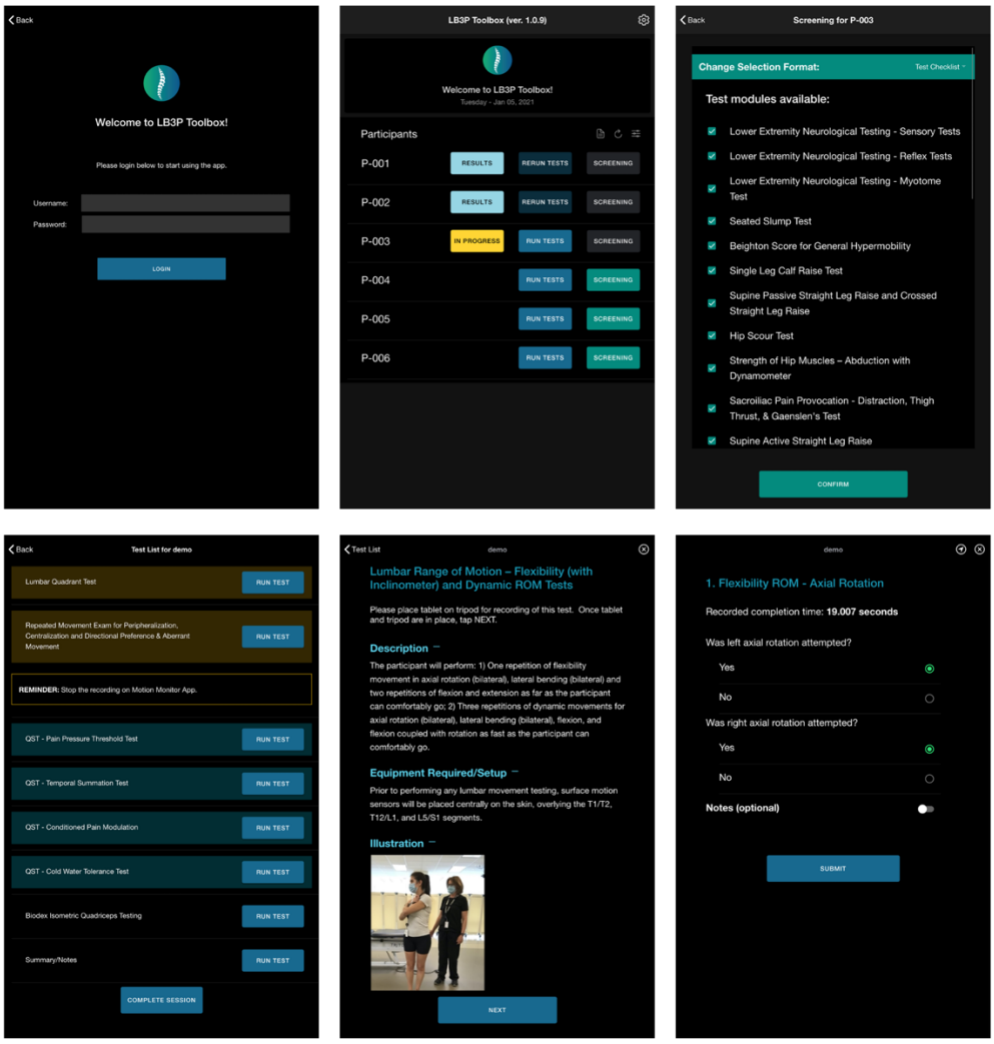


**Figure 16. Screenshots of the latest version of the in-clinic app.**

**The screenshots on the first row, from left to right, are login page, home page, and screening module. The screenshots on the second row, from left to right, are test list page, test description/information page, and case report form page.**

In addition to the improvements made to address app usability issues, changes were also made to expand the app's capabilities for use in other similar studies. The modifications included

adding different types of participants with their own set of tests, which required adjustments to the login and authentication module to allow users to access different studies within the same app.

Furthermore, for one of the studies, the app was developed to be able to communicate and process data from kinematics sensors using Bluetooth Low Energy connection. The in-clinic app was planned to have this capability, but the APIs needed to communicate with the sensors were not made available during the development of the initial app. The fact that the sensor integration framework design was feasible with the availability of communication protocols between the sensors and the app validated the approach taken in the development of the app.

Overall, these changes enabled the app to be used for a wider range of studies (Figure 17), demonstrating the flexibility and adaptability of the app's design. The incorporation of additional features such as sensor integration could potentially enhance the app's capabilities and broaden its scope for use in future studies.

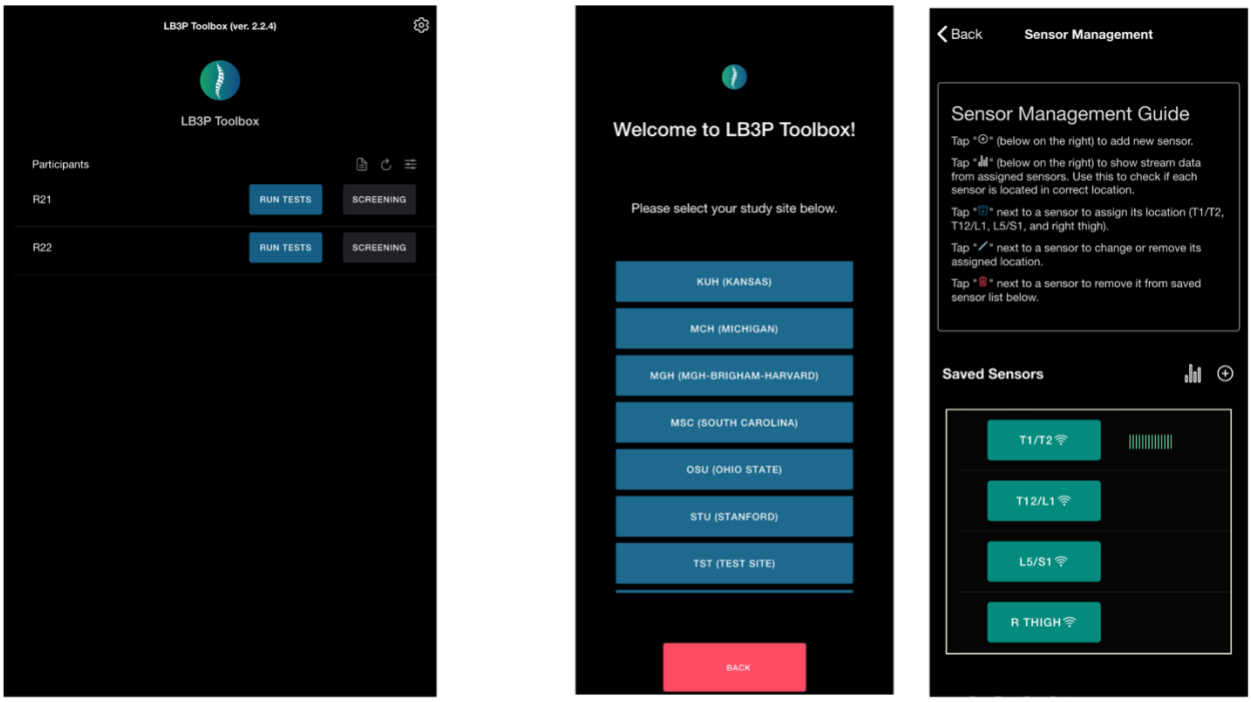


**Figure 17. Development of the in-clinic app for use in different studies.**

**The app was developed to be used in Repeatability study (A) and BEST (Biomarkers for Evaluating Spine Treatment) study (B). Sensor communication was developed in the in-clinic app for BEST study.**

### 4.3.2 At-Home App

An initial functional version of the at-home app was developed prior to the first usability assessment. The main core features and services developed in this version are login module and authentication service, home/main menu page, EMA module, reminder module, notes module, messaging module, account and settings modules, local data storage service, and remote communication service. A usability questionnaire module was also developed and incorporated inside the app, which enabled the patients to participate in the usability evaluation using the app.

The development of the initial version of the at-home app is described and discussed in Chapter 3.4.3.

The usability module was presented to the patients after they had completed the 7-day at-home assessment session. By filling out the usability questionnaire at the end of their assessment session, the patients had sufficient time to fully experience using the app. Although filling out the usability questionnaire was optional, the patients were encouraged to do so by the app. A total of 337 out of 522 patients filled out the MAUQ questionnaire, while 305 out of 522 patients filled out the open-ended usability questionnaire. The open-ended questionnaire module was presented after the patients had completed the MAUQ module, and the patients could opt not to fill out any questionnaire. There was no additional compensation for the patients if they filled out the questionnaires.

For the MAUQ questionnaire, there were two questions removed because they are not relevant to the app being assessed. The questions removed and the reason why they were removed were listed in Table 11.

**Table 11. Excluded MAUQ Questions for At-Home App Usability Assessment.**

| Removed question | Reason |
|---|---|
| The app improved my access to healthcare services. | The app doesn't provide any access to healthcare services. |
| The app helped me manage my health effectively. | The app is not designed to be used to manage the patient's health. |

#### 4.3.2.1 First Usability Evaluation and Development Updates

When the first usability assessment was conducted, there were 68 patients who had submitted their questionnaire responses. The average MAUQ usability score of the first assessment was 6.40 (SD=1.12). This result suggests that the initial functional version of the at-home app had good usability. The lowest scored component still had a relatively high average score of 5.74 (SD=1.52). The component with the lowest score in this usability assessment was "This mHealth app provides an acceptable way to receive healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment."

The answers to the open-ended questionnaire, except for the fourth question, were categorized as positive, negative, neutral, or unrelated. 86.29% of the answers were positive and 2.82% were negative. Based on these results, the overall impression of the app by the patients was good. Some patients suggested improvements for the app, such as the ability to fill out the EMA retrospectively. For example, one participant mentioned, *"...I missed a mid-day one, but once it reached time for the evening one, it would not let me fill out the mid-day one, even though I knew the information.*" However, this feedback was not incorporated because the purpose of EMA is to assess the patients in real-time, rather than retrospectively.

Several changes were made to the app based on feedback from the patients and study coordinators. The Account module was removed because the study account information was found to be not relevant for the patients, and its removal helped to avoid confusion. An FAQs module was added to provide patients with more information about the study. A screenshot of the FAQs module is provided in Figure 18. With the addition of the FAQs module and the removal of Account, the home page of the at-home app looked like the screenshot in Figure 19.

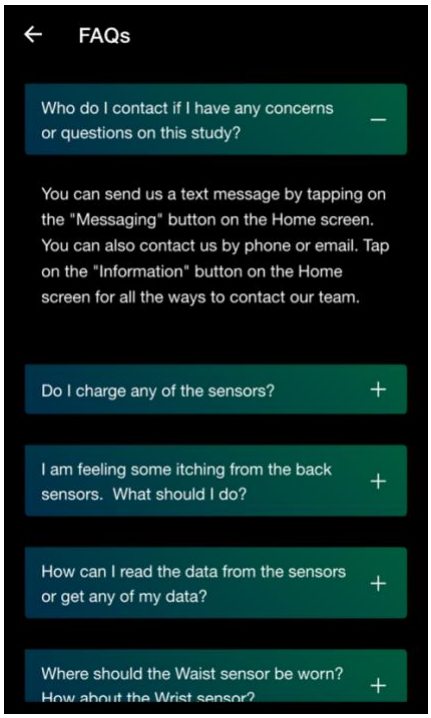


**Figure 18. FAQs module in the at-home app.**

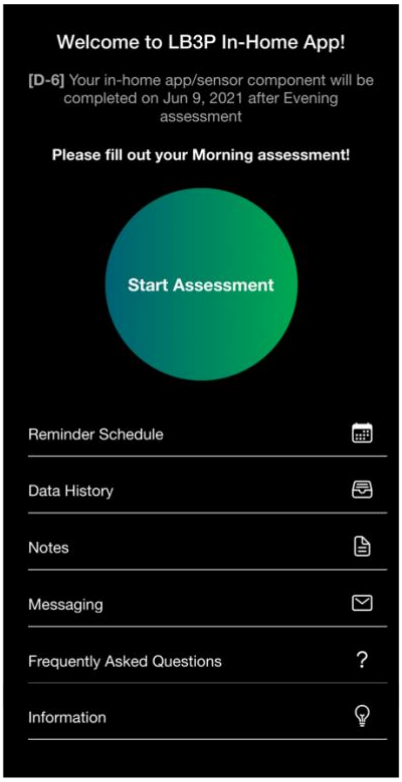


**Figure 19. Home page of the at-home app after changes were implemented.**

An assessment summary module was added to the app, which shows patients their compliance with filling out the EMA the day before. This module appears when patients open the app for the first time each day. Encouragement messages tailored to the patient's compliance were added to encourage them to follow the study protocols. The illustration of the assessment summary module was shown in Figure 20.


**Figure 20. Assessment review module in the at-home app.**

Additionally, a sensor deviation reporting module was added to the app, as the study coordinators wanted to know and to be able to monitor whether the patients were wearing their sensors. This module allows the patients to report when they are not wearing their sensors, so they can contact them and address any issues they may have. The module is shown in Figure 21.

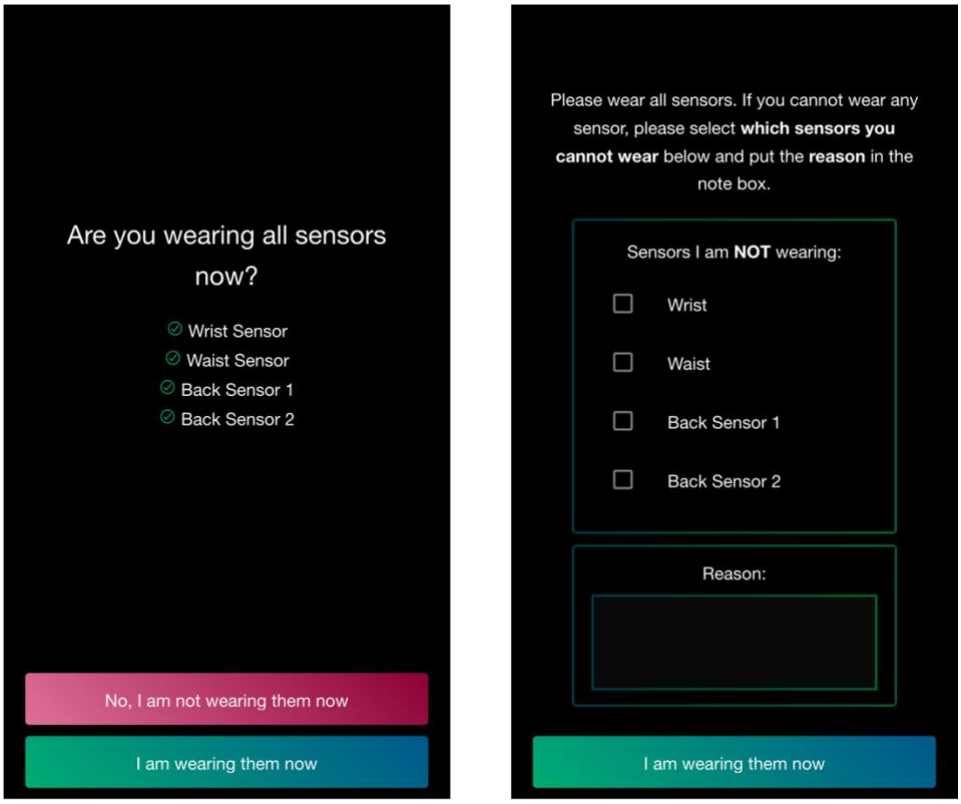


**Figure 21. Sensor deviation reporting module in the at-home app.**

#### 4.3.2.2 Second Usability Evaluation and Development Updates

After three months following the first at-home app usability evaluation, 62 patients filled out the usability questionnaire. The average score was 6.39 (SD=1.15), which did not change much from the first assessment. Of the open-ended questionnaires, 87.95% of answers were positive, and only 0.89% of answers were negative.

Some changes were made to the app based on feedback from patients and study coordinators. Study coordinators wanted to ensure that the patients know that the notes they filled in inside the app were not monitored during the weekend. To further assist the study coordinators in interacting with the patients, a new flagging feature was implemented in the

notes and messaging module. Any notes or messages containing any words of interest for sensor and skin issues would be flagged so that the study coordinators could focus more on those notes or messages. The flagging words include burn, itch, skin, allergy, sensor, irritation, red, tape, and rash. If the patients send any flagged notes and messages, the system will also send email notifications to the study coordinators.

These changes didn't impact the interface of the at-home app. Even though some changes were made in the note and messaging modules in the app, no visual changes needed to be developed. The patients would have the same experience in using the app before and after these changes. On the other hand, the clinician portal was updated to enable the study coordinator to see the flags on the patients' notes and messages as illustrated in Figure 22.

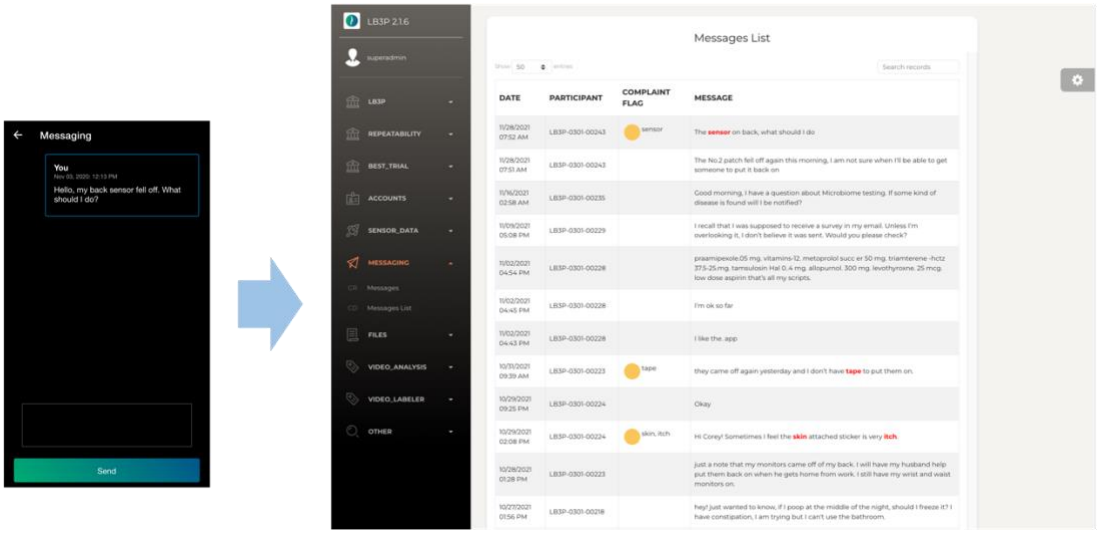


**Figure 22. Message flagging feature.**

#### 4.3.2.3 Further Usability Evaluation and Development Updates

The third usability assessment was conducted three months after the second assessment. Fifty-four patients completed the usability questionnaires, and an average score of 6.23 (SD=1.28) was calculated. Although slightly lower than the previous usability scores, this still indicates good usability. Of the open-ended questionnaire answers, 84.31% were positive and only 4.90% were negative. No changes were deemed necessary for this iteration.

A fourth analysis of the usability assessment was conducted three months after the third one. Eighty patients completed the questionnaires, and the average score was 6.24 (SD=1.43). This score did not show a significant difference from the previous score. 89.18% of answers to the open-ended questionnaires were positive, and only 3.73% of the answers were negative. Based on the usability assessment results, it can be seen that the scores plateaued around these numbers.

A fifth usability assessment analysis was conducted three months later, and the results were similar to previous assessments. Seventy-three patients completed the questionnaire, and the average score was 6.24 (SD=1.37), which is quite similar to the previous score. Of the open-ended questionnaire answers, 88.04% were positive, and 2.90% were negative. No changes were made to the app after the fourth and fifth usability evaluations. The app was able to serve its purposes to patients, study coordinators, and researchers who collected the data submitted in the at-home app. Some screenshots of the latest version of the at-home app can be seen in Figure 23. Overall, the at-home app had a good usability score even for the first functional version of the app.

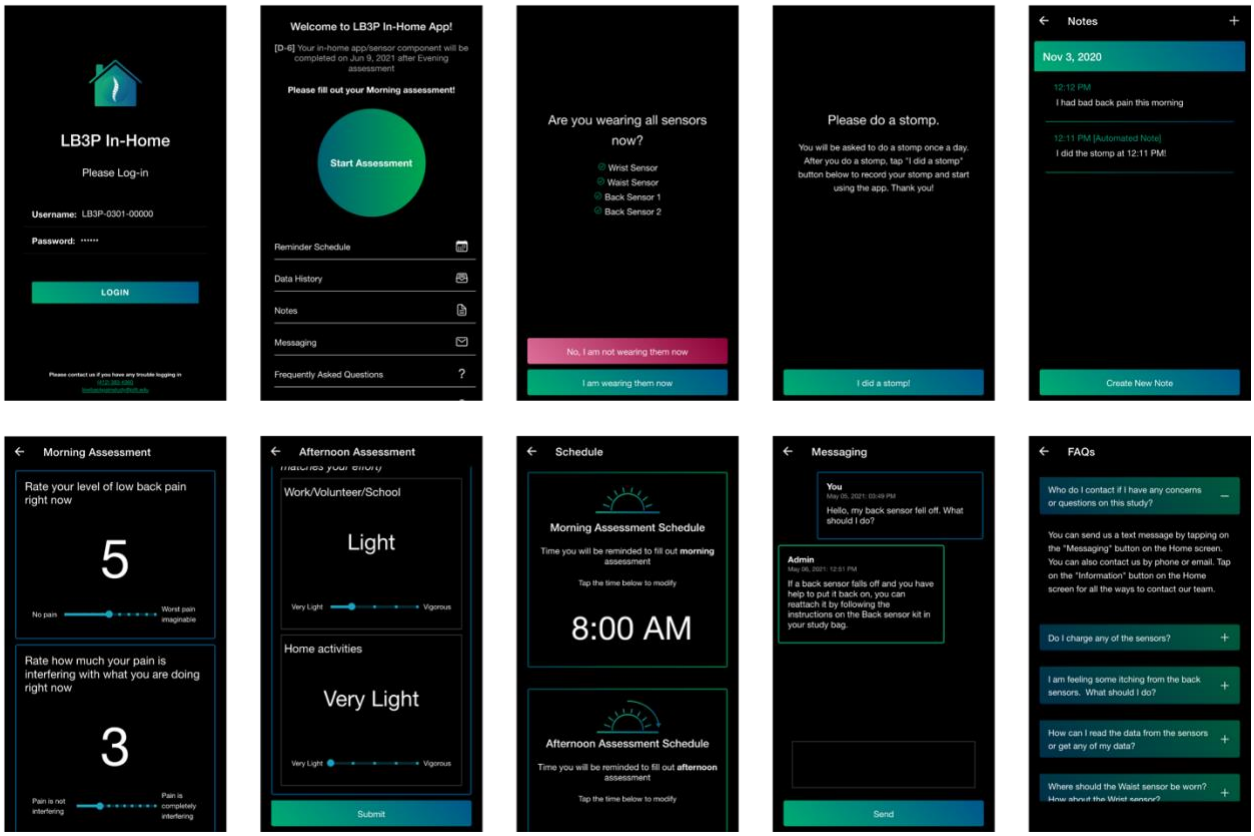


**Figure 23. Screenshots of the modules in the at-home app.**

**The screenshots on the first row, from left to right, are login page, home/main page, sensor deviation reporting module, stomp reporting module, and note module. The screenshots on the second row, from left to right, are morning assessment module, afternoon assessment module, reminder scheduling module, messaging module, and FAQs module.**

## 4.4 Discussion

### 4.4.1 Principal Results

In this study, an mHealth system was designed and developed for a comprehensive low back pain assessment. The system consists of several components that are linked into one connected system. One component is an in-clinic app that was developed to be used by physical therapists to perform structured in-clinic assessments for patients. Another component is an at-home app that was developed for patients to collect their EMA responses. Both apps were designed and developed with a focus on target users' needs. They were initially designed and developed based on the requirement analysis. Iterative revisions were made to both apps as usability was assessed and feedback from target users was investigated.

