Predictors of Medically Significant Responses to Mindfulness for Chronic Pain

Major Qualifying Project

A Major Qualifying Project submitted to the faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the Degree of Bachelor of Science

By:

Adeline Evans

Adina Palayoor

Report Submitted to:

Professor Carolina Ruiz, Advisor

Department of Computer Science and Data Science Program, WPI

Professor Benjamin Nephew, Co-Advisor

Department of Biology and Biotechnology and Neuroscience Program, WPI

Abstract

Chronic pain, affecting approximately 30% of the global population, presents complex challenges of physical discomfort accompanied by psychological distress. Current treatments involve a combination of practices such as lifestyle adjustments, opioid and non-opioid pharmacological therapies, psychological interventions, and integrative treatments. This project aimed to identify the characteristics of patients with chronic pain that are associated with a likelihood of achieving medically significant pain reduction through mindfulness-based treatment. Leveraging machine learning techniques, including Decision Trees, Random Forest, and XGBoost, this study delved into feature importance analysis for both classification and regression models to discern the key factors influencing treatment outcomes. The findings aspire to contribute to the nuanced understanding of pain treatment, potentially guiding future interventions toward more personalized and effective pain management strategies.

Table of Contents:

1. Introduction

Chronic pain is a widespread issue with intricate physical and psychological ramifications. It affects an estimated 30 percent of individuals worldwide making its treatment an important discussion in the medical field (Cohen et. al, 2021). The current efforts to treat chronic pain focus on managing pain and restoring physical function. These treatments involve a combination of practices such as lifestyle adjustments, opioid and non-opioid pharmacological therapies, psychological interventions, and integrative treatments.

As research continues in the area of pain management, mindfulness has become a growing topic because of its potential to treat chronic pain effectively. One of the main programs for mindfulness is called Mindfulness-Based Stress Reduction(MBSR), a structured program developed by Jon Kabat-Zinn in the 1980s. It is one of the most extensively studied forms of mindfulness training in the United States, making it an excellent candidate to be implemented in studies aiming to understand the effect of mindfulness on pain reduction. MSBR encompasses various mindfulness practices, focusing on enhancing the participant's ability to observe the immediate content of the experience(Goldin et.al, 2010). While MSBR does not explicitly aim to change the nature of thinking or emotional reactivity, it has demonstrated effectiveness in reducing stress and depression both of which are often connected with the severity or symptoms of chronic pain (Goldin et.al, 2010).

This Major Qualifying Project (MQP) explores the role of mindfulness as a treatment for chronic pain, delving into how mindfulness influences chronic pain and addressing the obstacles to its adoption. The dataset we are using comes from a study, “A Mind-Body Program for Older Adults With Chronic Low Back Pain - A Randomized Clinical Trial”, that researched how mindfulness affects the various pain scales and functional mobility in participants over the age of 65 with chronic lower pain(Morone, 2019). By looking into how the Mindfulness Based Stress Reduction program affected the pain levels and mobility of participants we can attempt to determine which attributes make a person more likely to respond well to this type of treatment. This may be in addition to other pain management strategies or as a substitute for pain treatments with adverse side effects. We applied various machine learning techniques and feature selection methods to identify the characteristics of a patient likely to respond well to mindfulness-based treatment, thereby potentially improving care plans for patients with chronic pain.

2. Background

2.1 Chronic Pain

2.1.1 Overview of Chronic Pain

Chronic pain is an aversive sensory and emotional experience intricately linked to actual or potential tissue damage. Within the medical community, it is widely acknowledged that the temporal threshold for characterizing pain as chronic is when it persists for a duration surpassing three months. It is one of the leading causes for individuals seeking medical attention, with conditions such as osteoarthritis, back pain, and headaches prominently featured among the primary factors (Cohen et al., 2021). Chronic pain differs from acute pain because it presents multifaceted physiological and psychological implications. This type of pain can be distinguished by three main categories: nociceptive, originating from tissue injury; neuropathic, emanating from nerve damage; and neoplastic, emerging from a sensitized nervous system.