The in-clinic app was successfully utilized by physical therapists to conduct the in-clinic assessments, and they had a positive impression of the app, finding it to be helpful. To date, the physical therapists have used the app to assess over 500 patients. With the app's general assessment functionality, other studies can use it for their own assessments, as the app can be tailored to meet their specific needs. In fact, one current study is using the app, which had been updated to incorporate kinematics sensors integration, for their assessment. This demonstrates the original framework with sensors integration can be implemented.

This study also found that the usability of the in-clinic app was high, with the latest usability evaluation score of 6.00 (SD=1.15). The feedback provided by the physical therapists through the questionnaires or directly to the author was very helpful in iterative development of the app. Several changes were made through multiple iterations to make the app more usable and helpful for the physical therapists.

Similar to the in-clinic app, the at-home app was also able to fulfill its purpose. It helped patients fill out and submit their EMA and allowed study coordinators to monitor and communicate with the patients. The at-home app also got a high usability score of 6.24 (SD=1.37) based on the latest usability evaluation and had an overall positive impression from patients. The app underwent multiple iterations of changes, with the last changes made after the second usability assessment. Although third, fourth, and fifth usability assessments were also conducted, no changes or adjustments were necessary to be developed for the app.

### 4.4.2 Limitations

One limitation of the development of the system is that there were no APIs provided to communicate with the sensors, which were acquired from a third-party provider. This lack of communication protocol between the app and the sensors prevented the integration between the two components. The planned framework was for the app to communicate and process kinematics data from the sensors, but the app was unable to do so due to the inability to communicate directly with the sensors. Although access and APIs for the sensors were provided at a later time, it was decided not to implement the integration with sensors for this study for consistency. However, this integration with the sensors was able to be implemented for a different study that also used this in-clinic app.

Another limitation was related to the participant recruitment for the usability assessment of the in-clinic app. Physical therapists were recruited by the clinic, and convenience sampling was used to select participants. Only three physical therapists were involved in the study, which is a small number and could limit the generalizability of the results. Moreover, the working relationship between the author and the physical therapists could introduce bias. After the second

usability assessment for the in-clinic app, the author decided to focus more on getting direct feedback from the physical therapists to expedite improvement and overall development, rather than asking them to fill out usability questionnaires.

For the usability assessment of the at-home app, not all patients filled out the usability questionnaires, and some even mentioned that they felt burdened by the many questions in the questionnaires. There was also a limitation related to the open-ended questionnaire, which was given to the patients to fill out. It would have been better to be able to interview the patients directly, as this would have resulted in a more extensive qualitative analysis.

### 4.4.3 Conclusions

This study highlights the successful development of an mHealth system for the comprehensive assessment of low back pain using a user-centered design approach. The mHealth system, especially the in-clinic and at-home apps, can deliver their purposes to their main target users: physical therapists and low back pain patients. One of the key findings of this study was the high usability scores of both the in-clinic and at-home apps. This was a crucial aspect of the development process, as it ensured that the apps were user-friendly and easy to use. Both the physical therapists and patients had positive impressions of the apps, and the usability assessments provided valuable feedback for the development team. The iterative design process enabled the development team to refine the apps and tailor them to the needs of their users.

There are several opportunities for further development and expansion of the mHealth system. One potential area for improvement is the in-clinic app, which has already demonstrated its flexibility and configurability. As such, it could be utilized for assessment purposes in other studies, with additional modules added and tailored for different uses.

Another potential area for development is the at-home app. While it has already proven effective in helping patients report their assessment, there is potential to expand its functionality to include treatment plans. The assessment data collected in this study was extensive and rich. This can be utilized to develop a machine learning component that can be used to create a personalized and adaptive treatment for chronic low back pain. Using this machine learning component, the at-home app can be improved to also incorporate a personalized and adaptive treatment component. This can further improve the system to not only assess, but also provide and deliver effective treatments for chronic low back pain patients.

# 5.0 Exploration for Future Works: Comparison of Perceived Physical Activity Levels with Calculated Activity Levels from Kinematics Sensors

## 5.1 Introduction

Physical activity is an important component of chronic low back pain management and prevention. Accurate measurement of physical activity levels is crucial for developing effective treatment for individuals with chronic low back pain. For example, measurement of one's physical activity level is important to formulate effective exercise interventions for that individual. For chronic low back pain phenotyping, physical activity is one of many characteristics that can be used to identify subgroups of chronic low back pain.

The mHealth system that is discussed in Chapter 3.0 and Chapter 4.0 was used to collect chronic low back pain assessment data from both in-clinic and at-home assessment to be used for phenotyping purposes. Extensive and comprehensive data were collected. For the at-home assessment data, kinematics data from wearable kinematics sensors and behavioral data, such as pain and physical activity, from patient self-reported EMA were collected. Kinematics sensors have become a popular method of measuring physical activity levels objectively. However, it is not clear how well these data from the sensors correlate with the individuals' perceived physical activity levels. The data collected by the mHealth system in this study can be used to investigate the correlation between the two types of data: objective physical activity level from the kinematics sensors and the perceived physical activity level from the patients' self-reported EMA.

## 5.2 Methods

### 5.2.1 Study Description

This chapter aims to discuss the correlation between objective physical activity level from the kinematics sensors and the perceived physical activity level from the patients' self-reported EMA. The data was collected using the mHealth system that is discussed in Chapter 3.0 and Chapter 4.0 during the at-home assessment session. Data from 12 individuals with chronic low back pain was collected. These 12 individuals are part of the participants recruited for the Repeatability research. To be eligible to participate in the study and be selected in this data analysis, the participants were required to be between 18 to 70 years old, be without exercise or activity restriction, and have chronic low back pain. Here, chronic low back pain is defined as back pain that has persisted at least three months and has resulted in pain on at least half the days in the past 6 months.

In this study, the participants were given the at-home app to be installed in their mobile phone. Both Android and iOS versions of the app are available to enable the participants to use their own phone, whether it's an Android phone or an iPhone. The at-home app was available in both Google Play Store and Apple App Store to make it easier for the participants to download and install the app. In the Google Play Store, the app is listed as LB3P In-Home (https://play.google.com/store/apps/details?id=org.harilab.lb3pinhome). In the Apple App Store, the app is listed as LB3P In-Home (https://apps.apple.com/us/app/lb3p-in-home/id1537128941). Both versions of the app are publicly available, but the app requires authentication to ensure only the study participants can use the app.

During the at-home assessment period, the participants were asked to report their behavioral information three times per day: morning, afternoon, and evening. The participants can use the EMA module in the app to fill out and submit their behavioral information. Each assessment has its own set of questions. Table 12 lists the assessment items and the corresponding response options in each EMA type (See Appendix C for the complete EMA forms).

**Table 12. Assessment Items in the At-Home EMA Module**

| EMA Type | Assessment item | Response Options |
|---|---|---|
| Morning EMA | Current level of low back pain | 0 (No pain) – 10 (Worst pain imaginable) |
| | Interfering pain level | 0 (No pain) – 10 (Worst pain imaginable) |
| | Sleep time | Time |
| | Wake up time | Time |
| Afternoon EMA | Current level of low back pain | 0 (No pain) – 10 (Worst pain imaginable) |
| | Interfering pain level | 0 (No pain) – 10 (Worst pain imaginable) |
| | Morning activities | • Sports/Exercise<br>• Hobbies<br>• Work, School, or Volunteer<br>• Home activities |

| | | |
|---|---|---|
| | Morning activities level | For every activity:<br>• Very light<br>• Light<br>• Moderate<br>• Moderate to vigorous<br>• Vigorous |
| Evening EMA | Current level of low back pain | 0 (No pain) – 10 (Worst pain imaginable) |
| | Interfering pain level | 0 (No pain) – 10 (Worst pain imaginable) |
| | Afternoon activities | • Sports/Exercise<br>• Hobbies<br>• Work, School, or Volunteer<br>Home activities |
| | Afternoon activities level | For every activity:<br>• Very light<br>• Light<br>• Moderate<br>• Moderate to vigorous<br>• Vigorous |
| | Is typical day? | • Yes<br>• No |

Participants' responses on the morning and afternoon activities level assessment items were used as the physical activity level that was compared to the activity level derived from the kinematics data from the sensors. Participants reported their morning and afternoon activities level every day for the whole 7-day assessment period.

During this assessment period, the participants were also asked to wear two wearable kinematics sensors (Lifeware Labs, LLC, Pittsburgh, PA) for 7-day period. The sensors were attached to the participants at T12/L1 interspinous space and L5/S1 interspinous space. The kinematics sensors record three-axis acceleration, angular velocity, and magnetic field raw data at a rate of 20 Hz. In the 7-day at-home assessment period, the participants were not instructed to do any specific physical exercises or activities. They can do their normal daily activities during this assessment period.

After the 7-day at-home assessment period, the participants send back the sensors. Sensor data was downloaded to a local computer directly. The data was then uploaded and stored in the study remote cloud storage. The sensor data is in space delimited format. It has elapsed time information since the sensors were turned on, accelerometer data in gravity ($1g = 9.80665 m/s^2$), gyroscope data in degree/second, and magnetometer data in gauss (1 gauss = $10^{-4}$ tesla). For this exploration study, only accelerometer data from the L1 sensors was calculated for the analysis. Accelerometer data was used because physical activity level can be calculated from accelerometer data. It is common to use an accelerometer to measure the movement and physical activity level of an individual. Activity-related information, such as frequency, intensity, and duration of movement, can be derived from accelerometer data, which can be used to estimate the physical activity level of an individual.

### 5.2.2 Data Analysis

The accelerometer data was processed to estimate the physical activity level of participants with chronic low back pain. The y-axis of the accelerometer data was used for analysis. In the case of individuals with chronic low back pain, the y-axis is commonly used for activity analysis due to the assumption that the vertical motion in this axis reflects changes in posture and the resultant mechanical stress on the lumbar spine. There are several algorithms and cutpoints that have been developed to translate accelerometer data into physical activity level estimates. One commonly used algorithm for calculating physical activity level from accelerometer data is the ActiGraph algorithm (Neishabouri et al., 2022). Overall, the following steps were taken to process the accelerometer data:

1. Timestamp generation: A timestamp was generated for each row of the raw accelerometer data. The time at which the participant first used the app was used as the base, and the remaining timestamps were calculated using the elapsed time information in the raw data.
2. Resampling: The accelerometer data was resampled to 60 Hz. Any resulting empty data rows were filled using the "forward fill" strategy, which assigns the value of the previous non-empty row to the empty row.
3. Activity count calculation: The publicly available Actigraph Python module, agcounts (https://github.com/actigraph/agcounts) (Neishabouri et al., 2022), was used to calculate the activity counts per minute. Activity counts refer to the number of movements detected by an accelerometer over a specified period of time.

4. Activity level estimation: The activity counts were then categorized into several intensity categories based on Freedson's cutpoint (Freedson et al., 1998) (Table 13). The activity level was coded to a numerical value. The coding is also listed in Table 13.
5. Data aggregation: To be comparable to the morning/afternoon time type in the EMA, the data was aggregated to hourly intervals between 7 am to 5 pm for the 7-day assessment period, in which 7 am to 12 pm represents the morning timespan and 12 pm to 5 pm represents the afternoon timespan. This involved transforming the activity count data into hourly data and calculating the number of counts for each 60s-epoch activity level in Freedson's cutpoint for each hour.

**Table 13. Freedson's Activity Level Cutpoint and Coding**

| Activity Level | Activity Counts Cutpoint | Coding |
|---|---|---|
| Sedentary | 0 - 99 | 1 |
| Light | 100 - 1951 | 2 |
| Moderate | 1952 – 5724 | 3 |
| Vigorous | 5725 - 9498 | 4 |
| Very vigorous | >9499 | 5 |

This process takes raw sensor data (snippet of raw sensor data can be seen in Appendix D.1) from the participants and process it to an aggregated activity level data (snippet of the processed data can be seen in Appendix D.2) for all participants. Other than the activity level data, the process also retained the 60s-epoch activity counts data calculated in the activity count

calculation step. This sensor data processing was performed using Python (https://www.python.org/) and several data processing and analysis libraries: Pandas (https://pandas.pydata.org/), SciPy (https://scipy.org/), and NumPy (https://numpy.org/). The process is illustrated in Figure 24.

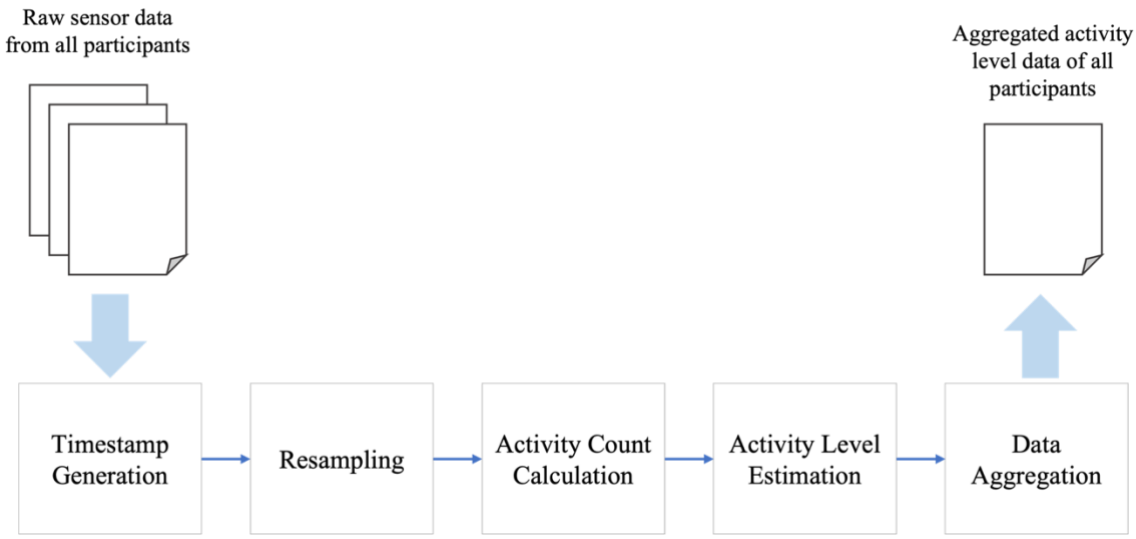


**Figure 24. Diagram of the sensor data processing.**

To extract the perceived activity level data, the following steps were taken:

1. Data Retrieval: EMA data for all participants was queried and retrieved from the study remote database.
2. Data filtering: The data was filtered to include only the responses on activity level in the morning and afternoon. Data from the first day of their at-home assessment period was excluded since the sensor data from the first day might not be complete.
3. Activity level coding: The activity level data was coded into a numerical value. The coding is listed in Table 14.

4. Data aggregation: The coded activity level data for each morning and afternoon from day 2 to day 7 of all participants were combined into a single data file. The resulting file contained the morning and afternoon activity level data for all participants over the 7-day assessment period.

**Table 14. Coding for Activity Level from EMA**

| Activity Level | Coding |
|---|---|
| Very light | 1 |
| Light | 2 |
| Moderate | 3 |
| Moderate to vigorous | 4 |
| Vigorous | 5 |

Similar to the sensor data processing, the processing of activity level data from EMA was performed using Python and several data processing and analysis libraries: Pandas, SciPy, and NumPy. The process is illustrated in Figure 25.

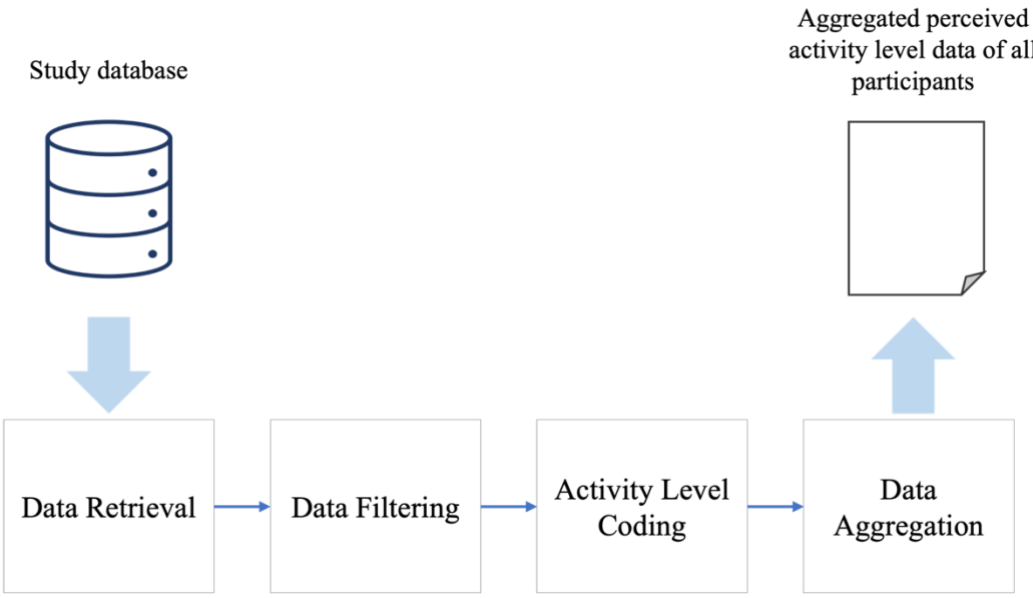


**Figure 25. Diagram of perceived activity level data processing from EMA.**

The processed activity level data from sensors and the processed perceived activity level data from EMA were further processed and aggregated into one data to be compared. Both data were put into an Excel (https://www.microsoft.com/en/microsoft-365/excel) file to be integrated in one sheet of data. For the perceived activity level data, one more step of data processing was done. Since the participants can report more than one activity level per report, if a row in the perceived activity level data had more than one response, the response with the higher activity level was selected. For example, if a participant reported 'very light' for home activities and 'moderate' for sports/exercise, the 'moderate' (coded as 3) response was selected for that row.

To be able to be compared with the perceived activity level data, which only covers broad timepoints of morning and afternoon, the activity level data from the sensors had to be further processed to align with the morning/afternoon granularity. This involved categorizing the timespan of each hourly activity level sensor data and processing the resultant values to align with the morning/afternoon timepoints. For categorization, activity level data from 7 am to 12

pm was categorized as morning and activity level data from 12 pm to 5 pm was categorized as afternoon. To calculate the representation of morning/afternoon activity level, several strategies were used:

1. Using the average of 60s-epoch activity level.
2. Using the most frequent 60s-epoch activity level.
3. Using the highest 60s-epoch activity level.
4. Using the average value of hourly activity counts.
5. Using the activity level calculated from the average value of hourly activity counts.
6. Using the highest value of hourly activity counts.
7. Using the activity level calculated from the highest value of hourly activity counts.
8. Using the weighted sum of average hourly proportion of each activity level and its corresponding ranking multiplier value.
9. Using the weighted sum of the highest hourly proportion of each activity level and its corresponding ranking multiplier value.
10. Using the weighted sum of average hourly proportion of each activity level and its corresponding custom multiplier value.
11. Using the weighted sum of the highest hourly proportion of each activity level and its corresponding custom multiplier value.

The integration process is illustrated in Figure 26. The activity level represented by the values calculated using those strategies were then compared to the perceived physical activity level data obtained from EMA. Spearman's correlation was used to determine the correlation between the perceived physical activity level and all the calculated activity level values. The

correlation analysis was performed using Python and several data processing and analysis libraries: Pandas, SciPy, and NumPy.

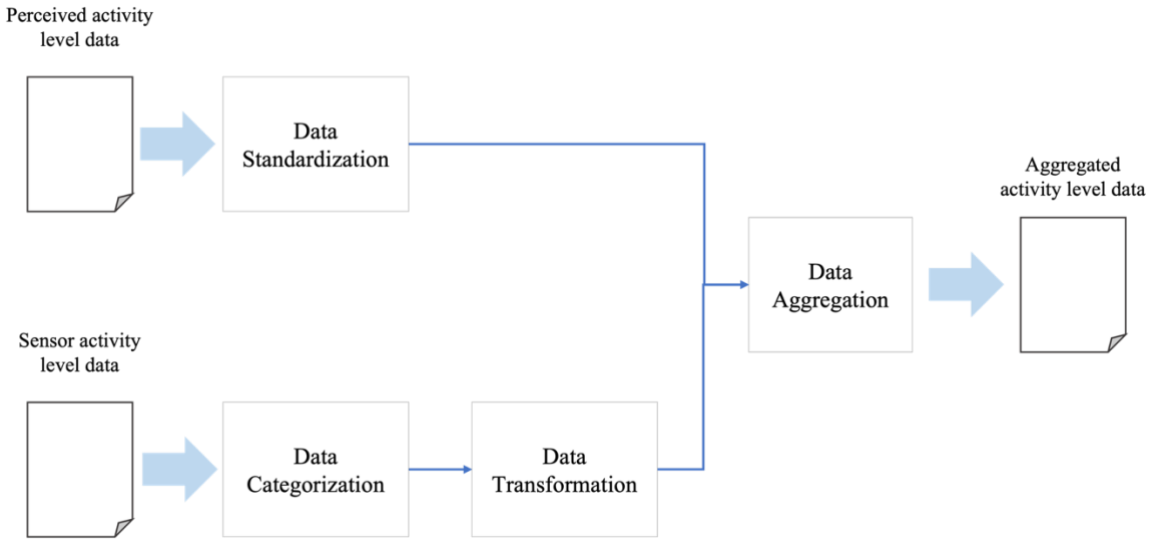


**Figure 26. Diagram of the integration process of the activity level data.**

## 5.3 Results

### 5.3.1 Data Collection

Sensor data and EMA data were collected from a total of 8 participants. Out of the 12 initial participants, 4 did not wear the kinematics sensors, and hence, the sensor data from these participants were not available for analysis. As a result, data from only 8 participants were used for the analysis of the sensor data.

It was expected that complete sensor data would be collected from the first full day of assessment until the last day of assessment, totaling six days. Since most participants started their seven-day assessment period in the late afternoon or evening, the first day of the assessment was excluded. The first full day of the assessment would be day 2. However, due to the sensors' battery performance, some sensors did not record complete six-day data. The average completeness[2] of the collected sensor data was 76.04% (SD = 15.71%).

For the EMA data, the participants were expected to report their perceived activity level twice each day: morning activity level in the afternoon EMA and afternoon activity level in the evening EMA. It was anticipated that 12 activity level reports would be collected from each participant during the six complete assessment days to align with the sensor data. The completeness of the reported EMA data appeared to be high, with an average completeness[3] of 92.71% (SD = 8.26%). The details of the completeness of the data collected can be seen in Table 15.

**Table 15. Completeness of Collected Data**

| | **Average Amount Collected** | **Average Completeness** |
|---|---|---|
| Sensor Data | 4.56 days (SD = 0.94) | 76.04% (SD = 15.71) |
| EMA | 11.13 reports (SD = 0.99) | 92.71% (SD = 8.26) |

[2] 100% complete means six-day sensor data was collected.

[3] 100% complete means 12 activity level reports were collected.

### 5.3.2 Data Processing

Data processing was performed for both the sensor data and the EMA data to prepare them for further analysis. Prior to processing of the sensor data, the code used for data processing was validated by comparing it with ActiLife (https://actigraphcorp.com/actilife/), a popular physical activity analysis software from ActiGraph. Actigraph data collected from sensors placed on the waist and sensor data from Lifeware sensors used in the study were used to compare the processing results. The data appeared to be highly comparable, with minimal differences observed, except for sedentary level counts. The difference in sedentary level counts was attributed to the fact that ActiLife does not calculate activity when the sensors are in sleep mode, whereas this study data processing algorithm accounted for all time points, including those when the sensors were inactive. This data processing algorithm recognized the activity level in these time points as sedentary. This led to a higher number of sedentary level counts in the algorithm's output. The result of the comparison can be seen in Table 16.

**Table 16. Results Comparison between the Data Processing Algorithm used in this Study with ActiLife**

| | ActiLife (Actigraph data placed on waist) | This Data Processing (Actigraph data placed on waist) | This Data Processing (Lifeware sensors data placed on L1) |
|---|---|---|---|
| Activity counts | 1358825 | 1360349 | 1223628 |
| Sedentary level counts | 4431 | 8685 | 8631 |
| Light level counts | 1169 | 1161 | 1152 |
| Moderate level counts | 263 | 263 | 215 |
| Vigorous level counts | 0 | 0 | 0 |
| Very vigorous level counts | 0 | 0 | 0 |

In this data analysis, sensor data were processed using the steps shown in Figure 24, and EMA data was processed using the steps shown in Figure 25. The distributions of the activity level taken from the processed EMA and sensor data can be seen in Figure 27 and Figure 28. The objective activity level data from the sensors and the perceived activity level data from the EMA were further processed and aggregated using the steps shown in Figure 26. The objective activity level data was transformed into eleven different metrics using the strategies discussed in Chapter 5.2.2. Data points that don't have activity level values from EMA and/or sensor data were excluded. A total of 73 data points were collected.

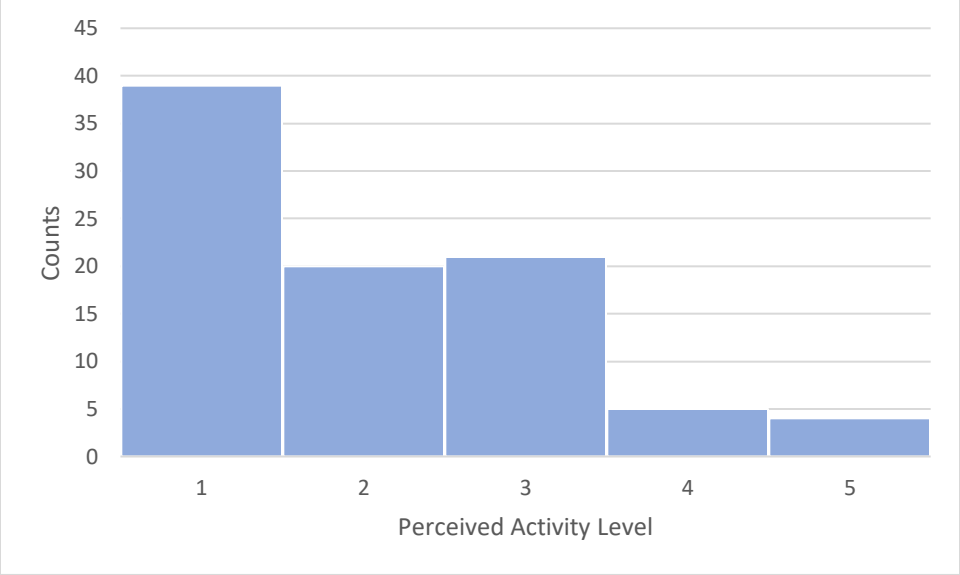


**Figure 27. Distribution of perceived activity level taken from EMA reports.**

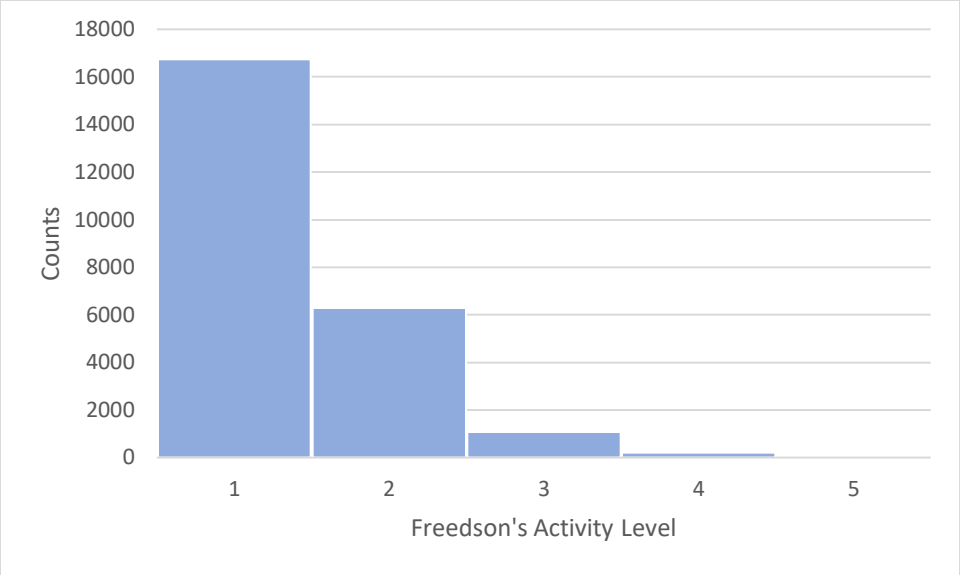


**Figure 28. Distribution of activity level derived from the sensor data.**

### 5.3.3 Correlation Analysis

Correlation analyses were performed to examine the relationship between the perceived morning/afternoon activity level and each of the eleven derived activity level representation from the sensor data. Further analyses to explore the relationship more deeply were also conducted. Relationships between each pair of variables in each time type (morning/afternoon) and for each participant were calculated and observed. In this analysis, the following cutpoints for correlation strength were used:

- Strong correlation: $r \geq 0.5$ or $r \leq -0.5$
- Moderate correlation: $0.3 \leq r < 0.5$ or $-0.5 < r \leq -0.3$
- Weak correlation: $-0.3 < r < -0.3$

#### 5.3.3.1 Perceived Activity Level vs Average of 60s-epoch Activity Level

The correlation between the perceived activity level and the average of 60s-epoch activity level was investigated. The average of 60s-epoch activity level refers to the average activity level that is calculated from the sensor data in each minute. The Spearman's correlation coefficient between the two variables was found to be 0.177 ($p = 0.134$), indicating a weak positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.767 ($p = 0.044$) and 0.653 ($p = 0.021$). Among the remaining participants, two had a moderate positive correlation and four had a negative correlation. The correlation coefficient calculation results for all participants are presented in Table 17.

**Table 17. Correlation Coefficient for Analysis 1**

**Perceived Activity Level vs Average of 60s-epoch Activity Level**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.177 | 0.134 |
| Participant A | 0.767 | 0.044 |
| Participant B | 0.378 | 0.402 |
| Participant C | -0.587 | 0.126 |
| Participant D | 0.653 | 0.021 |
| Participant E | 0.429 | 0.188 |
| Participant F | -0.028 | 0.940 |
| Participant G | -0.577 | 0.134 |
| Participant H | -0.502 | 0.140 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.267 ($p = 0.105$), indicating a weak positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.949 ($p = 0.051$) and 0.765 ($p = 0.076$). Among the remaining participants, one had a weak positive correlation, four had negative correlation, and one cannot be calculated because of uniformity in their data.

For afternoon data, the correlation coefficient between the two variables was found to be 0.018 ($p = 0.919$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.500 ($p = 0.667$), 0.530 ($p = 0.280$), and 0.632 ($p = 0.252$). Among the remaining participants, one had a weak positive correlation, three had negative correlation, and one cannot be calculated because of uniformity in their data. The correlation coefficient calculation results for morning/afternoon data are presented in Table 18.

**Table 18. Correlation Coefficient for Analysis 1 – Morning & Afternoon**

**Perceived Activity Level vs Average of 60s-epoch Activity Level**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.267 | 0.105 | 0.018 | 0.919 |
| Participant A | 0.949 | 0.051 | 0.500 | 0.667 |
| Participant B | - | - | -0.316 | 0.684 |
| Participant C | -0.316 | 0.684 | -0.949 | 0.051 |
| Participant D | 0.765 | 0.076 | 0.530 | 0.280 |

| | | | | |
|---|---|---|---|---|
| Participant E | 0.131 | 0.805 | 0.632 | 0.252 |
| Participant F | -0.438 | 0.385 | 0.258 | 0.742 |
| Participant G | -0.775 | 0.225 | - | - |
| Participant H | -0.738 | 0.155 | -0.447 | 0.450 |

#### 5.3.3.2 Perceived Activity Level vs Most Frequent 60s-epoch Activity Level

The correlation between the perceived activity level and the most frequent 60s-epoch activity level was investigated. The most frequent 60s-epoch activity level refers to the minutely activity level that is most prevalent in the morning/afternoon timespan. The Spearman's correlation coefficient between the two variables was found to be 0.150 ($p = 0.206$), indicating a weak positive correlation between the two variables. No participants had a moderate or strong positive correlation between the two variables. The correlation coefficient calculation results for all participants are presented in Table 19.

**Table 19. Correlation Coefficient for Analysis 2**

**Perceived Activity Level vs Most Frequent 60s-epoch Activity Level**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.150 | 0.206 |
| Participant A | 0.214 | 0.645 |
| Participant B | - | - |
| Participant C | 0.000 | 1.000 |
| Participant D | -0.099 | 0.759 |
| Participant E | 0.291 | 0.385 |

| | | |
|---|---|---|
| Participant F | - | - |
| Participant G | - | - |
| Participant H | - | - |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.215 (p = 0.195), indicating a weak positive correlation between the two variables. No participants had a moderate or strong positive correlation between the two variables. For afternoon data, the correlation coefficient between the two variables was found to be 0.074 (p = 0.672), indicating a weak positive correlation between the two variables. No participants had a moderate or strong positive correlation between the two variables. The correlation coefficient calculation results for morning/afternoon data are presented in Table 20.

There were a lot of correlation analyses that cannot be calculated because of data uniformity. Using most frequent 60s-epoch activity level to represent activity level in morning/afternoon time resulted in uniformity in activity level value (mean = 1.07, SD = 0.25). Most participants' morning/afternoon activity levels were mostly represented as sedentary.

**Table 20. Correlation Coefficient for Analysis 2 – Morning & Afternoon**

**Perceived Activity Level vs Most Frequent 60s-epoch Activity Level**

| **Participant** | **Morning** | | **Afternoon** | |
|---|---|---|---|---|
| | **r** | **p-value** | **r** | **p-value** |
| All participants | 0.215 | 0.195 | 0.074 | 0.672 |
| Participant A | - | - | 0.000 | 1.000 |

| | | | | |
|---|---|---|---|---|
| Participant B | - | - | - | - |
| Participant C | 0.000 | 1.000 | - | - |
| Participant D | -0.270 | 0.605 | 0.000 | 1.000 |
| Participant E | - | - | 0.186 | 0.764 |
| Participant F | - | - | - | - |
| Participant G | - | - | - | - |
| Participant H | - | - | - | - |

#### 5.3.3.3 Perceived Activity Level vs Highest 60s-epoch Activity Level

The correlation between the perceived activity level and the highest 60s-epoch activity level was investigated. The highest 60s-epoch activity level refers to the highest minutely activity level in the morning/afternoon timespan. Since morning/afternoon timespan covers a wide range of hours, the patient might submit the most memorable perceived activity level in that timespan, not the average or overall activity level in the morning/afternoon timespan. This activity level representation was chosen to try to pick a metric that corresponds to that kind of perceived activity level.

The Spearman's correlation coefficient between the two variables was found to be 0.140 ($p = 0.238$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.642 ($p = 0.120$), 0.632 (0.127), and 0.558 ($p = 0.059$). Among the remaining participants, two had a weak positive correlation and one had a negative correlation. The correlation coefficient calculation results for all participants are presented in Table 21.

**Table 21. Correlation Coefficient for Analysis 3**

**Perceived Activity Level vs Highest 60s-epoch Activity Level**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.140 | 0.238 |
| Participant A | 0.642 | 0.120 |
| Participant B | 0.632 | 0.127 |
| Participant C | -0.363 | 0.377 |
| Participant D | 0.558 | 0.059 |
| Participant E | 0.271 | 0.421 |
| Participant F | 0.000 | 1.000 |
| Participant G | -0.143 | 0.736 |
| Participant H | -0.188 | 0.603 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.250 ($p = 0.130$), indicating a weak positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.816 ($p = 0.184$) and 0.636 ($p = 0.175$). Among the remaining participants, two had a weak positive correlation, four had negative correlation, and one cannot be calculated because of uniformity in their data.

For afternoon data, the correlation coefficient between the two variables was found to be -0.032 ($p = 0.855$), indicating a weak negative correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.778 ($p = 0.222$) and 0.559 ($p = 0.327$). Among the remaining participants, one had a moderate positive correlation

and two had negative correlation. The correlation coefficient calculation results for morning/afternoon data are presented in Table 22.

There were some correlation analyses that cannot be calculated because of data uniformity. Some participants had consistent highest 60s-epoch activity level that resulted in uniformity in activity level value (mean = 3.03, SD = 0.83). Most participants' morning/afternoon activity levels were mostly represented as moderate.

**Table 22. Correlation Coefficient for Analysis 3 – Morning & Afternoon**

**Perceived Activity Level vs Highest 60s-epoch Activity Level**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.250 | 0.130 | -0.032 | 0.855 |
| Participant A | 0.816 | 0.184 | - | - |
| Participant B | - | - | 0.778 | 0.222 |
| Participant C | 0.000 | 1.000 | -0.707 | 0.293 |
| Participant D | 0.636 | 0.175 | 0.366 | 0.476 |
| Participant E | 0.139 | 0.793 | 0.559 | 0.327 |
| Participant F | - | - | -0.333 | 0.667 |
| Participant G | - | - | - | - |
| Participant H | -0.354 | 0.559 | 0.000 | 1.000 |

#### 5.3.3.4 Perceived Activity Level vs Average Hourly Activity Counts

The correlation between the perceived activity level and the average hourly activity counts level was investigated. The average hourly activity counts level refers to the average of activity counts that is calculated from the sensor data in each hour. The Spearman's correlation coefficient between the two variables was found to be 0.233 ($p = 0.047$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.692 ($p = 0.085$), 0.697 ($p = 0.082$) and 0.724 ($p = 0.008$). Among the remaining participants, one had a moderate positive correlation, one had a weak positive correlation, and four had negative correlation. The correlation coefficient calculation results for all participants are presented in Table 23.

**Table 23. Correlation Coefficient for Analysis 4**

**Perceived Activity Level vs Average Hourly Activity Counts**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.233 | 0.047 |
| Participant A | 0.692 | 0.085 |
| Participant B | 0.697 | 0.082 |
| Participant C | -0.587 | 0.126 |
| Participant D | 0.724 | 0.008 |
| Participant E | 0.381 | 0.247 |
| Participant F | -0.110 | 0.762 |
| Participant G | -0.247 | 0.555 |
| Participant H | -0.237 | 0.510 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.302 ($p = 0.065$), indicating a moderate positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.632 ($p = 0.368$) and 0.765 ($p = 0.076$). Among the remaining participants, one had a weak positive correlation, four had negative correlation, and one cannot be calculated because of uniformity in their data.

For afternoon data, the correlation coefficient between the two variables was found to be 0.095 ($p = 0.588$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 1.000 ($p = 0.000$), 0.738 ($p = 0.262$) and 0.618 ($p = 0.191$). Among the remaining participants, one had a weak positive correlation, three had negative correlation, and one cannot be calculated because of uniformity in their data. The correlation coefficient calculation results for morning/afternoon data are presented in Table 24.

**Table 24. Correlation Coefficient for Analysis 4 – Morning & Afternoon**

**Perceived Activity Level vs Average Hourly Activity Counts**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.302 | 0.065 | 0.095 | 0.588 |
| Participant A | 0.632 | 0.368 | 1.000 | 0.000 |
| Participant B | - | - | 0.738 | 0.262 |
| Participant C | -0.316 | 0.684 | -0.949 | 0.051 |
| Participant D | 0.765 | 0.076 | 0.618 | 0.191 |

| | | | | |
|---|---|---|---|---|
| Participant E | 0.131 | 0.805 | 0.264 | 0.668 |
| Participant F | -0.278 | 0.594 | -0.258 | 0.742 |
| Participant G | -0.258 | 0.742 | - | - |
| Participant H | -0.316 | 0.604 | -0.224 | 0.718 |

#### 5.3.3.5 Perceived Activity Level vs Activity Level from Average of Hourly Activity Counts

The correlation between the perceived activity level and the activity level from the average of hourly activity counts was investigated. The activity level from the average of hourly activity counts refers to the activity level classification of the average of hourly activity counts using Freedson's cutpoint. The Spearman's correlation coefficient between the two variables was found to be 0.095 ($p = 0.423$), indicating a weak positive correlation between the two variables. One participant had a strong positive correlation between the two variables: 0.528 ($p = 0.077$). Among the remaining participants, two had a moderate positive correlation, one had weak positive correlation, and two had negative correlation. The correlation coefficient calculation results for all participants are presented in Table 25.

**Table 25. Correlation Coefficient for Analysis 5**

**Perceived Activity Level vs Activity Level from Average of Hourly Activity Counts**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.095 | 0.423 |
| Participant A | - | - |
| Participant B | 0.322 | 0.481 |

| | | |
|---|---|---|
| Participant C | -0.531 | 0.176 |
| Participant D | 0.528 | 0.077 |
| Participant E | 0.451 | 0.163 |
| Participant F | 0.264 | 0.462 |
| Participant G | - | - |
| Participant H | -0.504 | 0.137 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.200 ($p = 0.228$), indicating a weak positive correlation between the two variables. One participant had a strong positive correlation between the two variables: 0.746 ($p = 0.088$). Among the remaining participants, one had a moderate positive correlation and one negative correlation.

For afternoon data, the correlation coefficient between the two variables was found to be -0.083 ($p = 0.637$), indicating a negative correlation between the two variables. One participant had a strong positive correlation between the two variables: 0.559 ($p = 0.327$). Among the remaining participants, one had a moderate positive correlation, one had weak positive correlation, and one had negative correlation. The correlation coefficient calculation results for morning/afternoon data are presented in Table 26.

There were a lot of correlation analyses that cannot be calculated because of data uniformity. Using activity level from the average of hourly activity counts to represent activity level in morning/afternoon time resulted in uniformity in activity level value (mean = 1.79, SD = 0.47). Most participants' morning/afternoon activity levels were mostly represented as light.

**Table 26. Correlation Coefficient for Analysis 5 – Morning & Afternoon**

**Perceived Activity Level vs Activity Level from Average of Hourly Activity Counts**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.200 | 0.228 | -0.083 | 0.637 |
| Participant A | - | - | - | - |
| Participant B | - | - | 0.272 | 0.728 |
| Participant C | - | - | -0.816 | 0.184 |
| Participant D | 0.746 | 0.088 | - | - |
| Participant E | 0.316 | 0.541 | 0.559 | 0.327 |
| Participant F | - | - | 0.333 | 0.667 |
| Participant G | - | - | - | - |
| Participant H | -0.745 | 0.148 | - | - |

### 5.3.3.6 Perceived Activity Level vs Highest Hourly Activity Counts

The correlation between the perceived activity level and highest hourly activity counts was investigated. The highest hourly activity counts refer to the highest activity counts in a one-hour timespan. Since morning/afternoon timespan covers a wide range of hours, the patient might submit the most memorable perceived activity level in that timespan, not the average or overall activity level in the morning/afternoon timespan. Similar like Section 5.3.3.3, this activity level representation was chosen to try to pick a metric that correspond to that kind of perceived activity level.

The Spearman's correlation coefficient between the two variables was found to be 0.254 (p = 0.030), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.767 (p = 0.044), 0.697 (p = 0.082), and 0.738 (p = 0.006). Among the remaining participants, one had a moderate positive correlation and one had negative correlation. The correlation coefficient calculation results for all participants are presented in Table 27.

**Table 27. Correlation Coefficient for Analysis 6**

**Perceived Activity Level vs Highest Hourly Activity Counts**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.254 | 0.030 |
| Participant A | 0.767 | 0.044 |
| Participant B | 0.697 | 0.082 |
| Participant C | -0.651 | 0.080 |
| Participant D | 0.738 | 0.006 |
| Participant E | 0.429 | 0.188 |
| Participant F | 0.028 | 0.940 |
| Participant G | -0.412 | 0.310 |
| Participant H | -0.086 | 0.814 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.330 (p = 0.043), indicating a moderate positive correlation between the two variables. Two participants

had a strong positive correlation between the two variables: 0.632 (p = 0.368) and 0.647 (p = 0.165). Among the remaining participants, two had a weak positive correlation.

For afternoon data, the correlation coefficient between the two variables was found to be 0.112 (p = 0.522), indicating a weak positive correlation between the two variables.

Three participants had a strong positive correlation between the two variables: 1.000 (p = 0.000), 0.738 (p = 0.262) and 0.765 (p = 0.076). Among the remaining participants, two had a weak positive correlation. The correlation coefficient calculation results for morning/afternoon data are presented in Table 28.

**Table 28. Correlation Coefficient for Analysis 6 – Morning & Afternoon**

**Perceived Activity Level vs Highest Hourly Activity Counts**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.330 | 0.043 | 0.112 | 0.522 |
| Participant A | 0.632 | 0.368 | 1.000 | 0.000 |
| Participant B | - | - | 0.738 | 0.262 |
| Participant C | -0.316 | 0.684 | -0.949 | 0.051 |
| Participant D | 0.647 | 0.165 | 0.765 | 0.076 |
| Participant E | 0.131 | 0.805 | 0.264 | 0.668 |
| Participant F | 0.123 | 0.816 | -0.258 | 0.742 |
| Participant G | -0.258 | 0.742 | - | - |
| Participant H | -0.105 | 0.866 | 0.224 | 0.718 |

#### 5.3.3.7 Perceived Activity Level vs Activity Level from Highest Hourly Activity Counts

The correlation between the perceived activity level and the activity level from the highest hourly activity counts was investigated. The activity level from the highest hourly activity counts refers to the activity level classification of the highest hourly activity counts using Freedson's cutpoint. The Spearman's correlation coefficient between the two variables was found to be 0.164 ($p = 0.165$), indicating a weak positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.833 ($p = 0.020$) and 0.512 ($p = 0.089$). Among the remaining participants, one had a moderate positive correlation and one had negative correlation. The correlation coefficient calculation results for all participants are presented in Table 29.

**Table 29. Correlation Coefficient for Analysis 7**

**Perceived Activity Level vs Activity Level from Highest Hourly Activity Counts**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.164 | 0.165 |
| Participant A | - | - |
| Participant B | 0.833 | 0.020 |
| Participant C | -0.619 | 0.102 |
| Participant D | 0.512 | 0.089 |
| Participant E | 0.451 | 0.163 |
| Participant F | -0.345 | 0.329 |
| Participant G | - | - |
| Participant H | 0.063 | 0.863 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.261 (p = 0.113), indicating a weak positive correlation between the two variables. One participant had a strong positive correlation between the two variables: 0.652 (p = 0.161) and 0.765 (p = 0.076). Among the remaining participants, one had a moderate positive correlation and one had weak correlation.

For afternoon data, the correlation coefficient between the two variables was found to be 0.000 (p = 1.000), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.943 (p = 0.057) and 0.559 (p = 0.327). The correlation coefficient calculation results for morning/afternoon data are presented in Table 30.

There were several correlation analyses that cannot be calculated because of data uniformity. Using activity level from the highest hourly activity counts to represent activity level in morning/afternoon time resulted in uniformity in activity level value (mean = 1.97, SD = 0.58). Most participants' morning/afternoon activity levels were represented as light.

**Table 30. Correlation Coefficient for Analysis 7 – Morning & Afternoon**

**Perceived Activity Level vs Activity Level from Highest Hourly Activity Counts**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.261 | 0.113 | 0.000 | 1.000 |
| Participant A | - | - | - | - |
| Participant B | - | - | 0.943 | 0.057 |

| | | | | |
|---|---|---|---|---|
| Participant C | - | - | -0.816 | 0.184 |
| Participant D | 0.652 | 0.161 | - | - |
| Participant E | 0.316 | 0.541 | 0.559 | 0.327 |
| Participant F | -0.283 | 0.587 | -0.577 | 0.423 |
| Participant G | - | - | - | - |
| Participant H | 0.186 | 0.764 | - | - |

#### 5.3.3.8 Perceived Activity Level vs Weighted Sum of Average Hourly Activity Level Proportion and Ranking Value

The correlation between the perceived activity level and the weighted sum of average hourly activity level proportion and ranking value was investigated. The ranking value used to calculate the activity level representation is the same value used in the coding of the Freedson's activity level shown in Table 13. This activity level representation was used to make the proportion of the higher activity level become more sensitive and influential to the activity level representation.