2.1.2 Pain Measures

There are many ways that studies have measured the changes in pain from before treatment to after treatment. These measurements could be in the format of questionnaires, surveys, scales, or mobility tests. In Table 1 below, we will summarize some of the main measurements used to determine not only pain levels for a patient but also various other factors that may be affecting chronic pain for the individual. These are the three main pain scales used, however, there are many more included in our dataset that we did not focus on here.

Pain Measurement	**Definition of Pain Scale**	**Rating**
Roland and Morris Disability Questionnaire(RMDQ)	Measure of functional limitations	A higher score indicates higher functional limitation, 2.5-5 was considered a clinically meaningful change in a previous study trying to determine the

		effect of mindfulness on chronic lower back pain(Frota et. al, 22)
Pain Numeric Rating(NRS)	Measure of pain levels	Higher score indicate worse pain, [0 = no pain to 20 = worst pain imaginable] (Correll, 2011)
McGill Pain Questionnaire	Measure of pain experience	Ranks various pain-related sensory adjectives as None, Mild, Moderate, and Severe (Melzack, 2005)
PEG	Pain Intensity and Interference Scale	3 questions ranged from 1-10, Average Pain, Pain Interference on Enjoyment of life, and Pain Interference with General Activity (Krebs et. al, 2009)
Patient-Reported Outcomes Measurement Information System(PROMIS)	Health Survey	Seven domains (physical functioning, anxiety, depression, fatigue, sleep disturbance, social functioning, and pain), and the pain domain has two subdomains (interference and intensity). Each of the seven domains has four 5-level items (i.e., 16 decrements each) (Hays et. al, 2019)

Table 1 Pain Scale Measures

2.1.3 Pain Management Strategies

Proper pain management strategies can significantly improve an individual's quality of life and functional mobility. Chronic pain does not pose a threat to one's survival and as a result, treatment is focused on restoring function and emotional well-being. Chronic pain is associated with specific alterations in the nervous system and significantly impacts the quality of life, but emotional support and good health can promote healing.

Numerous clinical trials and guidelines recommend a personalized, interdisciplinary approach to treatment. These treatments generally involve a combination of various practices such as lifestyle changes, opioids, non-opioid pharmacological therapies, psychological therapies, and integrative treatments and procedures(Cohen et.al, 2021)

The most commonly recommended pain management strategy is exercise. It may help improve sleep, aid in weight loss, trigger the release of endorphins, and reverse physical deconditioning(Geneen et. al, 2017). A study from the Cochrane Database of Systematic Reviews, with over 37,000 participants found that exercise is more effective for improving function than relieving pain(Geneen et. al, 2017). It has proved to be particularly beneficial for musculoskeletal and diffuse pain types but can also be used for neuropathic pain(Geneen et. al, 2017). There's no conclusive evidence supporting one specific exercise regimen, so a tailored program based on individual needs is the best option. For example, a low-level aerobic exercise for fibromyalgia or strength training for back pain could be recommended.

In terms of psychological intervention for chronic pain, Cognitive Behavioral Therapy(CBT), is often used as treatment. This therapy involves changing harmful beliefs, attitudes, and behaviors related to the condition(Lim et.al, 2018). Psychologists typically administer CBT but usually use it as part of a multidisciplinary approach with other chronic pain treatments. CBT has been studied extensively for various pain disorders. One systematic review showed that this type of therapy offers some short-term benefits compared to the usual treatment, however, it is important to note that it is most effective for individuals who have mood or anxiety disorders that are contributing to their pain(Lim et.al, 2018).

For some individuals battling chronic pain, physicians treat patients with various forms of non-opioid medications. The type of medication used is specific to the type and severity of the pain experienced by the patient. In cases where the patient has a disease-specific pain, opioids are sometimes used. Opioids are needed more when chronic pain is severe enough to require pain management daily and when alternative treatment options are inadequate(Geneen et. al, 2017). Physicians tend to not prescribe patients opioids, particularly in recent years, due to people developing side effects and tolerance that limit long-term benefits from opioids(Geneen et. al, 2017). It is more used for short-term pain management for individuals. The CDC recommends sustained-release and long-acting opioids only when the patient is opioid-tolerant.

2.2 Mindfulness

The concept of mindfulness has gained attention in various fields, including basic emotional research, clinical science, and neuroscience. In the United States, one of the more extensively studied forms of mindfulness training is Mindfulness-Based Stress Reduction(MSBR). It is a structured program developed by Kabat-Zinn in 1990 that involves various mindfulness practices.

MSBR involves various practices including formal meditation as well as informal meditation that is included with the participant's daily life. The goal of these practices is to enhance one’s ability to observe the immediate content of experience, such as thoughts, emotions, memories, mental images, and physical sensations (Goldin, et al, 2014). One of the specific forms of meditation introduced in MSBR includes focused attention(focusing on an object in the present moment) and open monitoring(observing any experience in the present moment without a specific focus on an object). While MSBR does not aim to explicitly change the nature of thinking or emotional reactivity, it has been shown to decrease stress, depression, and anxiety symptoms and improve immune functioning and attention control.

2.3 Mindfulness as Treatment for Chronic Pain

2.3.1 How Mindfulness Affects Chronic Pain

According to a randomized controlled trial of patients with chronic low back pain who underwent mindfulness-based stress reduction (MBSR), mindfulness can reduce pain intensity, pain-related distress, and pain interference with daily activities (Hilton et al., 2017). Mindfulness can also decrease the negative emotions associated with chronic pain, such as fear, anger, and sadness (La Cour & Petersen, 2015). Furthermore, mindfulness can increase the awareness of the body and its sensations and help people detach from their pain and accept it as a part of their experience. Mindfulness can also induce long-lasting changes in the brain regions involved in pain modulation, such as the anterior cingulate cortex, the insula, and the prefrontal cortex (Hilton et al., 2017). By doing so, mindfulness can change the way people perceive and relate to their pain, and reduce its impact on their quality of life. Mindfulness does not eliminate the source of chronic pain, but it can enhance the coping skills and resilience of people who suffer from it.

2.3.2 Obstacles Using Mindfulness as a Treatment

One of the challenges is the stigma that some patients and healthcare providers may have towards non-pharmacological treatments. Some patients may feel that they are not being taken seriously or that their pain is not validated if they do not receive a prescription for pain medication. Some healthcare providers may also be reluctant to recommend or refer patients to mindfulness-based interventions, due to a lack of knowledge, training, or confidence in their effectiveness (Morone, 2019). This is especially problematic for mindfulness, as it requires active participation and engagement from the patients, as well as trust and rapport with the providers. Therefore, both patients and providers must be educated on the evidence and mechanisms of mindfulness for chronic pain, as well as the potential benefits and limitations of this approach. This way, patients can make informed choices about their pain management options, and providers can offer them a comprehensive and multidisciplinary care plan that includes mindfulness as a viable alternative or adjunct to pharmacological treatments.

Another challenge is that mindfulness may only be equally effective for some patients with chronic pain, as there are individual differences and contextual factors that may influence the response to the intervention. Several studies have attempted to identify the predictors and moderators of mindfulness-based pain relief, such as the severity and nature of pain, the presence of comorbid psychological disorders, the level of resilience, the attribution of pain causes, and the perception of injustice. This project aims to use machine learning to identify which patients will likely respond well to using mindfulness to treat chronic pain. This machine learning analysis can help healthcare providers to optimize their treatment plans and to enhance the effectiveness and acceptability of mindfulness for chronic pain management.

2.4 Related Work

2.4.1 Mind-Body Program Study

To delve into the methods and analysis of our research, we will first understand the backbones of our project. The first one is the study, “A Mind-Body Program for Older Adults With Chronic Low Back Pain - A Randomized Clinical Trial” which was conducted in 2019 in which participants with chronic lower bain over the age of 65 were analyzed on their pain levels from the beginning to the end of the program(Morone, 2019). All participants are above the age of 65 and have functional limitations based on the Roland and Morris Disability Questionaire and chronic pain with over 3 months of moderate intensity. The participants were split into two groups: the intervention group received the 8-week Mindfulness-Based Stress Reduction program and the control group received the “10 Keys to Healthy Aging”. The primary outcomes were improvement in function and reduction in pain.