The Spearman's correlation coefficient between the two variables was found to be 0.185 ($p = 0.117$), indicating a weak positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.767 ($p = 0.044$) and 0.653 ($p = 0.021$). Among the remaining participants, two had a moderate positive correlation. The correlation coefficient calculation results for all participants are presented in Table 31.

**Table 31. Correlation Coefficient for Analysis 8**

**Perceived Activity Level vs Weighted Sum of Average Hourly Activity Level Proportion and Ranking Value**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.185 | 0.117 |
| Participant A | 0.767 | 0.044 |
| Participant B | 0.378 | 0.402 |
| Participant C | -0.587 | 0.126 |
| Participant D | 0.653 | 0.021 |
| Participant E | 0.429 | 0.188 |
| Participant F | -0.028 | 0.940 |
| Participant G | -0.577 | 0.134 |
| Participant H | -0.502 | 0.140 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.282 ($p = 0.086$), indicating a weak positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.949 ($p = 0.051$) and 0.765 ($p = 0.076$). Among the remaining participants, one had a weak positive correlation.

For afternoon data, the correlation coefficient between the two variables was found to be 0.018 ($p = 0.919$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.500 ($p = 0.667$), 0.530 ($p = 0.280$) and 0.632 ($p = 0.252$). Among the remaining participants, one had a weak positive correlation. The correlation coefficient calculation results for morning/afternoon data are presented in Table 32.

**Table 32. Correlation Coefficient for Analysis 8 – Morning & Afternoon**

**Perceived Activity Level vs Weighted Sum of Average Hourly Activity Level Proportion and Ranking Value**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.282 | 0.086 | 0.018 | 0.919 |
| Participant A | 0.949 | 0.051 | 0.500 | 0.667 |
| Participant B | - | - | -0.316 | 0.684 |
| Participant C | -0.316 | 0.684 | -0.949 | 0.051 |
| Participant D | 0.765 | 0.076 | 0.530 | 0.280 |
| Participant E | 0.131 | 0.805 | 0.632 | 0.252 |
| Participant F | -0.438 | 0.385 | 0.258 | 0.742 |
| Participant G | -0.775 | 0.225 | - | - |
| Participant H | -0.738 | 0.155 | -0.447 | 0.450 |

### 5.3.3.9 Perceived Activity Level vs Weighted Sum of Highest Hourly Activity Level Proportion and Ranking Value

The correlation between the perceived activity level and the weighted sum of the highest hourly activity level proportion and ranking value was investigated. The ranking value used to calculate the activity level representation is the same value used in the coding of the Freedson's activity level shown in Table 13. This activity level representation was used to make the proportion of the higher activity level become more sensitive and influential to the activity level

representation and to try to represent most memorable perceived activity level in morning/afternoon timespan.

The Spearman's correlation coefficient between the two variables was found to be 0.260 ($p = 0.026$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.840 ($p = 0.018$), 0.553 ($p = 0.198$), and 0.704 ($p = 0.011$). Among the remaining participants, one had a moderate positive correlation and one had negative correlation. The correlation coefficient calculation results for all participants are presented in Table 33.

**Table 33. Correlation Coefficient for Analysis 9**

**Perceived Activity Level vs Weighted Sum of Highest Hourly Activity Level Proportion and Ranking Value**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.260 | 0.026 |
| Participant A | 0.840 | 0.018 |
| Participant B | 0.553 | 0.198 |
| Participant C | -0.664 | 0.073 |
| Participant D | 0.704 | 0.011 |
| Participant E | 0.499 | 0.118 |
| Participant F | -0.056 | 0.878 |
| Participant G | 0.083 | 0.845 |
| Participant H | -0.317 | 0.372 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.357 (p

= 0.028), indicating a moderate positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.833 (p = 0.167) and 0.647 (p = 0.165). Among the remaining participants, two had a moderate positive correlation and one had weak positive correlation.

For afternoon data, the correlation coefficient between the two variables was found to be 0.076 (p = 0.665), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 1.000 (p = 0.000), 0.738 (p = 0.262) and 0.765 (p = 0.076). Among the remaining participants, one had a weak positive correlation. The correlation coefficient calculation results for morning/afternoon data are presented in Table 34.

**Table 34. Correlation Coefficient for Analysis 9 – Morning & Afternoon**

**Perceived Activity Level vs Weighted Sum of Highest Hourly Activity Level Proportion and Ranking Value**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.357 | 0.028 | 0.076 | 0.665 |
| Participant A | 0.833 | 0.167 | 1.000 | 0.000 |
| Participant B | - | - | 0.738 | 0.262 |
| Participant C | -0.316 | 0.684 | -0.949 | 0.051 |
| Participant D | 0.647 | 0.165 | 0.765 | 0.076 |
| Participant E | 0.393 | 0.441 | 0.264 | 0.668 |
| Participant F | -0.295 | 0.570 | -0.258 | 0.742 |
| Participant G | 0.258 | 0.742 | - | - |

| | | | | |
|---|---|---|---|---|
| Participant H | -0.316 | 0.604 | -0.224 | 0.718 |

#### 5.3.3.10 Perceived Activity Level vs Weighted Sum of Average Hourly Activity Level Proportion and Custom Value

The correlation between the perceived activity level the weighted sum of average hourly activity level proportion and custom value was investigated. Similar like the approach in Section 5.3.3.8, this activity level representation was used to make the proportion of the higher activity level become even more sensitive and influential to the activity level representation. Since the proportion of activity level is not distributed evenly as shown in Table 35, a custom scoring was made to make higher activity level become more influential. The custom value here is calculated as the division of average proportion of activity level 1 with average proportion of each activity level. The custom score is shown in the last column in Table 35.

**Table 35. Average Proportion of each Activity Level**

**Perceived Activity Level vs Weighted Sum of Average Hourly Activity Level Proportion and Custom Value**

| Activity Level | Average Proportion (%) | SD | Avg Proportion Level 1 / Avg Proportion |
|---|---|---|---|
| 1 | 0.68469101 | 0.25638427 | 1 |
| 2 | 0.2630618 | 0.21406408 | 2.60277629 |
| 3 | 0.04194757 | 0.09081831 | 16.3225446 |
| 4 | 0.0090824 | 0.05939366 | 75.3865979 |
| 5 | 0.00121723 | 0.01335975 | 562.5 |

The Spearman's correlation coefficient between the two variables was found to be 0.232 ($p = 0.048$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.823 ($p = 0.023$), 0.697 ($p = 0.082$), and 0.770 ($p = 0.003$). Among the remaining participants, one had a moderate positive correlation. The correlation coefficient calculation results for all participants are presented in Table 36.

**Table 36. Correlation Coefficient for Analysis 10**

**Perceived Activity Level vs Weighted Sum of Average Hourly Activity Level Proportion and Custom Value**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.232 | 0.048 |
| Participant A | 0.823 | 0.023 |
| Participant B | 0.697 | 0.082 |
| Participant C | -0.511 | 0.196 |
| Participant D | 0.770 | 0.003 |
| Participant E | 0.313 | 0.349 |
| Participant F | -0.055 | 0.880 |
| Participant G | -0.577 | 0.134 |
| Participant H | -0.184 | 0.610 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.281 ($p = 0.087$), indicating a weak positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.949 ($p = 0.051$) and 0.853 ($p = 0.031$).

Among the remaining participants, one had a moderate positive correlation and one had weak positive correlation.

For afternoon data, the correlation coefficient between the two variables was found to be 0.122 ($p = 0.484$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 1.000 ($p = 0.000$), 0.738 ($p = 0.262$) and 0.765 ($p = 0.076$). Among the remaining participants, one had a weak positive correlation. The correlation coefficient calculation results for morning/afternoon data are presented in Table 37.

**Table 37. Correlation Coefficient for Analysis 10 – Morning & Afternoon**

**Perceived Activity Level vs Weighted Sum of Average Hourly Activity Level Proportion and Custom Value**

| Participant | Morning | | Afternoon | |
|---|---|---|---|---|
| | r | p-value | r | p-value |
| All participants | 0.281 | 0.087 | 0.122 | 0.484 |
| Participant A | 0.949 | 0.051 | 1.000 | 0.000 |
| Participant B | - | - | 0.738 | 0.262 |
| Participant C | 0.316 | 0.684 | -0.949 | 0.051 |
| Participant D | 0.853 | 0.031 | 0.765 | 0.076 |
| Participant E | 0.131 | 0.805 | 0.264 | 0.668 |
| Participant F | -0.438 | 0.385 | -0.258 | 0.742 |
| Participant G | -0.775 | 0.225 | - | - |
| Participant H | -0.105 | 0.866 | -0.224 | 0.718 |

#### 5.3.3.11 Perceived Activity Level vs Weighted Sum of Highest Hourly Activity Level Proportion and Custom Value

The correlation between the perceived activity level the weighted sum of the highest hourly activity level proportion and custom value was investigated. This activity level representation is similar to the previous representation, but highest hourly proportion was used instead of the average in hope to get the most representative metric to correspond to the most memorable perceived activity level in morning/afternoon timespan.

The Spearman's correlation coefficient between the two variables was found to be 0.245 ($p = 0.036$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 0.708 ($p = 0.075$), 0.553 ($p = 0.198$), and 0.770 ($p = 0.003$). Among the remaining participants, one had a moderate positive correlation. The correlation coefficient calculation results for all participants are presented in Table 38.

**Table 38. Correlation Coefficient for Analysis 11**

**Perceived Activity Level vs Weighted Sum of Highest Hourly Activity Level Proportion and Custom Value**

| Participant | r | p-value |
|---|---|---|
| All participants | 0.245 | 0.036 |
| Participant A | 0.708 | 0.075 |
| Participant B | 0.553 | 0.198 |
| Participant C | -0.651 | 0.080 |
| Participant D | 0.770 | 0.003 |
| Participant E | 0.430 | 0.187 |

| | | |
|---|---|---|
| Participant F | -0.056 | 0.878 |
| Participant G | -0.412 | 0.310 |
| Participant H | -0.191 | 0.597 |

Further analysis was performed by separating the data into morning/afternoon categories. For morning data, the correlation coefficient between the two variables was found to be 0.308 ($p = 0.060$), indicating a moderate positive correlation between the two variables. Two participants had a strong positive correlation between the two variables: 0.500 ($p = 0.500$) and 0.853 ($p = 0.031$). Among the remaining participants, one had a weak positive correlation.

For afternoon data, the correlation coefficient between the two variables was found to be 0.143 ($p = 0.414$), indicating a weak positive correlation between the two variables. Three participants had a strong positive correlation between the two variables: 1.000 ($p = 0.000$), 0.738 ($p = 0.262$) and 0.765 ($p = 0.076$). Among the remaining participants, two had a weak positive correlation. The correlation coefficient calculation results for morning/afternoon data are presented in Table 39.

**Table 39. Correlation Coefficient for Analysis 11 – Morning & Afternoon**

**Perceived Activity Level vs Weighted Sum of Highest Hourly Activity Level Proportion and Custom Value**

| **Participant** | **Morning** | | **Afternoon** | |
|---|---|---|---|---|
| | **r** | **p-value** | **r** | **p-value** |
| All participants | 0.308 | 0.060 | 0.143 | 0.414 |
| Participant A | 0.500 | 0.500 | 1.000 | 0.000 |
| Participant B | - | - | 0.738 | 0.262 |

| | | | | |
|---|---|---|---|---|
| Participant C | -0.316 | 0.684 | -0.949 | 0.051 |
| Participant D | 0.853 | 0.031 | 0.765 | 0.076 |
| Participant E | 0.131 | 0.805 | 0.264 | 0.668 |
| Participant F | -0.295 | 0.570 | -0.258 | 0.742 |
| Participant G | -0.258 | 0.742 | - | - |
| Participant H | -0.105 | 0.866 | 0.224 | 0.718 |

## 5.4 Discussion

### 5.4.1 Principal Results

The overall correlation between activity level from sensor data and perceived activity level from EMA was found to be weak, ranging from the score of 0.095 to 0.260 (mean = 0.194, SD = 0.054). Assuming higher correlation with the perceived activity level means more accurate representation of activity level, correlation with each activity level representation was examined. The following representations have the highest correlation with the perceived activity level:

- average hourly activity counts ($r = 0.233$, $p = 0.047$),
- highest hourly activity counts ($r = 0.254$, $p = 0.030$),
- weighted sum of the highest hourly activity level proportion and ranking value ($r = 0.260$, $p = 0.026$),
- weighted sum of average hourly activity level proportion and custom value ($r = 0.232$, $p = 0.048$), and

- weighted sum of the highest hourly activity level proportion and custom value ($r = 0.245$, $p = 0.036$).

These results suggest that these activity level representations may be more useful for accurately assessing activity level perception using sensor data and EMA. Average hourly activity counts, one of the most straightforward approaches, performed better than most activity level representations. Highest hourly activity counts also performed better. The attempt to match the most memorable, which is usually the most exerting, perceived activity level by using the highest hourly activity seemed to work better compared to using the highest 60s-epoch activity level.

Three representations that utilize weighted hourly activity proportion performed better compared to the other approaches. Using weighted sum of the highest hourly activity level proportion and ranking value seemed capable of appropriating the most memorable perceived activity level. Similarly, the approach of using weighted sum of the average hourly activity level proportion and ranking value didn't perform as well. This approach might not be able to accurately represent the higher activity level. Meanwhile, using weighted sum of the average hourly activity level proportion and custom value performed better. Compared to weighting the average hourly activity value proportion with ranking value, weighting it with custom value can represent higher activity level better. This calculation gives more sensitivity to the higher activity level. Another representation, using weighted sum of the highest hourly activity level proportion and custom value performed better, similar to using weighted sum of the highest hourly activity level proportion and ranking value.

Meanwhile, the following activity level representations performed poorly:

- most frequent 60s-epoch activity level (0.150, p = 0.206),
- highest 60s-epoch activity level (0.140, p = 0.238),
- activity level from the average of hourly activity counts (0.095, p = 0.423), and
- activity level from the highest hourly activity counts (0.164, p = 0.165).

All representations that performed poorly use activity level as the representative value. Using activity level that was supposed to be attached to 60s-epoch data might not be a good representative for wider timespan like morning/afternoon. These representations resulted in many uniformities in the calculated activity level representation, making correlation computation incalculable. This happened especially for most frequent 60s-epoch activity level, activity level from the average of hourly activity counts, and activity level from the highest hourly activity counts. These activity level representations were also not sensitive to the variability of activity levels, especially the higher activity level.

Even though the overall correlation was weak, correlation of activity level from sensor data and perceived activity level for three participants (A, B, and D) was found to be strong. Using the five most accurate activity level representations, the average score for the correlation was 0.716 (SD = 0.081), suggesting that these participants may have a better perception of their activity level. These participants also had higher scores for both morning (mean = 0.731, SD = 0.140) and afternoon (mean = 0.824, SD = 0.133) activity levels.

After separating the data into morning and afternoon categories, and still using the five most accurate activity level representations, the morning category showed better correlation (mean = 0.316, p = 0.029) compared to the afternoon category (mean = 0.109, p = 0.026). This

moderate correlation of the morning timespan suggests that activity level perception may be more accurate for morning activity.

### 5.4.2 Limitations

There are several limitations in this study. The first one is that most of the correlations calculated in the result were not statistically significant. Small number of participants and data collected in the assessment affected the statistical significance negatively. Larger data is needed to perform better correlation analysis.

There was also the limitation on the sensor battery. Some sensors were not able to stay on for the whole 7-day period of the at-home assessment. Sensor data completeness for the morning/afternoon in the 7-day period was 76.04%. As comparison, 92.71% EMA data was collected during the at-home assessment period. Addressing the battery issue can increase the data for analysis.

In this analysis Freedson's cutpoints that is meant for general adult population was used. These cutpoints were used because of its popularity for research's use and also because no specific cutpoints for chronic low back pain population was found. It can be seen in Figure 27 and Figure 28 that the distribution of perceived activity level and Freedson's activity counts were different. 68% of the activity counts were in Freedson's level 1, meanwhile 66.29% of perceived activity level were divided within the first two level: 43.82% in level 1 and 22.47% in level 2. Splitting the first level of Freedson's cutpoint into two categories might make the activity level more representative of what the patients perceive. Other than that, physical activity level cutpoints from Smuck et. al. (Smuck, Tomkins-Lane, Ith, Jarosz, & Kao, 2017) that target

musculoskeletal pain and mobility-limited populations can also be considered as more representative cutpoints to be used for further analysis.

There is also limitation of the perceived activity level representation in the EMA. In the EMA, the patients were asked their activity level in the morning and afternoon. The timespan might be too wide for the patients to generalize their activity level during those timespans. The definition of morning and afternoon might be different for each person as well. Some might define morning to start at 6 am while the other might define morning to start at 9 am. Several assumptions were made to develop suitable representations for the activity level derived from the sensor data, that was hoped to address the patients' representation of morning/afternoon activity level.

### 5.4.3 Conclusions

The results of this exploratory study showed that, with limitation of small data, there were no clear correlation between the activity level from sensor data with the perceived activity level from EMA. Using these results, it is still recommended to use both data to complement each other. Even though activity level from sensor data offers objective monitoring of the patients' activity level, reported perceived activity level from EMA can offer different perspectives. There might be behavioral or psychological information that can be examined from the difference in perceived activity level with the objective activity level.

To address the limitations in this study, a comprehensive correlation analysis study can be planned with more participants and larger data to get more accurate correlation analysis results. It will be interesting to compare the results of the same analysis in a bigger study to the results observed from this exploratory study.

Deeper analysis can be conducted by accommodating personalized timespan for the morning and afternoon timespan for each patient. Uniform timespan was used in this analysis, which are 7 am to 12 pm for the morning and 12 pm to 5 pm for the afternoon. The representation of morning/afternoon used in this study might be different to representation of morning/afternoon for each person. If this is to be developed, the patient's personalized morning/afternoon timespan can be integrated within the EMA. There is also a need to validate and standardize what it means by activity level and how to score it. Patients need to know how they can score their perception of their activity level.

Finally, there will be a lot of participants recruited in the bigger study, and more data will be collected. It will be interesting to plan a study to use the data to define activity level cutpoints that are specialized for chronic low back pain population. Furthermore, activity types and scores that are specialized for chronic low back pain can also be explored and defined.

# 6.0 Exploration for Future Works: Preliminary Work on a Personalized and Adaptive Treatment Component as Part of the mHealth System

## 6.1 Introduction

One of the ultimate goals of chronic low back pain phenotyping is to develop effective and precise treatments for patients. A personalized and adaptive mHealth system can be used to accommodate precision treatment. This chapter presents preliminary work on the development of a personalized and adaptive mHealth system for chronic low back pain treatment and management. This chapter discusses a potential design for the personalized and adaptive mHealth system. Similar to Chapter 5.0, data collected from in-clinic and at-home assessments were utilized for the exploration study in this chapter.

Given the extensive and rich data collected from the assessments, exploring the potential of applying a machine learning model to process the data becomes more relevant to support the development of the personalized and adaptive mHealth system. To build a machine learning model, a large dataset is needed. One of the most common machine learning approaches, supervised machine learning, learns from a labeled dataset. A labeled dataset is a dataset that has been annotated or tagged with one or more descriptive labels or categories. To develop the dataset, labeling process is needed. Labeling refers to the process of assigning one or more descriptive tags or categories to each data point in a dataset. This labeling process is an important step in building the dataset for a supervised machine learning model.

This chapter discusses the development of a dataset builder component that was integrated in the mHealth system developed in this study. This dataset builder component was

developed as part of the clinician portal component (Chapter 3.4.1.3). This development was conducted to contribute toward further works in developing a personalized and adaptive mHealth system for chronic low back pain treatment and management.

## 6.2 Personalized and Adaptive mHealth System for Chronic Low Back Pain

### 6.2.1 Personalized and Adaptive Treatment System

A personalized and adaptive system is a system that is designed to provide customized support and services to the patients. This type of framework is especially important in healthcare, where each patient has their own unique characteristics and has different experiences with their condition and treatment. In the context of chronic low back pain, a personalized and adaptive framework would aim to provide tailored treatment plans or interventions to patients based on several factors that are relevant to their chronic low back pain conditions. Chronic low back pain phenotyping can help explore and investigate those relevant factors for each subgroup of chronic low back pain.

To be personalized, the system should provide customized treatment or intervention based on the patient's unique needs, preferences, and low back pain characteristics. This data can be collected from various sources, such as in-clinic assessments, wearable sensors, and patient-reported outcomes. The system then utilizes this data to create a personalized treatment plan that addresses the patient's individual needs and goals.

The adaptive part of the system means the system is responsive to changes in the patient's condition, treatment effectiveness, and other factors that may impact their care. For example, if a

patient's pain levels increase or they experience side effects from their medication, the framework may automatically adjust their treatment plan or provide additional support to help manage their symptoms. The personalized and adaptive treatment is illustrated in Figure 29. Overall, a personalized and adaptive treatment system has the potential to improve patient outcomes and satisfaction by providing individualized and adaptive care that addresses the unique needs and preferences of each patient.

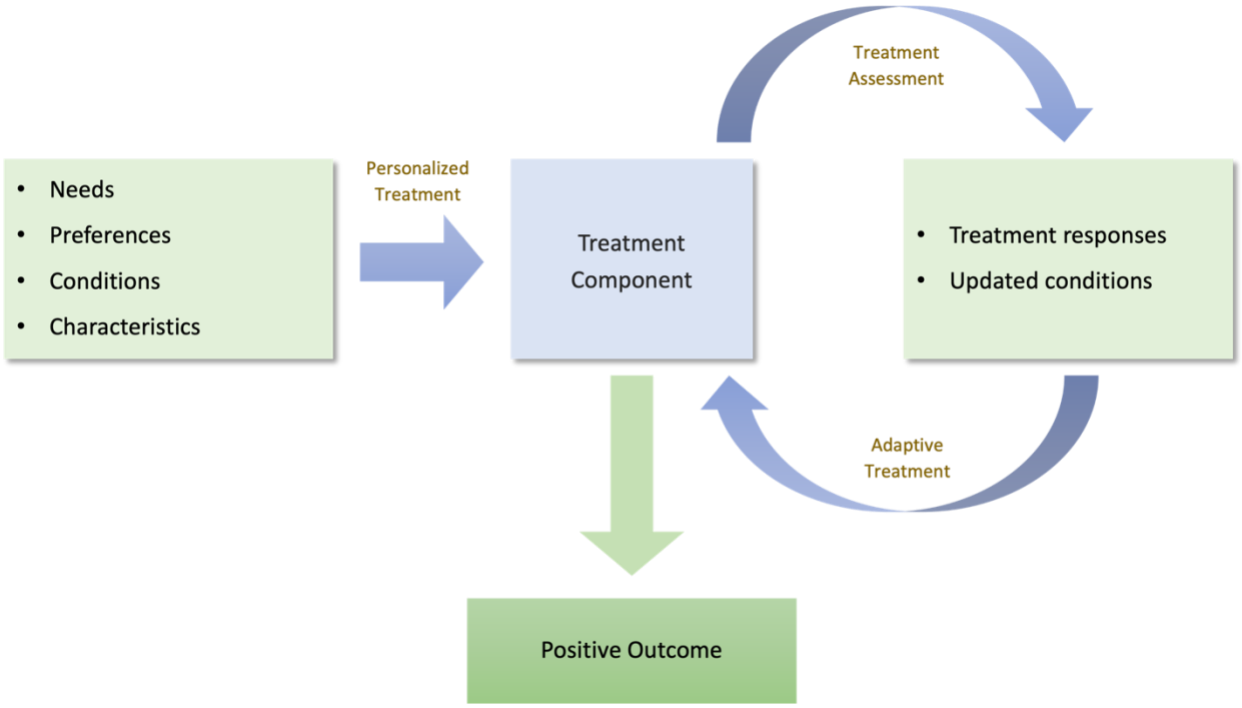


**Figure 29. Personalized and adaptive treatment component.**

### 6.2.2 Integrating Personalized and Adaptive Component to Assessment System

In line with the goal of developing a personalized and adaptive treatment mHealth system, the mHealth system for chronic low back pain assessment developed in this study

(Chapter 3.0) can be further developed to incorporate the personalized and adaptive treatment/intervention component. There are some approaches that can be used to integrate the treatment component and the assessment system. Lobelo and colleagues (Lobelo et al., 2016) proposed a framework for a mHealth-wearable sensors system that incorporates an intervention component. The framework was designed to support physical activity assessment, counseling, and intervention for reducing the risk of cardiovascular disease. The wearable sensors and mobile app in the framework collect physical activity data, which is transmitted to a digital ecosystem software platform for processing and standardization. Next, the processed data is sent to a clinical research center entity for analysis using a clinical outcome prediction algorithm. The resulting meaningful and summarized data is integrated into an EMR system that the healthcare team uses to make clinical decisions. Finally, a counseling and clinical intervention program is sent back to the patient via the app. This framework not only allows for the assessment and monitoring of a patient's condition but also enables the development and adjustment of personalized rehabilitation interventions. Lobelo's framework is illustrated in Figure 30.