During the prescreening, the researchers had some main reasons to include a person in the study and some exclusions to eliminate a person. Inclusions for the study were individuals over 65 years old, English speakers, with intact cognition, functional limitations, and self-reported moderate chronic pain levels. Exclusions for the study included previous participation in a mindfulness meditation program, serious underlying illness, non-ambulatory status, severely impaired mobility, visual or hearing impairments interfering with assessments, pain in other body parts more severe than chronic LBP, acute or terminal illness, or moderate to severe depressive symptoms. The study had 140 participants in the intervention group and 142 in the control group initially, with 132 and 138, respectively, completing the program. After 6 months, 118 participants remained in the intervention group and 135 in the control group.

The outcome measures included the Roland-Morris Disability Questionnaire (RMDQ) for functional limitations, the Numeric Pain Rating Scale (NRS) for pain assessment, the RAND-36 Health Status Inventory for quality of life, the Geriatric Depression Scale for depression, Chronic Pain Self-Efficacy Scale for self-efficacy, Coping Strategies Questionnaire for pain catastrophizing, and Mindful Attention Awareness Scale for self-reported mindfulness. Comorbidity data were assessed using the Cumulative Illness Rating Scale. After measuring pain and mobility at various time points throughout the study, analysis included comparing baseline characteristics of intervention and control groups using statistical tests. The intervention group, based on the Mindfulness-Based Stress Reduction program, included methods like body scan, sitting practice, walking meditation, and mindful stretching. Booster sessions for both groups were conducted monthly. Results showed that the mindfulness program improved short-term function and long-term current and most severe pain. However, the improvement in function was not sustained(Morone, 2019).

2.4.2 Machine Learning Methods

In this research project, we will utilize various machine learning methods to determine which attributes relating to a participant in the study will ultimately affect their ability to respond well to Mindfulness-Based Stress Reduction (MBSR). The application of machine learning in pain research is supported by existing literature, reflecting a growing trend in the field.

Machine learning, a subset of data science, is employed to detect patterns in data and make predictions, classify future data, observe potential subgroups, or extract information for deriving new knowledge (Lotsch et. al, 2018). Recent studies have demonstrated the efficacy of machine-learned technologies in general pain research, enabling the analysis and prediction of pain phenotypes from clinical data.

Supervised learning, a prominent approach in machine learning, involves symbolic and subsymbolic classifiers. Symbolic classifiers rely on domain experts interpreting conditions on features, while subsymbolic classifiers utilize machine-learning algorithms without a detailed understanding of their biomedical explanations (Lotsch et. al, 2018). Noteworthy supervised learning methods in pain research include Random Forest, Support Vector Machines, and K-nearest neighbors, all facilitating the exploration of datasets by reversing the analytical focus of classifier building and pattern detection.

One of the more crucial models for both machine learning methods as well as for feature selection strategies is the Random Forest model. This ensemble learning approach allows further understanding of the weight of various attributes on the outcome of the model. In the next section, we will review the advantages of using this approach and how we can incorporate it into our analysis of the dataset.

However, it is crucial to acknowledge the potential drawbacks of employing machine learning in pain research. Overfitting is a common concern, as the models may describe noise or irrelevant relationships rather than true patterns and correlations(Kernbach et. al, 2022). To mitigate this risk, the use of training, validation, and test datasets is essential to ensure the model's effectiveness on new data (Lotsch et. al, 2018). Additionally, the utility of these models is contingent on the availability of a substantial amount of data, emphasizing the importance of the successful enrollment of a large number of subjects in clinical studies for pain research.

In the context of our study on the changes in pain throughout an investigation into the effects of mindfulness on chronic pain, supervised machine learning methods offer a valuable opportunity. These methods can assist in identifying patterns in the dataset and pinpointing specific attributes that characterize a participant as a suitable candidate for MBSR.

3. Methodology

3.1 Data Description

The data for our study comes from a research study named "A Mind-Body Program for Older Adults With Chronic Low Back Pain - A Randomized Clinical Trial." This study aimed to track changes in pain and mobility among participants (Morone, 2019). Various pain scales, like the Numeric Pain Rating Scale, Roland and Morris Disability Questionnaire, and SF-36, were used to measure participants' pain levels at different points: before any treatment, at 8 weeks, 6 months, and one year. Data was collected for both the control and intervention groups.

Initially, there were 140 participants in the intervention group and 142 in the control group. By the 6-month mark, these numbers changed to 118 in the intervention group and 135 in the control group. The study covered 28 different measures for both groups, with each measure broken down into specific questions. The raw dataset includes separate files for each of these 28 scales, as listed in Table 2 below. This organized approach allows us to delve into the details of participants' pain and mobility experiences in a more detailed manner.

Data Scale	Acronym from Raw Dataset	Number of Attributes from Scale
Multidimensional Scale of Perceived Social Support	ss	12

SF-36 Physical Function Scale	sf	36
Chronic Pain Self Efficacy Scale	se	22
Roland and Morris Disability Questionnaire	RMDQ(rm)	26
Patient Global Impression of Change	pt	1
Profile of Mood States	pm	65
Physical Examination	pe	42
Pain Numeric Rating Scale (0-20)	pc	3
Short Physical Performance Battery	pb	2
National Adult Reading Test (NART)	na	61
Mindful Attention Awareness Scale (5 item)	msc	5
Mindful Attention Awareness Scale (15 item)	ms	15
Multidimensional Pain Inventory	mp	8
Mini-Mental Status Exam	mm	25
McGill Pain Questionnaire	mc	16
Modifiable Activity Questionnaire	ma	181
Informed Consent	ic	3
Health System Encounters	hs	32

Geriatric Depression Scale	GDS(gt)	31
Generalized Anxiety Disorder-7	ge	12
Five Facet Mindfulness Questionnaire	ff	39
Fear Avoidance and Beliefs Questionnaire-Scale 2	fa	4
Demographics and Other Factors	dm	17
CSQ Catastrophizing Scale	cs	6
Chronic Pain Acceptance Questionnaire	cp	20
Cumulative Illness Rating Scale	ci	14
Credibility Expectancy Questionnaire (CEQ)	ce	6
CAMCI	cc	1

Table 2 Various Scales from Raw Dataset

3.2 Data Preprocessing

The original dataset, stemming from the comprehensive paper on the impact of Mindfulness on Chronic Lower Back Pain, was mostly robust due to its organization of pain scales into smaller datasets. Each pain scale, as detailed in the table above, was segregated, accompanied by a dictionary page within the dataset file explaining each metric. While this structure provided a systematic overview of pain assessments, it necessitated further data preprocessing to merge these datasets into a unified dataset to implement analytical models.

Our data integration process was executed in Python, leveraging libraries such as NumPy and Pandas. The combination of these datasets occurred based on the participants' unique identification numbers, ensuring the consolidation of diverse pain assessments into a cohesive dataset. Furthermore, each smaller dataset reflected data collected at distinct time points: the study's initial time point, 8 weeks, 6 months, and 12 months. Consequently, the dataset featured individual rows corresponding to each time point for each participant, measuring diverse pain and mobility measures. A categorical marker denoting group allocation (control or intervention) was also incorporated.

The utilization of Python enabled the creation of separate data frames stratified by time points and group allocations. For instance, distinct data frames were crafted to show initial time point measurements with subsequent time points for both control and intervention groups. Following data frame creation, a pivotal step involved computing a target variable, representing the difference in pain measurements from a specific time point to the study's initiation. This calculation, exemplified by the Roland and Morris Disability Questionnaire (RMDQ) total score, provided a quantitative metric for pain variations throughout the study.

In the preliminary analysis, emphasis was placed on evaluating pain differences between the study's initiation and the 8-week and 6-month time points. The RMDQ total score served as the target variable, although alternative pain measurements could be substituted for further exploration. The purpose behind creating these structured data frames with well-defined target variables was to deploy them as training data for model development. This approach aimed not only to determine significant differences in pain across dimensions but also to understand attributes having the greatest influence on model outcomes.

3.3 Modeling

There are numerous models we will be implementing to find which one has the highest predictive power and can show whether or not there were significant changes in pain levels for participants in the study. We will use both Random Forest Classifier and XGBoost as a machine learning method as well as a feature selection strategy. Additionally, we will use Decision Trees to model to determine which attributes of a patient's health can help predict their response to treatment. Other machine learning methods such as linear and logistic regression will also be implemented to determine if they have stronger abilities to show the patterns in our dataset. By the end of our modeling strategies, the ultimate intention is to find attributes of a participant that make them more likely to respond well to mindfulness as a treatment for chronic pain.