**Figure 30. Labelo's framework for m-Health-wearables system with intervention.**

Another framework that can be used to incorporate personalized and adaptive treatment component into an mHealth system is Just-in-Time Adaptive Intervention (JITAI). JITAI is commonly used in mHealth interventions to provide personalized and timely support to

individuals in a convenient and accessible way. JITAI aims to provide the right support, at the right time, by adapting to the individual's internal and/or contextual variables changes (Nahum-Shani et al., 2015, 2018). The JITAI framework consists of four components: decision points, intervention options, tailoring variables, and decision rules (Nahum-Shani et al., 2018). Intervention options refer to the possible treatments or interventions that can be given at any decision point. Tailoring variables consist of individual-specific information used to determine when and what intervention to provide. Decision rules are a set of guidelines that specify which intervention or treatment to offer, to whom, and when. The decision rules component can be developed using machine learning techniques. Other design principal components of JITAI are distal outcome and proximal outcomes. Distal outcomes are the primary goals of the intervention and proximal outcomes are the short-term goals that can lead toward achieving the distal outcomes.

Linking JITAI with Lobelo's framework will create an adaptive intervention component that would be automatically implemented based on the decision rules without having to go through the healthcare team. Data process and decision rules analysis will be done inside the mHealth app. It can lead to faster and more seamless adjustment to the treatment program for the patients, which accommodates treatment delivery in real-time based on the individual's moment-to-moment needs and context. As patients engage with the treatment recommendations provided by the framework, the system would continue to collect assessment data and adjust its recommendations based on each patient's responses. Over time, this feedback loop would allow the framework to become increasingly personalized and adaptive, providing even more effective support to patients. Design of personalized and adaptive treatment mHealth system using the integration of Lobelo's framework and JITAI is illustrated in Figure 31.

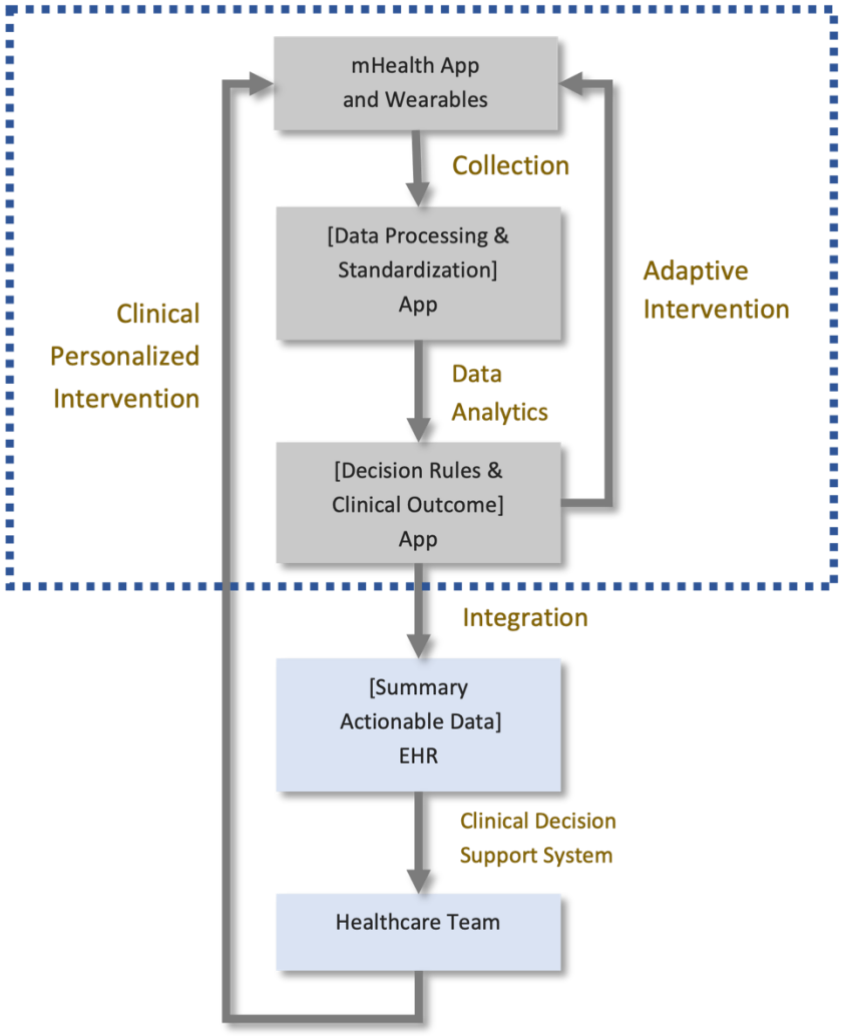


**Figure 31. Personalized and adaptive treatment mHealth system using Lobelo's and JITAI framework.**

### 6.2.3 Design of Machine Learning Component

Decision rules component plays a crucial role in any personalized and adaptive treatment system. Machine learning models can be employed to develop this component. In the present study, the data collected during the assessments can be used to train machine learning algorithms

and create a model that can serve as the decision rules component. This model can process the tailoring variables and recommend the most appropriate intervention or treatment option for each patient.

To contribute toward the development of a machine learning component that can serve as the decision rules component in the personalized and adaptive mHealth system for chronic low back pain treatment, this chapter explored two machine learning frameworks that can classify or predict motions as intermediate outcomes, which can be used in the bigger machine learning component. Using the data collected from the in-clinic and at-home assessments, a supervised and semi-supervised machine learning frameworks were designed.

#### 6.2.3.1 Supervised Machine Learning for Motion Classification

One approach explored for the machine learning component in the personalized and adaptive mHealth system for chronic low back pain treatment was a supervised machine learning framework for motion classification. This approach involves training a machine learning model using videos and sensor data collected during in-clinic assessments to recognize specific motions performed by patients.

The ability to accurately recognize and classify motions is crucial in developing effective treatment plans for chronic low back pain patients. Understanding how patients move and perform daily activities can help healthcare professionals identify and target specific movement patterns that may be contributing to pain and disability. By using supervised machine learning for motion classification, the mHealth system can provide personalized feedback and intervention based on the patient's unique movement patterns and limitations.

This machine learning framework requires processing kinematics data into labeled data. The labeled data is used in a classifier training to create a machine learning model that can

predict a motion or movement pattern from kinematics data. To help with the labeling process, visual data from video is used as a guide for labeling a time segment in the kinematics data. Instead of labeling the kinematics data directly, the labelers can label time segments in the video that later can be synced with time segments in the kinematics data. The design of this framework is illustrated in Figure 32.

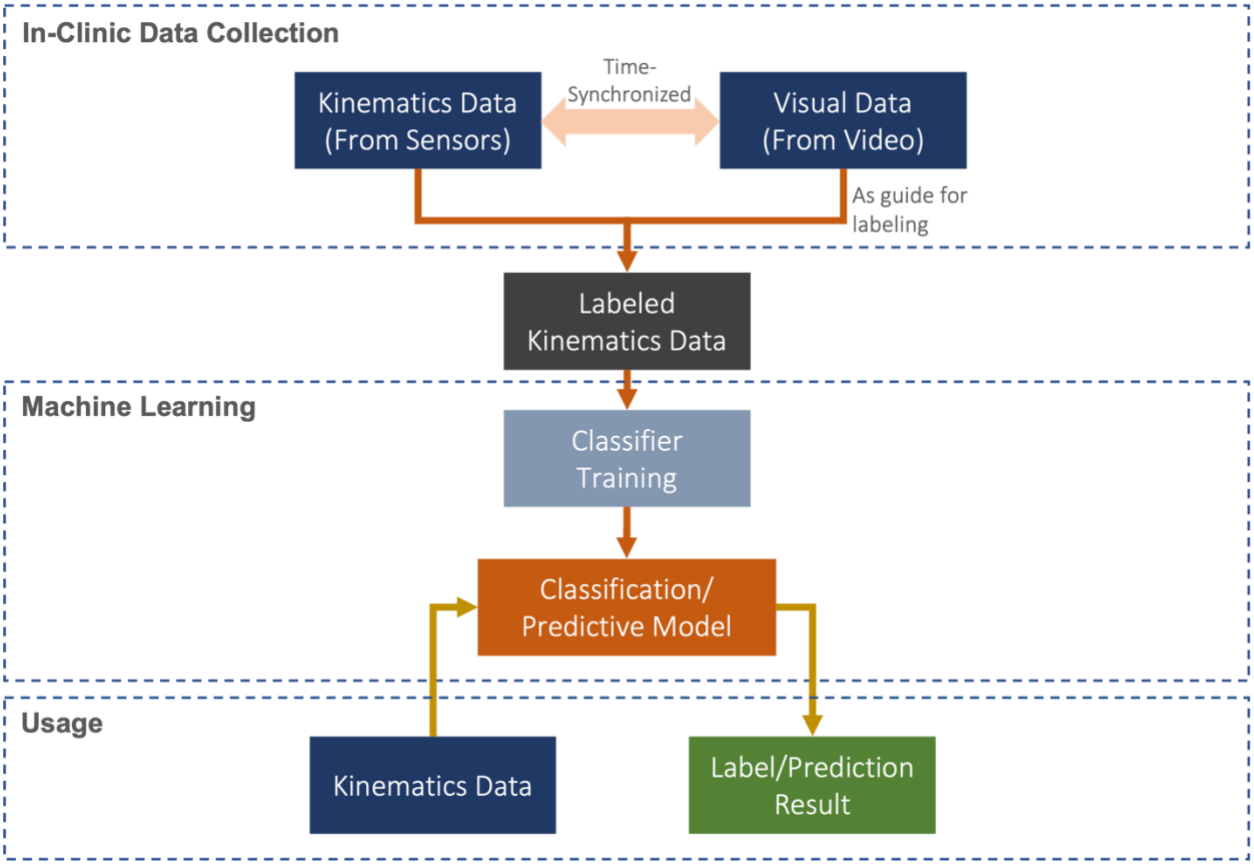


**Figure 32. Supervised machine learning framework for motion classification.**

### 6.2.3.2 Semi-Supervised Machine Learning for Motion Classification

Similar to the first framework, this semi-supervised approach involves training a machine learning model using videos and sensor data collected during in-clinic assessments to recognize

specific motions performed by patients. The difference is that a semi-supervised approach is used. One strength of semi-supervised learning compared to supervised learning is that it can leverage unlabeled data in addition to labeled data for training. This can be particularly advantageous in situations where labeling data is expensive or time-consuming.

In this study, kinematics data from the at-home assessment was collected. The difference between kinematics data from in-clinic assessment and at-home assessment is that there is no video available for the latter, which can be used to help visually label the data. In this approach, the at-home kinematics data can be leveraged to be used in the classifier training without requiring labelers to manually label it.

This machine learning framework requires processing kinematics data, with the help of visual data, into labeled data. This labeled data is used in a classifier training to create an intermediate machine learning model. This model is used to classify the unlabeled data from the at-home kinematics data. The classification result is used to label the at-home kinematics data. The classifier training uses the now-labeled at-home kinematics data to train the final model. The design of this framework is illustrated in Figure 33.

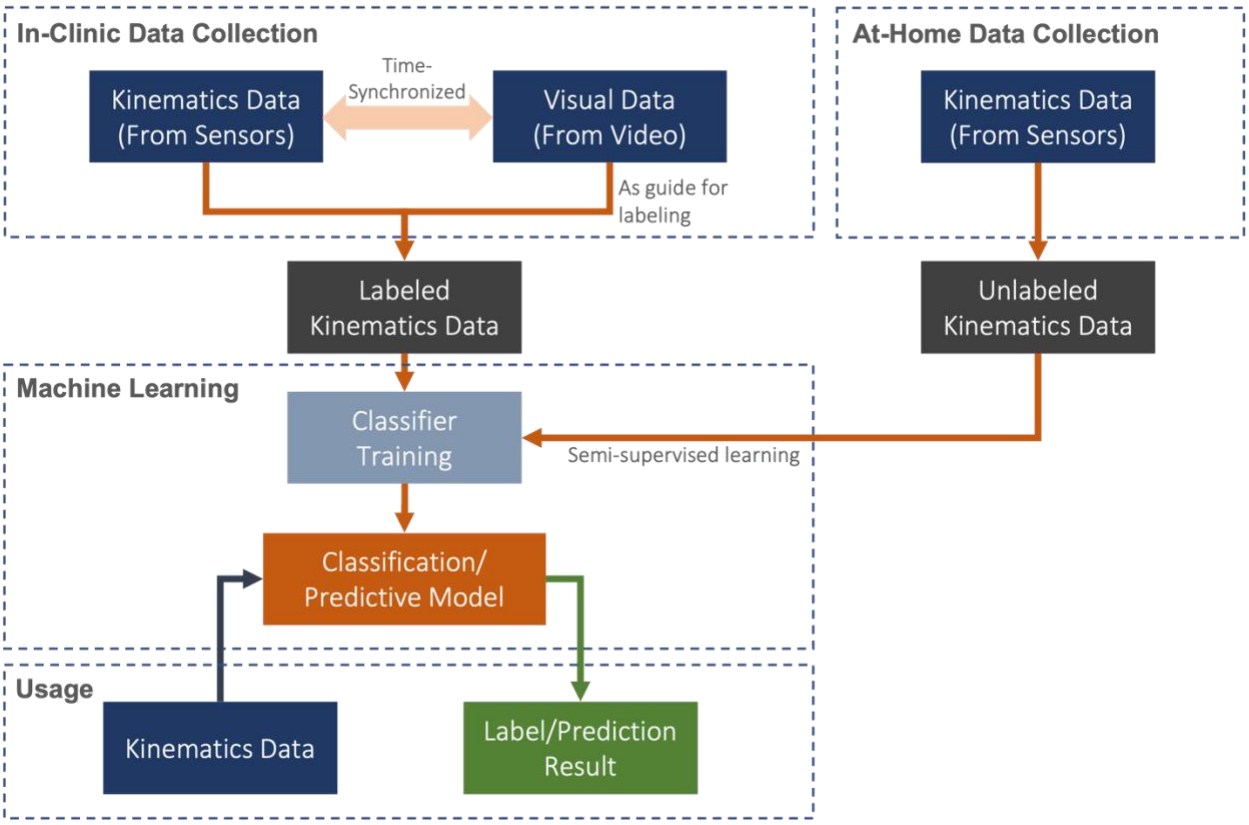


**Figure 33. Semi-supervised machine learning framework for motion classification.**

## 6.3 Development of the Dataset Builder Component

### 6.3.1 Requirement Analysis

In this study, a component to help the labeling process described in Section 6.2.3 was designed. This component is responsible for creating labeled dataset that will be used to train the machine learning model. To ensure that the dataset builder component meets the needs of the system, several requirements have been identified.

- Access: The dataset builder should be accessible to and only to the labelers that have approval to access the data. Ideally, it should be integrated into an existing component of the mHealth system, rather than requiring the development of a new component that would add complexity to the system.
- Video Retrieval and Display: The dataset builder component should be able to retrieve and display video data from the mHealth system, making it easy for users to identify and label motion segments.
- Time Segment Selection: The component should allow users to select time segments in the video, enabling them to label the specific motion segments of interest.
- Label Creation and Assignment: The dataset builder component should enable users to create and manage labels for the identified motion segments.
- User-Friendly Display: The component should display the data and labels assigned by the user in a clear and concise manner, making it easy for users to review and update their assigned labels.
- Data Management and Storage: The dataset builder component should be able to manage and store the labeled dataset, ensuring that the data is safe and easily accessible.

### 6.3.2 Design of the Dataset Builder Component

Based on the identified requirements, several services and features were planned as the building blocks of the dataset builder component. The dataset builder component was designed to have:

- login and access module,
- video and data retrieval service,
- video selection page,
- video player,
- time segment selector,
- label creation and selection module,
- labeled data section,
- data management service, and
- remote communication service.

Since the clinician portal already had some of the required features, such as a login module, data retrieval and management service, and remote communication service with the study database through the backend component, the dataset builder component was integrated as part of the clinician portal. This approach led to faster development and prevented unnecessary complexity in the overall mHealth system. The dataset builder was designed to utilize the clinician's portal login module, and authentication rules were updated to allow certified labelers to access the component. The video and data retrieval service was designed to make use of the existing data retrieval service in the clinician portal, aiming for efficient development and more

integrated connections between services. The video selection page was designed to enable labelers to choose which videos they wanted to label. The video player, time segment selector, label creation and selection module, and labeled data section were integrated into the video labeler module, which was designed as the main module for labeling videos. The data management service was designed to collect labeled data and send it to the remote database through the communication protocol provided by the backend. The design is illustrated in Figure 34.

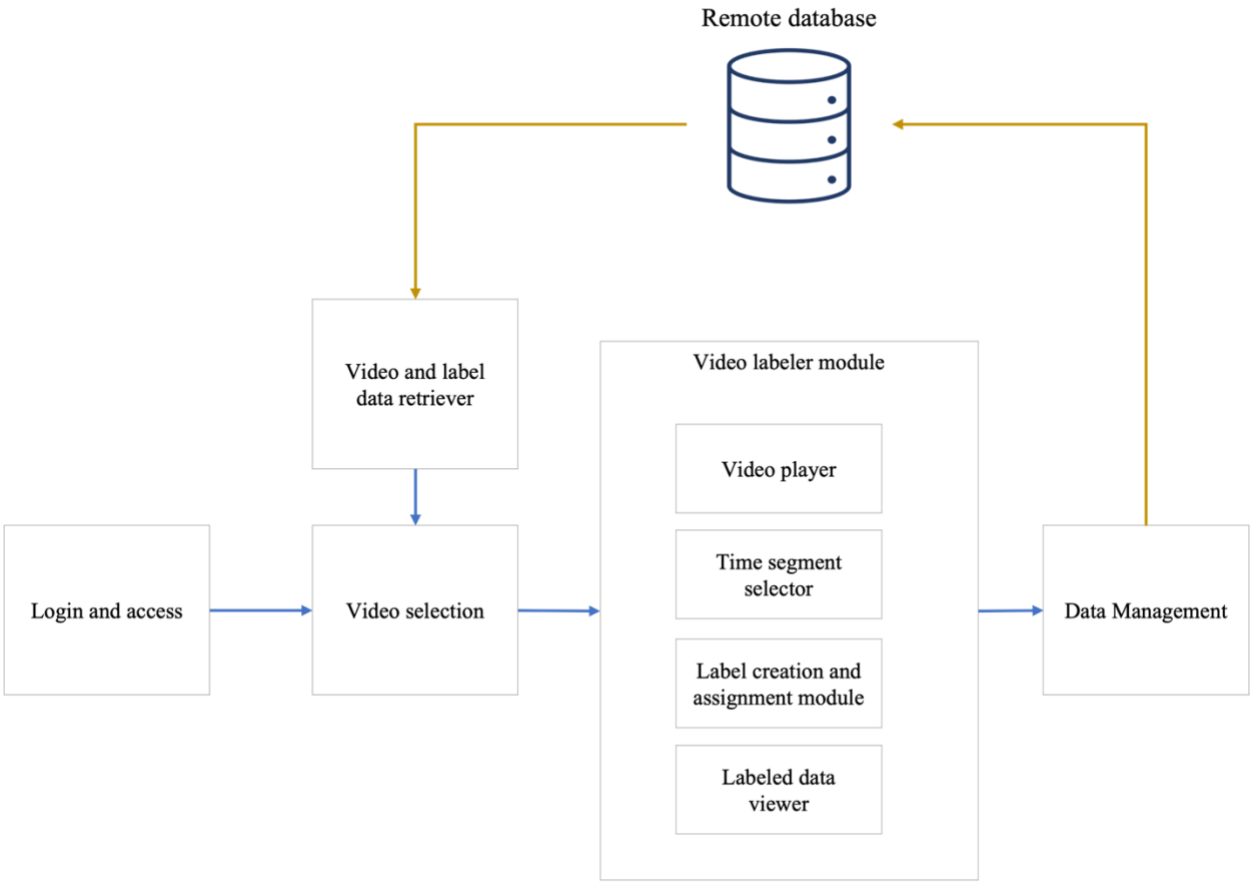


**Figure 34. Dataset builder component design.**

### 6.3.3 Development of the Dataset Builder Component

The dataset builder component, which creates labeled motion data for the machine learning component of the mHealth system, was developed within the clinician portal. A new access role was added in the clinician portal authorization rules to enable labelers to access the dataset builder component in the portal. This role allows certified labelers to access the component and start labeling the videos. Labelers can access the portal by using their account to login (Figure 35) to the clinician portal.

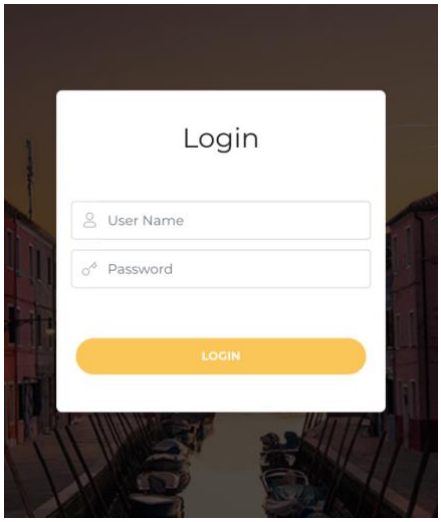


**Figure 35. Login page in the clinician portal.**

The video and data retrieval feature was also developed to allow the component to retrieve the videos and labels for the labeler. The existing data retrieval service in the clinician portal was adjusted to accommodate the video and data retrieval that is used by the dataset builder component. The video selection page was also developed to create an interface for selecting the videos to label. The test result page in the clinician portal was used as a template to

create the page, which allows labelers to easily navigate a familiar structure and choose the videos they want to label. The labelers need to select the patient/study participant before selecting the test video they want to label. The participant and video selection pages are shown in Figure 36.

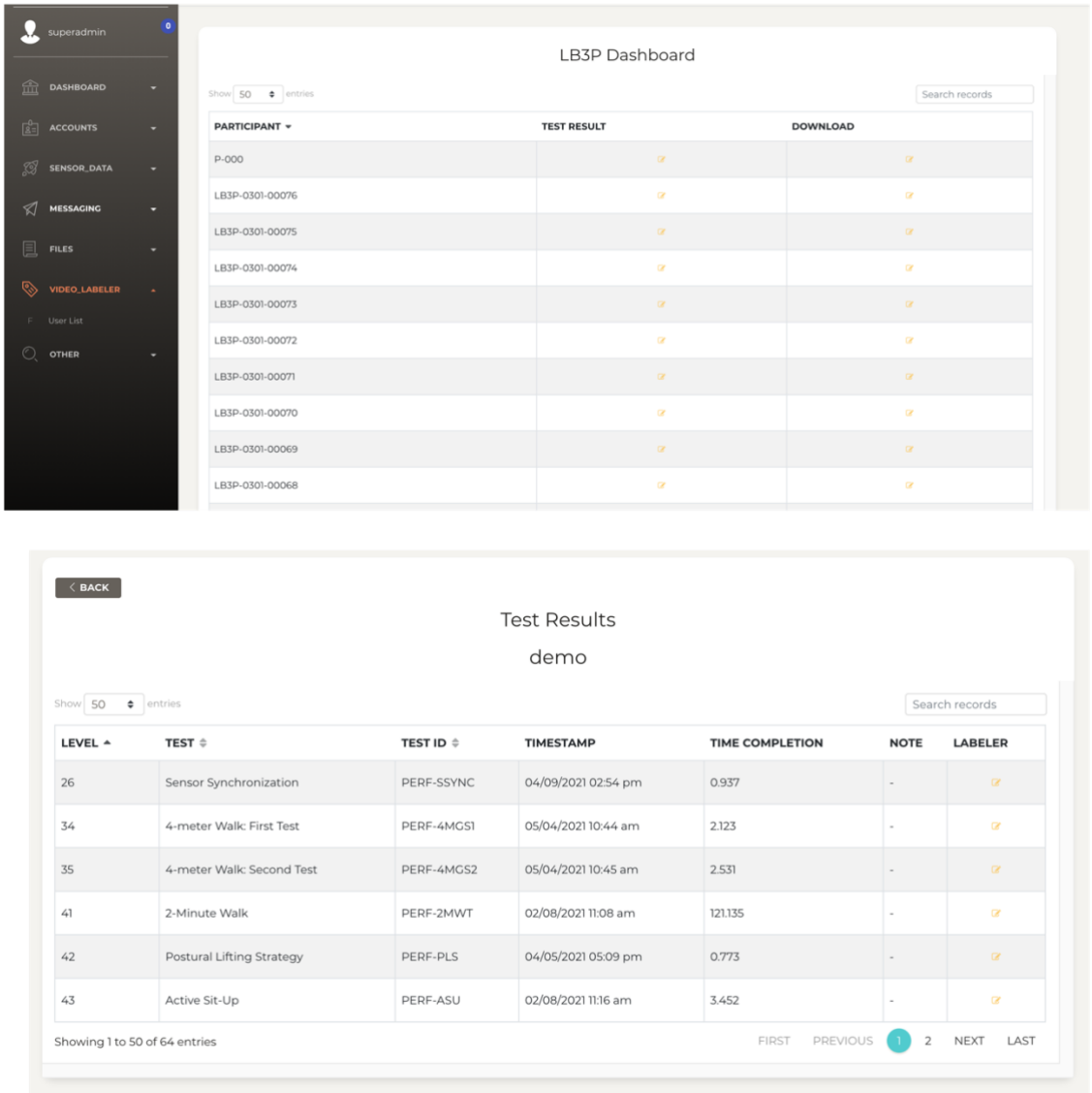


**Figure 36. Video selection page for the dataset builder component.**

The video labeler module was also developed as the main module of the dataset builder component that will be used for the labeling process. This module consists of the video player, time segment selector, label creation and assignment module, and labeled data viewer, which were integrated to create a seamless labeling process. The video player displays the video and allows the labeler to play and pause it as needed. The time segment selector enables the labeler to select a single timepoint to label a time segment in the video. To simplify the process and to ensure that all time segments in the video are labeled, time segment selector only requires the labeler to select one timepoint for one time segment. Once a timepoint is selected, a time segment between the previous timepoint, or start time if this is the first timepoint selected, and the timepoint selected would be used as the time segment to be labeled.

In the label creation and assignment module, the labeler can select a label from a provided set of labels or create a new one if necessary. This module allows the labeler to assign a label to the time segment selected. Finally, the labeled data viewer section displays the labeled time segments, allowing the labeler to review and edit their assignments as needed. If a mistake is made or the label needs to be changed, the labeler can delete the labeled data from this section. The video labeler module is illustrated in Figure 37.

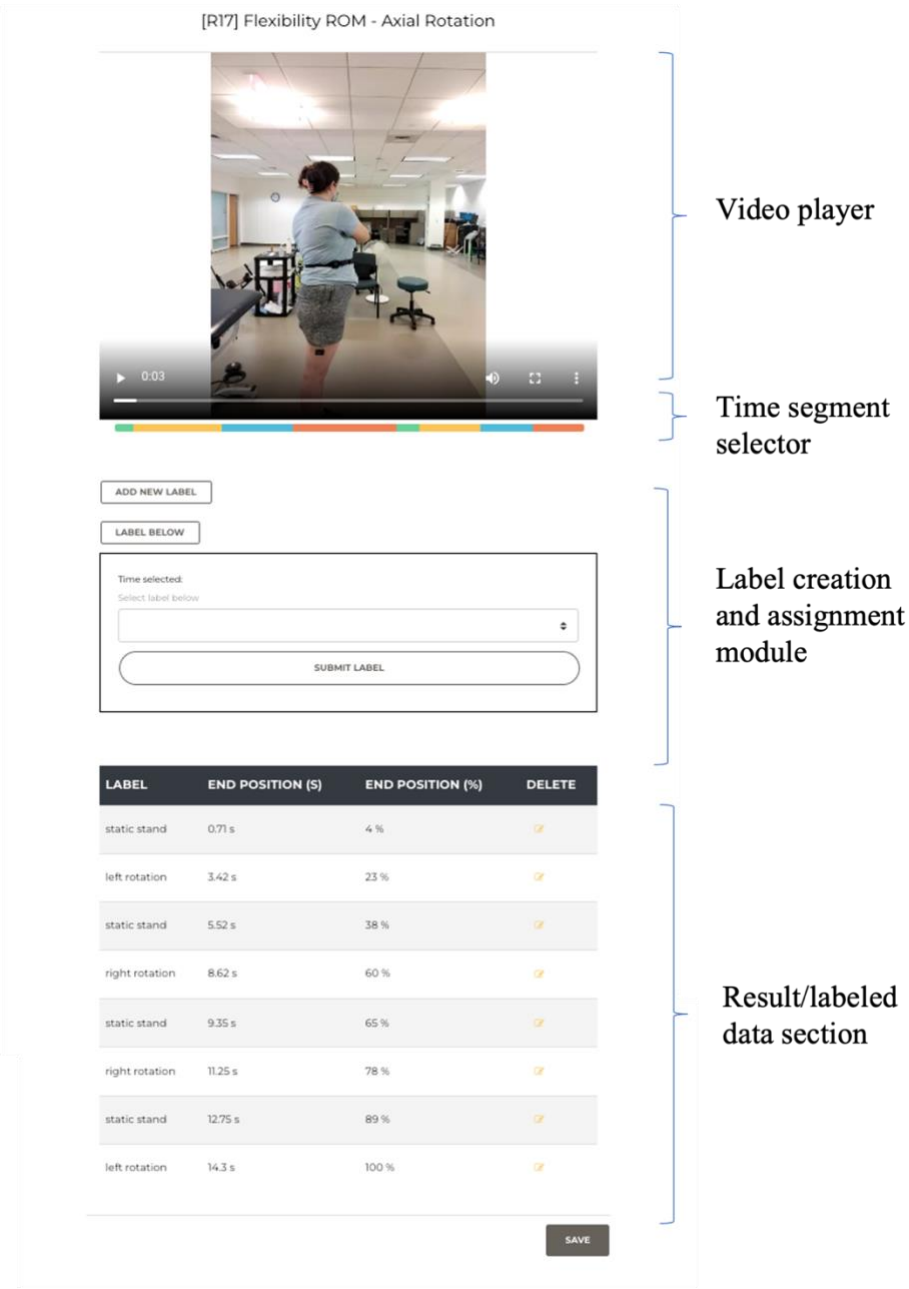


**Figure 37. Video labeler module.**

Finally, the data management service was developed to collect, process, and transmit the labeled data to the remote database. A new table was created in the database and new APIs were developed in the backend component to facilitate communication between the dataset builder component and the backend.

Furthermore, the dataset builder component has already been used by a number of researchers/labelers, and one labeler was able to create a dataset of motion-labeled time segments from 24 participants, with six functional performance test videos each. The labels used in this process includes motions such as neutral stand, backward bend, forward bend, left bend, left rotation, right bend, right rotation, and so on. This demonstrates that the dataset builder was successfully developed and can be used to create a dataset of motion-labeled data. A snippet of the labeled data can be seen in Appendix E.

Additionally, the dataset builder component was adapted to create a component for scoring a test or video session, not just the time segments in the session. This component was developed to score a functional performance test called the Postural Lifting Strategy test. Instead of allowing the user to select a time segment and assign a label, this component prompts the user to submit scores in several categories, such as spine neutral score, flexion dominance score, base of support score, aberrant movement score, and box proximity to body score. This test/video scoring component can also contribute to the development of a dataset for machine learning use. The screenshot of this component is shown in Figure 38.

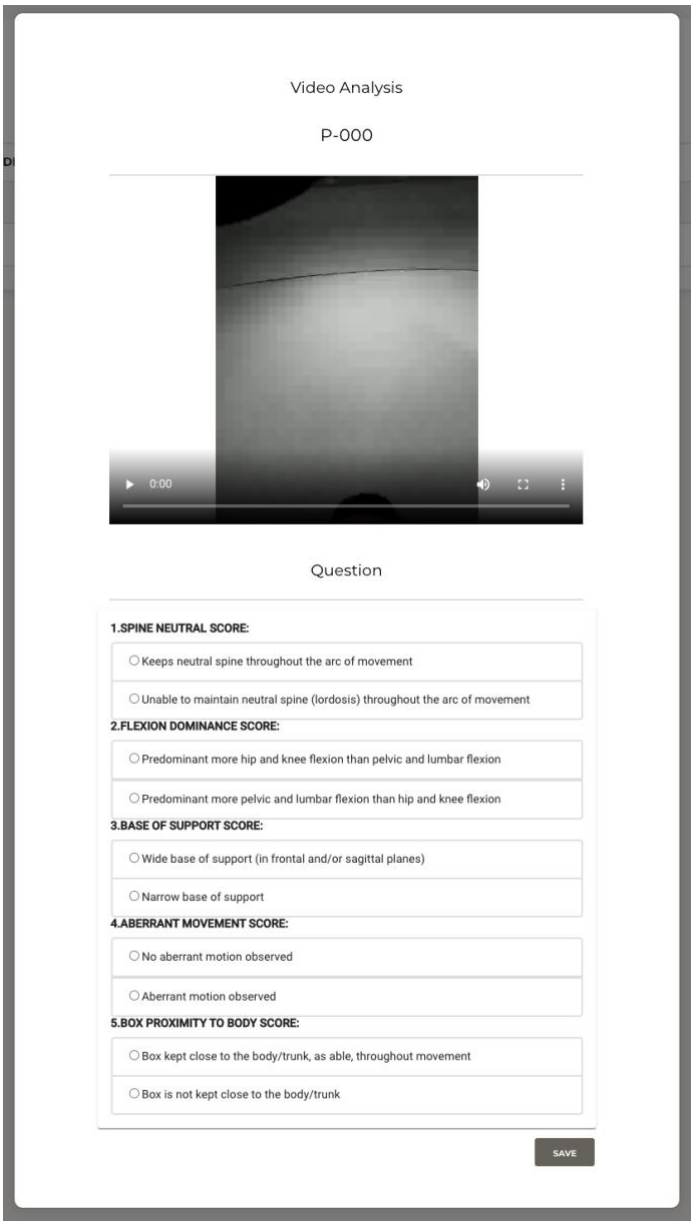


**Figure 38. Test/video scoring component that adapted the dataset builder component.**

## 6.4 Discussion

### 6.4.1 Principal Results

The principal results of this study include the design and implementation of several machine learning frameworks that are specifically tailored to the needs of the mHealth system developed in this study. A supervised and a semi-supervised machine learning frameworks were designed to be able to recognize motions from kinematics data. These machine learning frameworks require a set of labeled data. The machine learning frameworks were designed to make use of the availability of visual data. The frameworks use the visual data to help with the labeling, making the process easier for the labelers.

Additionally, the dataset builder component was successfully developed to facilitate the labeling of time segments in videos, allowing for the creation of labeled motion data that can be used to train machine learning models. This dataset builder component can be used in either supervised or semi-supervised machine learning frameworks that have been designed. The development of the dataset builder component and its successful use in labeling multiple videos is an important step towards the development of personalized and adaptive treatments for chronic low back patients.

The dataset builder component can also be adapted to other uses. By the time of the writing, a new component was developed using dataset builder component as the base. That component was created to enable scoring of functional performance tests by utilizing visual data from videos, similar to the dataset builder component.

### 6.4.2 Limitations

While the current study has made significant progress towards developing a machine learning model for motion classification, there are still several limitations to be addressed. Firstly, the machine learning component itself still requires further development. There is a need to define the approach to be used for machine learning, such as selecting the appropriate algorithm, feature selection, and optimizing the model. Further development and testing are required to determine the most effective approach to be used in this context.

Additionally, while the dataset builder component has been successfully developed, it is still under continuous development to implement integration with other data sources, such as sensor data. The dataset should contain segments of kinematics data with their assigned label. As of now, the dataset contains time segments with their assigned label. Synchronization between the time segment and the time in the kinematics data still needs to be done. A process to replace the time segments with the kinematics data segments with the help of the synchronized time information still needs to be developed.

Another limitation of the current study is that usability evaluation has not been performed yet. Usability study needs to be conducted to examine the usability of the dataset builder component and also to find if there are any usability issues. A usability study can provide valuable feedback on how to improve the user interface and overall usability of the system.

### 6.4.3 Conclusions

This study shows promises and potentials of utilizing machine learning toward the development of personalized and adaptive treatments for chronic low back pain patients. Preliminary steps were done in this study in the form of machine learning frameworks and a dataset builder component.

Further work needs to be done to further develop the dataset builder. One of the most essential features to work on is processing the labeled time segment data into labeled kinematics segment data. This can be achieved by making use of the timestamp information in both data. Another work that can be done to continue this study is to conduct usability evaluation of this dataset builder component. The component should have high usability and any usability issues should be identified and addressed. Also, further design analysis should be done to design the overall machine learning component. Feasibility analysis should be conducted to ensure that the machine learning component can be developed, and that any necessary data is available.

In conclusion, this study presented the development of a dataset builder component as part of the machine learning frameworks that was also designed in this study. Although, further works still need to be done, the development done in this study contributes to the potential development of the personalized and adaptive treatment system for chronic low back pain patients.

# 7.0 Summary and Discussion

## 7.1 Results Summary

To discuss the results of the studies in this dissertation, the specific aims of dissertation are revisited.

### 7.1.1 Specific Aim 1: To design and develop an mHealth system for comprehensive chronic low back pain assessment.

The development of an mHealth system for chronic low back pain assessment, which consists of an in-clinic app, at-home app, clinician portal, and backend and database, was conducted. The mHealth system was able to handle extensive and rich data from both the structured in-clinic assessment as well as the assessment in the patients' daily life settings.

The in-clinic app was able to help the physical therapist do the in-clinic assessment. The in-clinic app provides the physical therapists with the necessary assessment resources in one place. The use of the in-clinic app streamlines the assessment process and minimizes the effort and burden of the physical therapists.

The at-home app was able to accommodate the patients to report their EMA. This feature enhances the accuracy and reliability of the data while reducing the potential for recall bias. The at-home app also provides services and modules that can help the patients with their at-home assessment.

The clinician portal was also able to be utilized by study coordinators and researchers to perform data management tasks. The remote database was able to store the extensive and rich assessment data while the backend provided services and supports for data management and communication between the other components.

Overall, the mHealth system streamlines the assessment process and make the assessment more effective and efficient. All components were well integrated and connected, enabling seamless and automatic data integration from both in-clinic and at-home assessments.

### 7.1.2 Specific Aim 2: To evaluate the usability of the mHealth system.

Physical therapists found the in-clinic app helpful for conducting assessments, and as of the time of writing, over 500 patients have been assessed using it. Two usability evaluations were conducted to assess the usability of the in-clinic app. After each evaluation, several updates and revisions were made to address identified usability issues and concerns. In the latest evaluation, the in-clinic app received a high usability score of 6.00 (SD=1.15).

Five usability evaluations from a total of 337 chronic low back pain patients were conducted to assess the usability of the at-home app. After the first and second usability evaluations, several updates and revisions were made to address the identified usability issues. No significant usability issues were found after the third, fourth, and fifth usability evaluations. The usability scores plateaued after the second usability evaluation. In the last usability evaluation, the at-home app received a high usability score of 6.24 (SD=1.37).

The results showed that both in-clinic and at-home app had high usability score. Both apps can be used by the target users for their intended purposes.

### 7.1.3 Specific Aim 3: To investigate and compare activity level from subjective patient-reported EMA and objective kinematics sensor data.

To investigate the correlation between activity level from sensor data and perceived activity level from EMA, several representations were developed to represent activity level from sensor data since the common 60s-epoch activity level data needs to be transformed to represent activity level in morning/afternoon timespan, the timespan that represents perceived activity level from EMA. Eleven activity level representations were developed in aim to best represent the activity level in morning/afternoon timespan.

The overall correlation between activity level from sensor data and perceived activity level from EMA was found to be weak, ranging from the score of 0.095 to 0.260 (mean = 0.194, SD = 0.054). Five activity representations were selected as they scored highest correlation with the perceived activity level. Those activity representations are average hourly activity counts, highest hourly activity counts, weighted sum of the highest hourly activity level proportion and ranking value, weighted sum of average hourly activity level proportion and custom value, and weighted sum of the highest hourly activity level proportion and custom value ($r = 0.245$, $p = 0.036$). The use of weighted activity level proportion gives more sensitivity to the higher activity level.

Meanwhile, four activity level representations performed poorly, which are most frequent 60s-epoch activity level, highest 60s-epoch activity level, activity level from the average of hourly activity counts, and activity level from the highest hourly activity counts. Using activity level that was supposed to be attached to 60s-epoch data might not be a good representative for wider timespan like morning/afternoon. These activity level representations were also not sensitive to the variability of activity levels, especially the higher activity level.

Even though the overall correlation was weak, correlation of activity level from sensor data and perceived activity level for three participants (A, B, and D) was found to be strong. Using the five most accurate activity level representations, the average score for the correlation was 0.716 (SD = 0.081), suggesting that these participants may have a better perception of their activity level.

After separating the data into morning and afternoon categories, and still using the five most accurate activity level representations, the morning category showed better correlation (mean = 0.316, p = 0.029) compared to the afternoon category (r = 0.10, p = 0.026). This moderate correlation of the morning timespan suggests that activity level perception may be more accurate for morning activity.

### 7.1.4 Specific Aim 4: To design and develop dataset collection tools as preliminary work toward a personalized and adaptive intervention component in the mHealth system.

The results of this study include the design of machine learning frameworks tailored to the mHealth system's needs, including supervised and semi-supervised frameworks to recognize motions from kinematics data. These frameworks use visual data to facilitate labeling and make the process easier for labelers. The study also successfully developed the dataset builder component to label time segments in videos and create labeled motion data for training machine learning models. This component can be used in either supervised or semi-supervised frameworks and is an important step towards personalized and adaptive treatments for chronic low back pain patients. The dataset builder component can also be adapted for other uses, and a new component was developed using it to score functional performance tests with visual data from videos.

## 7.2 Limitations

There are several limitations identified for this study. First, for the development of the mHealth app, access to use communication protocol of the kinematics sensors was not made available. The mHealth system was developed without direct integration with the kinematics sensors, which was different from the mHealth framework that was designed.

For the usability evaluations, even though both in-clinic and at-home scored a high usability score, it would be beneficial to conduct direct interview with the chronic low back pain patients who use the at-home app. The usability evaluations embedded in the app were used instead to accommodate for rapid development of the app. Also, the usability of clinician portal needs to be evaluated. It was not evaluated in this study because it was designed for internal use at the time of development.

For the comparison of objective and subjective activity level, several limitations are identified. The first one is that most of the correlations calculated in the result were not statistically significant. There was also the limitation on the sensor battery that prevented complete collection of the sensor data. And, to classify the activity level, Freedson's cutpoints that is meant for general adult population were used, instead of cutpoints that made specifically for chronic low back pain population. There is also limitation of the perceived activity level representation in the EMA. Deeper analysis should be done to standardize the perception of activity level and the timespan it was asked.

## 7.3 Conclusions

The mHealth system for chronic low back pain was developed and can be used to assess chronic low back pain patients in clinical settings or daily life settings. The usability scores of the in-clinic and at-home app were high. Any usability issues identified during the usability evaluation were addressed. Further development of the system can be done in multiple directions. As of now, the mHealth system is also being used in another study that requires physical and functional performance test assessments. The mHealth system can be adjusted to collect more relevant data, such as patient reported outcome. Another development direction is toward the development of the personalized and adaptive treatment mHealth system for chronic low back pain.

Several preliminary works were done to contribute to the development in this direction. In this study, activity level from sensor data was compared to perceived activity level from EMA. Although the study found that the correlation was weak, this can open a new perspective in treating those two different data. For example, the results can be used to inform on which data is more suitable to be used in machine learning component to identify an outcome. For now, based on the results, the use of combination of the two data can give more information and context. Further analysis with more participants and larger data should be investigated to achieve clearer results. Some modifications can be made to the activity level component in the EMA to give the patients clearer understanding of what activity level means and how they should generalize their perception for a specific time period. It will also be interesting to plan a further study to use the data to define activity level cutpoints that are specialized for chronic low back pain population.

A dataset builder component was developed as part of the machine learning component that was also designed in this study. This development shows the promises and potentials of the machine learning component, including the dataset builder, in contributing to the development of a personalized and adaptive treatments for chronic low back pain patients. Although the dataset builder was able and was already used to generate labeled data, further development is still needed. One crucial area that requires attention is the conversion of labeled time segment data into labeled kinematics segment data, which can be achieved by utilizing timestamp information from both datasets. Additionally, a usability evaluation of the dataset builder component should be planned to identify and address any potential usability issues.

Overall, the development of the mHealth system for chronic low back pain was developed successfully. The mHealth system developed in this study plays a significant role in the chronic low back pain phenotyping research. The mHealth system is also usable and useful, assessed from its high usability score and its current usage in assessing more than 500 participants so far. Further works, as mentioned, are still needed to do to contribute toward the development of the personalized and adaptive treatment for chronic low back pain.

## Appendix A Usability Questionnaires Used for the In-Clinic App Evaluation

### Appendix A.1 mHealth App Usability Questionnaire (MAUQ) for Standalone Mobile Health App for Health Care Providers

The following MAUQ form (Zhou et al., 2019) was retrieved from https://ux.hari.pitt.edu/v2/portal/#/.

**mHealth App Usability Questionnaire (MAUQ)**

for Standalone Mobile Health App for Health Care Providers

| # | Statements | N/A | 1 2 3 4 5 6 7 |
|---|---|---|---|
| 1. | The app was easy to use. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 2. | It was easy for me to learn to use the app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 3. | The navigation was consistent when moving between screens. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 4. | The interface of the app allowed me to use all the functions (such as entering information, responding to reminders, viewing information) offered by the app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |

| | | | |
|---|---|---|---|
| 5. | Whenever I made a mistake using the app, I could recover easily and quickly. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 6. | I like the interface of the app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 7. | The information in the app was well organized, so I could easily find the information I needed. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 8. | The app adequately acknowledged and provided information to let me know the progress of my action. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 9. | I feel comfortable using this app in social settings. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 10. | The amount of time involved in using this app has been fitting for me. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 11. | I would use this app again. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 12. | Overall, I am satisfied with this app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 13. | The app would be useful for my healthcare practice. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 14. | The app improved my access to delivering healthcare services. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 15. | The app helped me manage my patients' health effectively. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |

| | | | |
|---|---|---|---|
| 16. | This app has all the functions and capabilities I expected it to have. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 17. | I could use the app even when the Internet connection was poor or not available. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 18. | This mHealth app provides an acceptable way to deliver healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |

In this questionnaire, 1 - strongly disagree, 2 – disagree, 3 – somewhat disagree, 4 – neither agree nor disagree, 5 – somewhat agree, 6 – agree, 7 – strongly agree

## Appendix A.2 Custom Open-Ended Usability Evaluation Form

Please answer the following questions based on how you feel after using the app today.

| |
|---|
| How easy do you think it is to perform tasks using this app? |
| |
| How quickly do you think you can perform tasks using this app after the training? |
| |

| |
|---|
| How pleasant is it to use the app to perform the tasks? |
| |
| What do you think can be improved from the app? |
| |
| What is your overall impression in using the app? |
| |

## Appendix B Usability Questionnaires Used for the At-Home App Evaluation

### Appendix B.1 mHealth App Usability Questionnaire (MAUQ) for Standalone Mobile Health App Used by Patients

The following MAUQ form (Zhou et al., 2019) was retrieved from https://ux.hari.pitt.edu/v2/portal/#/.

**mHealth App Usability Questionnaire (MAUQ)**

for Standalone mHealth Apps Used by Patients

| # | Statements | N/A | 1 2 3 4 5 6 7 |
|---|---|---|---|
| 1. | The app was easy to use. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 2. | It was easy for me to learn to use the app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 3. | The navigation was consistent when moving between screens. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 4. | The interface of the app allowed me to use all the functions (such as entering information, responding to reminders, viewing information) offered by the app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |

| | | | |
|---|---|---|---|
| 5. | Whenever I made a mistake using the app, I could recover easily and quickly. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 6. | I like the interface of the app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 7. | The information in the app was well organized, so I could easily find the information I needed. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 8. | The app adequately acknowledged and provided information to let me know the progress of my action. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 9. | I feel comfortable using this app in social settings. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 10. | The amount of time involved in using this app has been fitting for me. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 11. | I would use this app again. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 12. | Overall, I am satisfied with this app. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 13. | The app would be useful for my health and well-being. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 14. | The app improved my access to healthcare services. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 15. | The app helped me manage my health effectively. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |

| | | | |
|---|---|---|---|
| 16. | This app has all the functions and capabilities I expected it to have. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 17. | I could use the app even when the Internet connection was poor or not available. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |
| 18. | This mHealth app provides an acceptable way to receive healthcare services, such as accessing educational materials, tracking my own activities, and performing self-assessment. | ☐ | DISAGREE ☐ ☐ ☐ ☐ ☐ ☐ ☐ AGREE |

In this questionnaire, 1 - strongly disagree, 2 – disagree, 3 – somewhat disagree, 4 – neither agree nor disagree, 5 – somewhat agree, 6 – agree, 7 – strongly agree

## Appendix B.2 Custom Open-Ended Usability Evaluation Form

Please answer the following questions based on how you feel after using the app today.

| |
|---|
| How easy do you think it is to perform tasks using this app? |
| |
| How quickly do you think you can perform tasks using this app after the training? |
| |

| |
|---|
| How pleasant is it to use the app to perform the tasks? |
| |
| What do you think can be improved from the app? |
| |
| What is your overall impression in using the app? |
| |

## Appendix C EMA Forms in the At-Home App

**Morning Assessment (12 AM – 12 PM; By default, reminder is set at 8 AM)**

- Rate your level of low back pain right now
    - 0 (No pain) – 10 (Worst pain imaginable)
- Rate how much your pain is interfering with what you are doing right now
    - 0 (Pain is not interfering) – 10 (Pain is completely interfering)
- What time did you fall asleep?
- What time did you wake up?

**Afternoon Assessment (12 PM – 6 PM; By default, reminder is set at 1 PM)**

- Rate your level of low back pain right now
    - 0 (No pain) – 10 (Worst pain imaginable)
- Rate how much your pain is interfering with what you are doing right now
    - 0 (Pain is not interfering) – 10 (Pain is completely interfering)
- What activities did you do this morning? (Checkbox, can select more than one)
    - Sports/Exercise
    - Hobbies
    - Work, School, or Volunteer
    - Home activities

- How much effort did your activities require?
    - Sports/Exercise: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous
    - Hobbies: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous
    - Work, School, or Volunteer: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous
    - Home activities: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous

**Evening Assessment (6PM – 12 AM; By default, reminder is set at 7 PM)**

- Rate your level of low back pain right now
    - 0 (No pain) – 10 (Worst pain imaginable)
- Rate how much your pain is interfering with what you are doing right now
    - 0 (Pain is not interfering) – 10 (Pain is completely interfering)
- What activities did you do this afternoon? (Checkbox, can select more than one)
    - Sports/Exercise
    - Hobbies
    - Work, School, or Volunteer
    - Home activities
- How much effort did your activities require?
    - Sports/Exercise: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous
    - Hobbies: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous

    - Work, School, or Volunteer: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous
    - Home activities: Very Light, Light, Moderate, Moderate to Vigorous, Vigorous
- Was this a typical day for you?
    - Yes/No

# Appendix D Sensor Data

## Appendix D.1 At-Home Raw Sensor Data

Below is a snippet of the raw sensor data in space delimited format that was retrieved from the at-home sensors.

```
day hr min sec Ax(g) Ay(g) Az(g) Gx(deg/s) Gy(deg/s) Gz(deg/s) Mx(Gauss) My(Gauss) Mz(Gauss)
. . . . -0.00341600 -0.00012200 1.00125396 -23.74749947 10.98999977 0.40250000 0.20874001 -1.02508008 -0.03640000
. . . . 0.40894398 -0.04184600 1.33589995 463.19000244 -162.75000000 -46.27000046 0.20664001 -1.02955997 -0.03822000
. . . . -0.00341600 -0.00036600 0.97807395 -9.95750046 2.34500003 0.47250000 0.20314001 -1.02157998 -0.04116000
. . . . -0.00073200 0.00000000 0.97038800 -10.06250000 2.36249995 0.29750001 0.20706001 -1.02242005 -0.04410000
. . . . -0.00231800 0.00073200 0.97075397 -9.76500034 2.09999990 0.12250000 0.20650001 -1.02690005 -0.03738000
. . . . -0.00024400 0.00024400 0.97087598 -9.73000050 2.25749993 0.38499999 0.20328000 -1.02382004 -0.03640000
. . . . -0.00036600 -0.02989000 0.97075397 -9.81750011 2.15249991 0.36750001 0.20342000 -1.02564001 -0.04074000
. . . . -0.00073200 0.00097600 0.97172999 -10.13249969 2.48499990 0.33250001 0.20454000 -1.02564001 -0.04718000
. . . . -0.00268400 -0.00061000 0.97599995 -9.90499973 2.32750010 0.28000000 0.20734000 -1.02437997 -0.03444000
. . . . 0.00024400 -0.00061000 0.97416997 -9.52000046 2.46749997 0.42000002 0.21084000 -1.02634001 -0.03584000
. . . . 0.00268400 0.00195200 0.98051399 -9.69499969 1.92499995 0.17500000 0.20888001 -1.02802002 -0.03584000
. . . . 0.00012200 0.00000000 0.98441797 -9.83500004 2.20499992 0.42000002 0.20300001 -1.02620006 -0.04046000
. . . . -0.00622200 0.00341600 0.98210001 -9.99250031 2.94000006 0.34999999 0.20580001 -1.02410007 -0.03458000
. . . . 0.03050000 -0.00024400 0.98832196 -9.62500000 3.71000004 0.47250000 0.20118001 -1.02578008 -0.03374000
. . . . -0.26364198 0.17397200 1.15961003 -1.22500002 1.24249995 -0.14000000 0.20342000 -1.02354002 -0.03836000
. . . . -0.00488000 0.00109800 0.99417800 -10.72749996 2.43250012 0.42000002 0.20398000 -1.02157998 -0.04088000
. . . . -0.00988200 -0.00561200 0.97002196 -9.78250027 1.83749998 0.21000001 0.20944001 -1.01766002 -0.03556000
0 0 0 1 0.00353800 -0.00012200 0.97575599 -10.01000023 1.94250000 0.31500000 0.20566000 -1.01682007 -0.04200000
```

. . . . -0.00463600 0.00207400 0.97551197 -10.14999962 2.25749993 0.24500000 0.20790000 -1.01528001 -0.03808000

. . . . 0.00634400 0.00036600 0.97246200 -9.92249966 2.53749990 0.28000000 0.21028000 -1.01626003 -0.04004000

. . . . 0.00244000 -0.00024400 0.97538996 -10.43000031 2.45000005 0.26249999 0.20412001 -1.01864004 -0.04578000

. . . . 0.00695400 -0.00329400 0.96355599 -9.57250023 9.18750000 -0.29750001 0.20062001 -1.01654005 -0.04844000

. . . . 0.00292800 -0.00402600 0.96928996 -10.04500008 2.25749993 0.50749999 0.20020001 -1.01598001 -0.04886000

. . . . 0.15884399 -0.13578600 0.82972199 -29.64500046 -14.56000042 -23.37999916 0.19656001 -1.01472008 -0.04116000

. . . . -0.07930000 0.07576200 0.98612601 -28.92749977 -0.52499998 24.41250038 0.19712001 -1.01639998 -0.04004000

. . . . -0.09979600 0.06807600 1.22951603 -18.34000015 -2.29250002 -11.84749985 0.19936000 -1.00743997 -0.04060000

. . . . -0.20520400 -0.15579399 1.68555200 55.89500046 46.37500000 44.39749908 0.19208001 -0.99218005 -0.04508000

. . . . -0.24058400 -0.05002000 1.54061592 70.33249664 28.94499969 4.34000015 0.18746001 -0.95550001 -0.06720001

. . . . -0.20898600 -0.09442800 1.28417194 63.34999847 8.36499977 3.46499991 0.18844001 -0.91882002 -0.06636000

. . . . -0.17018999 -0.19812800 1.03016794 15.20750046 -28.26250076 -5.07499981 0.20174001 -0.88410002 -0.04536000

. . . . -0.14396000 -0.28438199 0.94366997 -45.51750183 -87.77999878 -21.82250023 0.22988001 -0.85427999 -0.02730000

. . . . -0.09345200 -0.50325000 0.88279200 -42.36750031 -108.13249969 -13.44000053 0.25760001 -0.84840000 -0.00140000

. . . . -0.21716000 -0.14847399 0.37734598 26.02250099 -8.48750019 35.47249985 0.28224000 -0.82754004 -0.01302000

. . . . -0.24900199 -0.25034401 0.49068400 -52.88499832 -49.89250183 -32.06000137 0.29708001 -0.82222003 -0.01568000

. . . . -0.11394800 -0.26901001 0.53265196 -16.45000076 -26.96750069 0.98000002 0.30954000 -0.82376003 -0.01694000

. . . . -0.05892600 -0.06209800 0.40150198 -11.28750038 34.17750168 18.42749977 0.32494000 -0.82110000 -0.02478000

. . . . -0.17153199 -0.04428600 0.63073999 -79.74749756 -6.68499994 -23.32749939 0.34622002 -0.82978004 -0.03542000

. . . . -0.00890600 -0.00268400 0.73382998 -59.58750153 14.14000034 -6.24749994 0.36848000 -0.84490001 -0.05194000

. . . . -0.06075600 -0.03416000 0.78946197 -49.96250153 42.96250153 9.06499958 0.39074001 -0.85427999 -0.06188000

. . . . -0.06148800 0.23985200 0.72553396 -16.52000046 41.72000122 -23.22249985 0.41160002 -0.85218000 -0.08848000

. . . . 0.07356600 0.05246000 0.88901401 -15.60999966 42.08750153 -17.93750000 0.40642002 -0.85190004 -0.11508001

. . . . 0.01476200 0.19215000 0.92488199 -50.83750153 55.72000122 -15.29500008 0.39214000 -0.84854001 -0.13566001

. . . . 0.15713599 0.08003200 1.04810202 -64.90750122 45.08000183 -54.72249985 0.36372000 -0.85806000 -0.12740001

. . . . 0.02501000 0.21618399 1.04492998 -60.13000107 28.13999939 -34.09000015 0.33698002 -0.86674005 -0.11886001

. . . . 0.15006000 0.11382600 1.15399802 -36.10250092 2.08249998 -49.10499954 0.31360000 -0.88004005 -0.08946000

. . . . 0.11382600 0.11736400 1.12325394 19.00499916 33.72249985 13.40499973 0.29750001 -0.88480002 -0.07588000

. . . . 0.11285000 0.10418800 1.01979792 -7.71750021 3.39499998 -2.08249998 0.28308001 -0.88550001 -0.05838000

. . . . 0.13529800 0.01122400 1.06127799 -3.44749999 5.26749992 24.79750061 0.26670000 -0.87934005 -0.03878000

. . . . 0.01427400 0.15762399 0.94684196 -26.09250069 -28.29750061 -15.73250008 0.25676000 -0.87052000 -0.01946000

## Appendix D.2 Processed Data from The At-Home Sensor Data

Below is a snippet of the processed sensor data in comma delimited or comma-separated value (csv) format.

ParticipantID,date,day,hour,Lvl1,Lvl2,Lvl3,Lvl4,Lvl5,total

R02,07/13/2021,1,7,59,1,0,0,0,60

R02,07/13/2021,1,8,35,25,0,0,0,60

R02,07/13/2021,1,9,32,27,1,0,0,60

R02,07/13/2021,1,10,20,37,3,0,0,60

R02,07/13/2021,1,11,20,38,2,0,0,60

R02,07/13/2021,1,12,15,41,4,0,0,60

R02,07/13/2021,1,13,16,44,0,0,0,60

R02,07/13/2021,1,14,13,47,0,0,0,60

R02,07/13/2021,1,15,16,44,0,0,0,60

R02,07/13/2021,1,16,5,50,5,0,0,60

R02,07/14/2021,2,7,59,1,0,0,0,60

R02,07/14/2021,2,8,20,37,3,0,0,60

R02,07/14/2021,2,9,28,30,2,0,0,60

R02,07/14/2021,2,10,59,1,0,0,0,60

R02,07/14/2021,2,11,33,23,4,0,0,60

R02,07/14/2021,2,12,43,15,2,0,0,60

R02,07/14/2021,2,13,33,20,7,0,0,60

R02,07/14/2021,2,14,8,25,27,0,0,60

R02,07/14/2021,2,15,25,32,3,0,0,60

R02,07/14/2021,2,16,33,27,0,0,0,60

R02,07/15/2021,3,11,22,34,4,0,0,60

R02,07/15/2021,3,12,33,26,1,0,0,60

R02,07/15/2021,3,13,30,30,0,0,0,60

R02,07/15/2021,3,14,53,6,1,0,0,60

R02,07/15/2021,3,15,60,0,0,0,0,60

R02,07/15/2021,3,16,60,0,0,0,0,60

R02,07/16/2021,4,7,57,3,0,0,0,60

R02,07/16/2021,4,8,37,23,0,0,0,60

R02,07/16/2021,4,9,32,28,0,0,0,60

R02,07/16/2021,4,10,31,29,0,0,0,60

R02,07/16/2021,4,11,1,5,0,0,0,6

R02,07/16/2021,4,12,0,0,0,0,0,0

R02,07/16/2021,4,13,0,0,0,0,0,0

R02,07/16/2021,4,14,0,0,0,0,0,0

R02,07/16/2021,4,15,0,0,0,0,0,0

R02,07/16/2021,4,16,0,0,0,0,0,0

R02,07/17/2021,5,7,0,0,0,0,0,0

R02,07/17/2021,5,8,0,0,0,0,0,0

R02,07/17/2021,5,9,0,0,0,0,0,0

R02,07/17/2021,5,10,0,0,0,0,0,0

R02,07/15/2021,3,7,60,0,0,0,0,60

R02,07/15/2021,3,8,41,19,0,0,0,60

R02,07/15/2021,3,9,27,33,0,0,0,60

R02,07/15/2021,3,10,21,37,2,0,0,60

R02,07/17/2021,5,11,0,0,0,0,0,0

R02,07/17/2021,5,12,0,0,0,0,0,0

R02,07/17/2021,5,13,0,0,0,0,0,0

R02,07/17/2021,5,14,0,0,0,0,0,0

## Appendix E Labeled Data from Dataset Builder Component

Below is a snippet of the labeled data created using the dataset builder component.


"id","participant_id","test_name","starting_time_range_in_second","ending_time_range_in_second","starting_timestamp","ending_timestamp","label","ending_percentage_of_video_duration","video_duration"

674549,"R01","Dynamic Motion - Axial Rotation",0,0.4,"2021-07-09T09:59:00","2021-07-09T09:59:00","static stand",5,7.398

674550,"R01","Dynamic Motion - Axial Rotation",0.4,1.02,"2021-07-09T09:59:00","2021-07-09T09:59:01","left rotation",13,7.398

674551,"R01","Dynamic Motion - Axial Rotation",1.02,1.58,"2021-07-09T09:59:01","2021-07-09T09:59:01","right rotation",21,7.398

674552,"R01","Dynamic Motion - Axial Rotation",1.58,2.13,"2021-07-09T09:59:01","2021-07-09T09:59:02","left rotation",28,7.398

674553,"R01","Dynamic Motion - Axial Rotation",2.13,2.74,"2021-07-09T09:59:02","2021-07-09T09:59:02","right rotation",37,7.398

674554,"R01","Dynamic Motion - Axial Rotation",2.74,3.23,"2021-07-09T09:59:02","2021-07-09T09:59:03","left rotation",43,7.398

674555,"R01","Dynamic Motion - Axial Rotation",3.23,3.76,"2021-07-09T09:59:03","2021-07-09T09:59:03","right rotation",50,7.398

674556,"R01","Dynamic Motion - Axial Rotation",3.76,4,"2021-07-09T09:59:03","2021-07-09T09:59:04","static stand",54,7.398

674557,"R01","Dynamic Motion - Axial Rotation",4,4.5,"2021-07-09T09:59:04","2021-07-09T09:59:04","right rotation",60,7.398

674558,"R01","Dynamic Motion - Axial Rotation",4.5,5.12,"2021-07-09T09:59:04","2021-07-09T09:59:05","left rotation",69,7.398

674559,"R01","Dynamic Motion - Axial Rotation",5.12,5.62,"2021-07-09T09:59:05","2021-07-09T09:59:05","right rotation",75,7.398

674560,"R01","Dynamic Motion - Axial Rotation",5.62,6.17,"2021-07-09T09:59:05","2021-07-09T09:59:06","left rotation",83,7.398

674561,"R01","Dynamic Motion - Axial Rotation",6.17,6.68,"2021-07-09T09:59:06","2021-07-09T09:59:06","right rotation",90,7.398

674562,"R01","Dynamic Motion - Axial Rotation",6.68,7.21,"2021-07-09T09:59:06","2021-07-09T09:59:07","left rotation",97,7.398

674563,"R01","Dynamic Motion - Axial Rotation",7.21,7.4,"2021-07-09T09:59:07","2021-07-09T09:59:07","static stand",100,7.398

674580,"R01","Dynamic Motion - Flexion",0,0.83,"2021-07-09T10:00:15","2021-07-09T10:00:15","forward bend",14,5.898

674581,"R01","Dynamic Motion - Flexion",0.83,1.86,"2021-07-09T10:00:15","2021-07-09T10:00:16","backward bend",31,5.898

674582,"R01","Dynamic Motion - Flexion",1.86,2.86,"2021-07-09T10:00:16","2021-07-09T10:00:17","forward bend",48,5.898

674583,"R01","Dynamic Motion - Flexion",2.86,3.79,"2021-07-09T10:00:17","2021-07-09T10:00:18","backward bend",64,5.898

674584,"R01","Dynamic Motion - Flexion",3.79,4.75,"2021-07-09T10:00:18","2021-07-09T10:00:19","forward bend",80,5.898

674585,"R01","Dynamic Motion - Flexion",4.75,5.57,"2021-07-09T10:00:19","2021-07-09T10:00:20","backward bend",94,5.898

674586,"R01","Dynamic Motion - Flexion",5.57,5.9,"2021-07-09T10:00:20","2021-07-09T10:00:20","static stand",100,5.898

674566,"R01","Dynamic Motion - Lateral Bending",0,0.32,"2021-07-09T09:59:33","2021-07-09T09:59:33","static stand",4,7.331

674567,"R01","Dynamic Motion - Lateral Bending",0.32,0.97,"2021-07-09T09:59:33","2021-07-09T09:59:33","left bend",13,7.331

674568,"R01","Dynamic Motion - Lateral Bending",0.97,1.51,"2021-07-09T09:59:33","2021-07-09T09:59:34","right bend",20,7.331

674569,"R01","Dynamic Motion - Lateral Bending",1.51,2.05,"2021-07-09T09:59:34","2021-07-09T09:59:35","left bend",27,7.331

## Bibliography


Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-Centered Design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, *37*(4), 445–456.

Allegri, M., Montella, S., Salici, F., Valente, A., Marchesini, M., Compagnone, C., … Fanelli, G. (2016). Mechanisms of low back pain: a guide for diagnosis and therapy. *F1000Research*, *5*, 1530. https://doi.org/10.12688/f1000research.8105.2

Armey, M. F., Schatten, H. T., Haradhvala, N., & Miller, I. W. (2015). Ecological momentary assessment (EMA) of depression-related phenomena. *Current Opinion in Psychology*, *4*, 21–25. https://doi.org/10.1016/J.COPSYC.2015.01.002

Banos, O., Villalonga, C., Damas, M., Gloesekoetter, P., Pomares, H., & Rojas, I. (2014). PhysioDroid: Combining Wearable Health Sensors and Mobile Devices for a Ubiquitous, Continuous, and Personal Monitoring. *Scientific World Journal*, *2014*. https://doi.org/10.1155/2014/490824

Baxter, G., Bell, A., Allen, J., & Ravey, J. (1991). Low Level Laser Therapy: Current Clinical Practice in Northern Ireland. *Physiotherapy*, *77*(3), 171–178. https://doi.org/10.1016/S0031-9406(10)61696-3

Bell, K. M., Onyeukwu, C., McClincy, M. P., Allen, M., Bechard, L., Mukherjee, A., … Irrgang, J. J. (2019). Verification of a portable motion tracking system for remote management of physical rehabilitation of the knee. *Sensors (Switzerland)*, *19*(5). https://doi.org/10.3390/s19051021

Benson, L. C., Clermont, C. A., Bošnjak, E., & Ferber, R. (2018). The use of wearable devices for walking and running gait analysis outside of the lab: A systematic review. *Gait & Posture*, *63*, 124–138. https://doi.org/10.1016/J.GAITPOST.2018.04.047

Bisio, I., Lavagetto, F., Marchese, M., & Sciarrone, A. (2015). A smartphone-centric platform for remote health monitoring of heart failure. *International Journal of Communication Systems*, *28*(11), 1753–1771. https://doi.org/10.1002/dac.2778

Bonato, P. (2010). Wearable sensors and systems. *IEEE Engineering in Medicine and Biology Magazine*, *29*(3), 25–36. https://doi.org/10.1109/MEMB.2010.936554

Brooke, J. (1996). SUS - A quick and dirty usability scale. *Usability Evaluation in Industry*, *189*. Retrieved from http://liacs.leidenuniv.nl/~verbeekfj/courses/hci/SUS-questionaire.pdf

Carmody, T. P. (2001). Psychosocial subgroups, coping, and chronic low-back pain. *Journal of Clinical Psychology in Medical Settings*, *8*(3), 137–148.

https://doi.org/10.1023/A:1011309401598/METRICS

Choi, S., Han, S. I., Jung, D., Hwang, H. J., Lim, C., Bae, S., … Kim, D. H. (2018). Highly conductive, stretchable and biocompatible Ag–Au core–sheath nanowire composite for wearable and implantable bioelectronics. *Nature Nanotechnology*, *13*(11), 1048–1056. https://doi.org/10.1038/s41565-018-0226-8

Chowdhary, K., Yu, D. X., Pramana, G., Mesoros, M., Fairman, A., Dicianno, B. E., & Parmanto, B. (2022). User-Centered Design to Enhance mHealth Systems for Individuals With Dexterity Impairments: Accessibility and Usability Study. *JMIR Human Factors*, *9*(1), e23794. https://doi.org/10.2196/23794

Collins, L. M., Murphy, S. A., & Bierman, K. L. (2004, September). A conceptual framework for adaptive preventive interventions. *Prevention Science*. Springer. https://doi.org/10.1023/B:PREV.0000037641.26017.00

Couture, B., Lilley, E., Chang, F., Smith, A. D., Cleveland, J., Ergai, A., … Collins, S. A. (2018). Applying user-centered design methods to the development of an mhealth application for use in the hospital setting by patients and care partners. *Applied Clinical Informatics*, *9*(2), 302–312. https://doi.org/10.1055/S-0038-1645888/ID/OR170198RA-23

Delitto, A., Erhard, R. E., Bowling, R. W., DeRosa, C. P., & Greathouse, D. G. (1995). A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Physical Therapy*, *75*(6), 470–489. https://doi.org/10.1093/PTJ/75.6.470

Delitto, A., George, S. Z., Van Dillen, L. R., Whitman, J. M., Sowa, G., Shekelle, P., … Orthopaedic Section of the American Physical Therapy Association. (2012). Low back pain. *The Journal of Orthopaedic and Sports Physical Therapy*, *42*(4). https://doi.org/10.2519/jospt.2012.42.4.a1

Dicianno, B. E., Parmanto, B., Fairman, A. D., Crytzer, T. M., Yu, D. X., Pramana, G., … Petrazzi, A. A. (2015). Perspectives on the Evolution of Mobile (mHealth) Technologies and Application to Rehabilitation. *Physical Therapy*, *95*(3), 397–405. https://doi.org/10.2522/PTJ.20130534

Dinh-Le, C., Chuang, R., Chokshi, S., & Mann, D. (2019, September 11). Wearable health technology and electronic health record integration: Scoping review and future directions. *JMIR MHealth and UHealth*. JMIR Publications Inc. https://doi.org/10.2196/12861

Dobkin, B. H., & Dorsch, A. (2011). The promise of mHealth: Daily activity monitoring and outcome assessments by wearable sensors. *Neurorehabilitation and Neural Repair*, *25*(9), 788–798. https://doi.org/10.1177/1545968311425908

Fairbank, J., Gwilym, S. E., France, J. C., Daffner, S. D., Dettori, J., Hermsmeyer, J., & Andersson, G. (2011). The role of classification of chronic low back pain. *Spine*, *36*(21 SUPPL.). https://doi.org/10.1097/BRS.0B013E31822EF72C

Fairman, A. D., Yih, E. T., McCoy, D. F., LoPresti, E. F., McCue, M. P., Parmanto, B., & Dicianno, B. E. (2016). Iterative Design and Usability Testing of the iMHere System for Managing Chronic Conditions and Disability. *International Journal of Telerehabilitation*, *8*(1), 11–20. https://doi.org/10.5195/ijt.2016.6194

Farao, J., Malila, B., Conrad, N., Mutsvangwa, T., Rangaka, M. X., & Douglas, T. S. (2020). A user-centred design framework for mHealth. *PLOS ONE*, *15*(8), e0237910. https://doi.org/10.1371/JOURNAL.PONE.0237910

Feuerstein, M., & Beattie, P. (1995). Biobehavioral factors affecting pain and disability in low back pain: mechanisms and assessment. *Physical Therapy*, *75*(4), 267–280. https://doi.org/10.1093/PTJ/75.4.267

Free, C., Phillips, G., Felix, L., Galli, L., Patel, V., & Edwards, P. (2010a). The effectiveness of M-health technologies for improving health and health services: A systematic review protocol. *BMC Research Notes*, *3*(1), 250. https://doi.org/10.1186/1756-0500-3-250

Free, C., Phillips, G., Felix, L., Galli, L., Patel, V., & Edwards, P. (2010b). The effectiveness of M-health technologies for improving health and health services: A systematic review protocol. *BMC Research Notes*, *3*. https://doi.org/10.1186/1756-0500-3-250

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise*, *30*(5), 777–781. https://doi.org/10.1097/00005768-199805000-00021

French, S. D., Cameron, M., Walker, B. F., Reggars, J. W., & Esterman, A. J. (2006). Superficial heat or cold for low back pain. *Cochrane Database of Systematic Reviews*, *2011*(2). https://doi.org/10.1002/14651858.CD004750.PUB2/INFORMATION/EN

Furlan, A. D., Brosseau, L., Imamura, M., & Irvin, E. (2002). Massage for low-back pain: A systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine*, *27*(17), 1896–1910. https://doi.org/10.1097/00007632-200209010-00017

Gay, V., & Leijdekkers, P. (2007). *A health monitoring system using smart phones and wearable sensors*. *International Journal of ARM* (Vol. 8). Retrieved from https://www.researchgate.net/publication/228343525

Georgakakis, E., Nikolidakis, S. A., Vergados, D. D., & Douligeris, C. (2011). An analysis of bluetooth, Zigbee and bluetooth low energy and their use in WBANs. In *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering* (Vol. 55 LNICST, pp. 168–175). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-20865-2_22

George, S. Z., Goertz, C., Hastings, S. N., & Fritz, J. M. (2020). Transforming low back pain care delivery in the United States. *Pain*, *161*(12), 2667. https://doi.org/10.1097/J.PAIN.0000000000001989

Ghika, J., Wiegner, A. W., Fang, J. J., Davies, L., Young, R. R., & Growdon, J. H. (1993).

Portable System for Quantifying Motor Abnormalities in Parkinson's Disease. *IEEE Transactions on Biomedical Engineering*, *40*(3), 276–283. https://doi.org/10.1109/10.216411

Goldstein, S. P., Evans, B. C., Flack, D., Juarascio, A., Manasse, S., Zhang, F., & Forman, E. M. (2017). Return of the JITAI: Applying a Just-in-Time Adaptive Intervention Framework to the Development of m-Health Solutions for Addictive Behaviors. *International Journal of Behavioral Medicine*, *24*(5), 673–682. https://doi.org/10.1007/s12529-016-9627-y

Guzmán, J., Esmail, R., Karjalainen, K., Malmivaara, A., Irvin, E., & Bombardier, C. (2001). Multidisciplinary rehabilitation for chronic low back pain: Systematic review. *British Medical Journal*, *322*(7301), 1511–1516. https://doi.org/10.1136/bmj.322.7301.1511

Hartvigsen, J., Hancock, M. J., Kongsted, A., Louw, Q., Ferreira, M. L., Genevay, S., … Woolf, A. (2018). What low back pain is and why we need to pay attention. *The Lancet*, *391*(10137), 2356–2367. https://doi.org/10.1016/S0140-6736(18)30480-X

Hayden, J. A., Ellis, J., Ogilvie, R., Malmivaara, A., & van Tulder, M. W. (2021). Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews*, *2021*(9). https://doi.org/10.1002/14651858.CD009790.PUB2/INFORMATION/EN

Henschke, N., Ostelo, R. W. J. G., van Tulder, M. W., Vlaeyen, J. W. S., Morley, S., Assendelft, W. J. J., & Main, C. J. (2010). Behavioural treatment for chronic low-back pain. *Cochrane Database of Systematic Reviews*, *2011*(2). https://doi.org/10.1002/14651858.CD002014.pub3

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, *15*(1), 1–39. https://doi.org/10.1348/135910709X466063

Heymans, M. W., van Tulder, M. W., Esmail, R., Bombardier, C., & Koes, B. W. (2004). Back schools for non-specific low-back pain. *The Cochrane Database of Systematic Reviews*, (4). https://doi.org/10.1002/14651858.CD000261.PUB2

Holthe, T., Halvorsrud, L., Karterud, D., Hoel, K. A., & Lund, A. (2018). Usability and acceptability of technology for community-dwelling older adults with mild cognitive impairment and dementia: A systematic literature review. *Clinical Interventions in Aging*, *13*, 863–886. https://doi.org/10.2147/CIA.S154717

Hoy, D., Bain, C., Williams, G., March, L., Brooks, P., Blyth, F., … Buchbinder, R. (2012). A systematic review of the global prevalence of low back pain. *Arthritis and Rheumatism*, *64*(6), 2028–2037. https://doi.org/10.1002/art.34347

Istepanian, R., Laxminarayan, S., & Pattichis, C. (2007). *M-health: Emerging mobile health systems*. Retrieved from https://books.google.com/books?hl=en&lr=&id=2dAYdPigj5wC&oi=fnd&pg=PR8&ots=kbWMs7wBhB&sig=x4k39s8dIYA4mzGXgi5WeOS34Vc

Istepanian, R. S. H., Jovanov, E., & Zhang, Y. T. (2004). Introduction to the special section on m-Health: Beyond seamless mobility and global wireless health-care connectivity. *IEEE Transactions on Information Technology in Biomedicine*. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/TITB.2004.840019

Jovanov, E. (2005). Wireless technology and system integration in body area networks for m-Health applications. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings* (Vol. 7 VOLS, pp. 7158–7160). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/iembs.2005.1616158

Jovanov, E., Milenkovic, A., Otto, C., & De Groen, P. C. (2005). A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, *2*(1), 6. https://doi.org/10.1186/1743-0003-2-6

Karayannis, N. V., Jull, G. A., & Hodges, P. W. (2012). Physiotherapy movement based classification approaches to low back pain: Comparison of subgroups through review and developer/expert survey. *BMC Musculoskeletal Disorders*, *13*(1), 1–15. https://doi.org/10.1186/1471-2474-13-24/TABLES/5

Khadilkar, A., Milne, S., Brosseau, L., Robinson, V., Saginur, M., Shea, B., … Wells, G. (2005). Transcutaneous electrical nerve stimulation (TENS) for chronic low-back pain. *Cochrane Database of Systematic Reviews*, (3). https://doi.org/10.1002/14651858.cd003008.pub2

Koes, B. W., Van Tulder, M. W., & Thomas, S. (2006, June 17). Diagnosis and treatment of low back pain. *British Medical Journal*. BMJ Publishing Group. https://doi.org/10.1136/bmj.332.7555.1430

Kumar, D., Jeuris, S., Bardram, J. E., & Dragoni, N. (2020). Mobile and Wearable Sensing Frameworks for mHealth Studies and Applications. *ACM Transactions on Computing for Healthcare* , *2*(1). https://doi.org/10.1145/3422158

Kyerematen, N. A., Brousse, T., & Pech, D. (2017, January 10). Microsupercapacitors as miniaturized energy-storage components for on-chip electronics. *Nature Nanotechnology*. Nature Publishing Group. https://doi.org/10.1038/nnano.2016.196

Liew, M. S., Zhang, J., See, J., & Ong, Y. L. (2019). Usability Challenges for Health and Wellness Mobile Apps: Mixed-Methods Study Among mHealth Experts and Consumers. *JMIR Mhealth Uhealth 2019;7(1):E12160 Https://Mhealth.Jmir.Org/2019/1/E12160*, *7*(1), e12160. https://doi.org/10.2196/12160

Lim, H. R., Kim, H. S., Qazi, R., Kwon, Y. T., Jeong, J. W., & Yeo, W. H. (2020, April 1). Advanced Soft Materials, Sensor Integrations, and Applications of Wearable Flexible Hybrid Electronics in Healthcare, Energy, and Environment. *Advanced Materials*. Wiley-VCH Verlag. https://doi.org/10.1002/adma.201901924

Lin, W. C., Burke, L., Schlenk, E. A., & Yeh, C. H. (2019). Use of an Ecological Momentary Assessment Application to Assess the Effects of Auricular Point Acupressure for Chronic Low Back Pain. *CIN - Computers Informatics Nursing*, *37*(5), 276–282.

https://doi.org/10.1097/CIN.0000000000000478

Litt, M. D., Cooney, N. L., & Morse, P. (1998). Ecological momentary assessment (EMA) with treated alcoholics: Methodological problems and potential solutions. *Health Psychology*, *17*(1), 48–52. https://doi.org/10.1037/0278-6133.17.1.48

Lobelo, F., Kelli, H. M., Tejedor, S. C., Pratt, M., McConnell, M. V., Martin, S. S., & Welk, G. J. (2016, May 1). The Wild Wild West: A Framework to Integrate mHealth Software Applications and Wearables to Support Physical Activity Assessment, Counseling and Interventions for Cardiovascular Disease Risk Reduction. *Progress in Cardiovascular Diseases*. W.B. Saunders. https://doi.org/10.1016/j.pcad.2016.02.007

Lucas, J. W., Connor, E. M., & Bose, J. (2021). Back, lower limb, and upper limb pain among U.S. adults, 2019. https://doi.org/10.15620/CDC:107894

Marras, W. S. (2012). The complex spine: The multidimensional system of causal pathways for low-back disorders. *Human Factors*, *54*(6), 881–889. https://doi.org/10.1177/0018720812452129

May, M., Junghaenel, D. U., Ono, M., Stone, A. A., & Schneider, S. (2018). Ecological Momentary Assessment Methodology in Chronic Pain Research: A Systematic Review. *Journal of Pain*, *19*(7), 699–716. https://doi.org/10.1016/j.jpain.2018.01.006

Mukhopadhyay, S. C. (2015). Wearable sensors for human activity monitoring: A review. *IEEE Sensors Journal*. IEEE. https://doi.org/10.1109/JSEN.2014.2370945

Munos, B., Baker, P. C., Bot, B. M., Crouthamel, M., de Vries, G., Ferguson, I., … Wang, P. (2016). Mobile health: the power of wearables, sensors, and apps to transform clinical trials. *Annals of the New York Academy of Sciences*, *1375*(1), 3–18. https://doi.org/10.1111/NYAS.13117

Murphy, K. R., Han, J. L., Yang, S., Hussaini, S. M., Elsamadicy, A., Parente, B., … Lad, S. P. (2017). Prevalence of Specific Types of Pain Diagnoses in a Sample of United States Adults. *Pain Physician*, *20*(2), 257–268. Retrieved from https://pubmed.ncbi.nlm.nih.gov/28158163/

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*, *34*(0), 1209–1219. https://doi.org/10.1037/hea0000306

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Neishabouri, A., Nguyen, J., Samuelsson, J., Guthrie, T., Biggs, M., Wyatt, J., … Guo, C. C. (2022). Quantification of acceleration as activity counts in ActiGraph wearable. *Scientific Reports 2022 12:1*, *12*(1), 1–8. https://doi.org/10.1038/s41598-022-16003-x

O'Reilly, G. A., & Spruijt-Metz, D. (2013). Current mHealth Technologies for Physical Activity Assessment and Promotion. *American Journal of Preventive Medicine*, *45*(4), 501–507. https://doi.org/10.1016/J.AMEPRE.2013.05.012

Ostelo, R. W., van Tulder, M. W., Vlaeyen, J. W., Linton, S. J., Morley, S., & Assendelft, W. J. (2005). Behavioural treatment for chronic low-back pain. *Cochrane Database of Systematic Reviews*, (1). https://doi.org/10.1002/14651858.cd002014.pub2

Parmanto, B., Lewis, Jr., A. N., Graham, K. M., & Bertolet, M. H. (2016). Development of the Telehealth Usability Questionnaire (TUQ). *International Journal of Telerehabilitation*, *8*(1), 3. https://doi.org/10.5195/IJT.2016.6196

Parmanto, B., Pramana, G., Yu, D. X., Fairman, A. D., Dicianno, B. E., & McCue, M. P. (2013). iMHere: A Novel mHealth System for Supporting Self-Care in Management of Complex and Chronic Conditions. *JMIR Mhealth Uhealth 2013;1(2):E10 Https://Mhealth.Jmir.Org/2013/2/E10*, *1*(2), e2391. https://doi.org/10.2196/MHEALTH.2391

Parreira, P., Heymans, M. W., van Tulder, M. W., Esmail, R., Koes, B. W., Poquet, N., … Maher, C. G. (2017). Back Schools for chronic non-specific low back pain. *The Cochrane Database of Systematic Reviews*, *8*(8). https://doi.org/10.1002/14651858.CD011674.PUB2

Peck, J., Urits, I., Peoples, S., Foster, L., Malla, A., Berger, A. A., … Viswanath, O. (2021). A Comprehensive Review of Over the Counter Treatment for Chronic Low Back Pain. *Pain and Therapy*, *10*(1), 69–80. https://doi.org/10.1007/S40122-020-00209-W/METRICS

Pramana, G., Parmanto, B., Kendall, P. C., & Silk, J. S. (2014). The SmartCAT: An m-Health Platform for Ecological Momentary Intervention in Child Anxiety Treatment. *Https://Home.Liebertpub.Com/Tmj*, *20*(5), 419–427. https://doi.org/10.1089/TMJ.2013.0214

Quirk, A. D., Johnson, M. E., Anderson, D. E., Smuck, M., Sun, R., Matthew, R., … Bowden, A. E. (2022). Biomechanical Phenotyping of Chronic Low Back Pain: Protocol for BACPAC. *Pain Medicine*. https://doi.org/10.1093/pm/pnac163

Reid, M. C., Williams, C. S., & Gill, T. M. (2005). Back pain and decline in lower extremity physical function among community-dwelling older persons. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *60*(6), 793–797. https://doi.org/10.1093/gerona/60.6.793

Rodrigues, C. P., da Silva, R. A., Neto, E. N., Andraus, R. A. C., Fernandes, M. T. P., & Fernandes, K. B. P. (2017). Analysis of functional capacity in individuals with and without chronic lower back pain. *Acta Ortopedica Brasileira*, *25*(4), 143–146. https://doi.org/10.1590/1413-785220172504156564

Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Dieguez, A., Carry, M., … Travers, J. (2016). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of Biomedical Informatics*, *60*, 243–251.

https://doi.org/10.1016/J.JBI.2016.02.002

Setiawan, I. M. A., Zhou, L., Alfikri, Z., Saptono, A., Fairman, A. D., Dicianno, B. E., & Parmanto, B. (2019). An Adaptive Mobile Health System to Support Self-Management for Persons With Chronic Conditions and Disabilities: Usability and Feasibility Studies. *JMIR Form Res 2019;3(2):E12982 Https://Formative.Jmir.Org/2019/2/E12982*, *3*(2), e12982. https://doi.org/10.2196/12982

Shiffman, S. (2009). Ecological Momentary Assessment (EMA) in Studies of Substance Use. *Psychological Assessment*, *21*(4), 486–497. https://doi.org/10.1037/A0017074

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Https://Doi.Org/10.1146/Annurev.Clinpsy.3.022806.091415*, *4*, 1–32. https://doi.org/10.1146/ANNUREV.CLINPSY.3.022806.091415

Shokri, S., Ward, S., Anton, P.-A. M., Siffredi, P., & Papetti, G. (2020). Recent Advances in Wearable Sensors with Application in Rehabilitation Motion Analysis. Retrieved from http://arxiv.org/abs/2009.06062

Smuck, M., Tomkins-Lane, C., Ith, M. A., Jarosz, R., & Kao, M. C. J. (2017). Physical performance analysis: A new approach to assessing free-living physical activity in musculoskeletal pain and mobility-limited populations. *PLOS ONE*, *12*(2), e0172804. https://doi.org/10.1371/JOURNAL.PONE.0172804

Solhan, M. B., Trull, T. J., Jahng, S., & Wood, P. K. (2009). Clinical Assessment of Affective Instability: Comparing EMA Indices, Questionnaire Reports, and Retrospective Recall. *Psychological Assessment*, *21*(3), 425–436. https://doi.org/10.1037/A0016869

Song, Y., Min, J., & Gao, W. (2019, November 26). Wearable and Implantable Electronics: Moving toward Precision Therapy. *ACS Nano*. American Chemical Society. https://doi.org/10.1021/acsnano.9b08323

Spieker, S., Jentgens, C., Boose, A., & Dichgans, J. (1995). Reliability, specificity and sensitivity of long-term tremor recordings. *Electroencephalography and Clinical Neurophysiology/ Electromyography*, *97*(6), 326–331. https://doi.org/10.1016/0924-980X(95)00208-3

Steinmetz, A. (2022). Back pain treatment: a new perspective. *Therapeutic Advances in Musculoskeletal Disease*, *14*. https://doi.org/10.1177/1759720X221100293

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2002). Patient non-compliance with paper diaries. *BMJ*, *324*(7347), 1193–1194. https://doi.org/10.1136/BMJ.324.7347.1193

Teng, X. F., Poon, C. C. Y., Zhang, Y. T., & Bonato, P. (2008). Wearable Medical Systems for p-Health. *IEEE Reviews in Biomedical Engineering*, *1*, 62–74. https://doi.org/10.1109/RBME.2008.2008248

Thilarajah, S., Clark, R. A., & Williams, G. (2016). Wearable sensors and Mobile Health

(mHealth) technologies to assess and promote physical activity in stroke: a narrative review. *Brain Impairment*, *17*(1), 34–42. https://doi.org/10.1017/BRIMP.2016.1

Tricoli, A., Nasiri, N., & De, S. (2017). Wearable and Miniaturized Sensor Technologies for Personalized and Preventive Medicine. *Advanced Functional Materials*, *27*(15), 1605271. https://doi.org/10.1002/adfm.201605271

Truchon, M. (2001). Determinants of chronic disability related to low back pain: towards an integrative biopsychosocial model. *Disability and Rehabilitation*, *23*(17), 758–767. https://doi.org/10.1080/09638280110061744

Urits, I., Burshtein, A., Sharma, M., Testa, L., Gold, P. A., Orhurhu, V., … Kaye, A. D. (2019). Low Back Pain, a Comprehensive Review: Pathophysiology, Diagnosis, and Treatment. *Current Pain and Headache Reports*, *23*(3), 1–10. https://doi.org/10.1007/S11916-019-0757-1/METRICS

Usability 101: Introduction to Usability. (n.d.). Retrieved March 20, 2021, from https://www.nngroup.com/articles/usability-101-introduction-to-usability/

User-Centered Design Basics | Usability.gov. (n.d.). Retrieved March 20, 2021, from https://www.usability.gov/what-and-why/user-centered-design.html

Van Duijvenbode, I. C. D., Jellema, P., Van Poppel, M. N. M., & Van Tulder, M. W. (2008). Lumbar supports for prevention and treatment of low back pain. *Cochrane Database of Systematic Reviews*, *2008*(2). https://doi.org/10.1002/14651858.CD001823.pub3

Van Middelkoop, M., Rubinstein, S. M., Kuijpers, T., Verhagen, A. P., Ostelo, R., Koes, B. W., & Van Tulder, M. W. (2011, January). A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *European Spine Journal*. Springer. https://doi.org/10.1007/s00586-010-1518-3

Van Middelkoop, M., Rubinstein, S. M., Verhagen, A. P., Ostelo, R. W., Koes, B. W., & van Tulder, M. W. (2010). Exercise therapy for chronic nonspecific low-back pain. *Best Practice & Research Clinical Rheumatology*, *24*(2), 193–204. https://doi.org/10.1016/J.BERH.2010.01.002

Vos, T., Allen, C., Arora, M., Barber, R. M., Brown, A., Carter, A., … Zuhlke, L. J. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1545–1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Wang, E. H., Zhou, L., Chen, S. H. K., Hill, K., & Parmanto, B. (2018). An mhealth platform for supporting clinical data integration into augmentative and alternative communication service delivery: User-centered design and usability evaluation. *JMIR Rehabilitation and Assistive Technologies*, *5*(2), e9009. https://doi.org/10.2196/rehab.9009

Waterman, B. R., Belmont, P. J., & Schoenfeld, A. J. (2012). Low back pain in the United

States: incidence and risk factors for presentation in the emergency setting. *The Spine Journal : Official Journal of the North American Spine Society*, *12*(1), 63–70. https://doi.org/10.1016/J.SPINEE.2011.09.002

Welhausen, C. A., & Bivens, K. M. (n.d.). mHealth Apps and Usability: Using User-Generated Content to Explore Users' Experiences with a Civilian First Responder App.

Wu, A., March, L., Zheng, X., Huang, J., Wang, X., Zhao, J., … Hoy, D. (2020). Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. *Annals of Translational Medicine*, *8*(6), 299–299. https://doi.org/10.21037/ATM.2020.02.175

Zhou, L., Bao, J., Setiawan, I. M. A., Saptono, A., & Parmanto, B. (2019). The mHealth app usability questionnaire (MAUQ): Development and validation study. *Journal of Medical Internet Research*, *21*(4), e11500. https://doi.org/10.2196/11500