3.3.1 Decision Trees

Decision Trees are a machine learning method that makes decisions by splitting data based on various features, creating a tree-like structure. Each node within the tree represents the exact split of the data and each branch shows the outcome of that particular split. The end node, or leaf node, has a class label that could be binary(for classification) or numerical(for regression). They are highly interpretable as well as good for both classification and regression tasks. This model works by carefully pruning the data into various subsets and determining, for each level of the tree, the most effective way to separate the different classes based on measures such as the Gini index, entropy, or variance.

In 2014, there was a paper published, "The clinical decision analysis using decision tree", which attempted to use the decision tree model to be able to create a tool that would help doctors make clinical decisions in terms of maximizing effectiveness and minimizing harm. The tool was called, CDA or “clinical decision analysis” and the aim was to overcome uncertainty and help have more confidence in making medical decisions. The study concluded that this tool was found to be fairly effective in determining “environment variables that need to be considered when deciding on a clinical setting”(Bae, 2014). This relates to our goals in this project of determining indicators of a patient’s health that are associated with them responding well to treatment.

3.3.2 Random Forest

The Random Forest technique is a widely used machine learning algorithm that is a combination of the output of multiple decision trees to reach a single result. It leverages randomness in both the data and feature selection to improve predictive accuracy and generalization. It has proven to be adept in both classification and regression problems making it an ideal candidate for our analysis in which we are aiming to have a model that can either have results in terms of classification and regression. Random Forest is considered an ensemble learning method because it takes the predictions of multiple individual models to make more accurate predictions. Additionally, the Random Forest Classifier itself can be used as a feature selection strategy. It can pick out attributes of a dataset that are most important to the model output and evaluation. This can also be helpful for our analysis because it can allow us to pick out the pain measures that are most important to determining how our target variable is changing.

A paper from 2010, “The American College of Rheumatology Preliminary Diagnostic Criteria for Fibromyalgia and Measurement of Symptom Severity” used the Random Forest model to determine the development of fibromyalgia, a chronic disorder characterized by widespread pain, by developing criteria for diagnosing a person with the disease as well as constructing a symptom severity scale(Wolfe et. al, 2010). Both the motivation behind the study as well as the methods in the research aligns with our goals in our study. Given a dataset including 829 previously diagnosed fibromyalgia patients with data on their widespread pain index(WPI), researchers used the random forest algorithm initially to determine variable importance. Similarly, we will be utilizing the random forest classifier to determine the pain measures included in the raw dataset from the study on chronic lower back pain to determine what makes a particular person likely to respond well to treatment and inversely, what may make a person not respond well to treatment. The second way we will be implementing the Random Forest classifier is to create a model that can predict whether or not there is a significant change in pain levels from the initial time point of the study to any other time point. The study on the diagnostic criteria for fibromyalgia was done to essentially recommend a new case definition of the disease(Wolfe et. al, 2010). This was through an analysis of patient pain levels and other diagnostic variables. While in our research we are not trying to redefine chronic pain, this study supports our goal to determine whether mindfulness is an effective remedy for chronic pain and if so, what characteristics of a patient's health make them a good candidate for treatment

3.3.3 XGBoost

Extreme Gradient Boosting (XGBoost) is a powerful machine learning method for predictive modeling. Like Random Forest, it takes the predictions of weaker models like decision trees to then create a stronger predictive model making it an ensemble learning method. The reason XGBoost tends to have strong predictive power is because it takes in a range of techniques allowing it to process and learn from more than one algorithm therefore increasing its speed and accuracy. We will be using XGBoost in a similar way to Random Forest- to model the changes in pain over time in the study.

A previous study from 2022, “Machine learning versus logistic regression for prognostic modeling in individuals with non-specific neck pain” discussed the overall accuracy of the XGBoost model in predicting pain in patients (Liew et. al, 2022). It found that in comparison to linear and logistic regression as well as various other machine learning methods, XGBoost had the best predictive power. This study is similar to our research in that it was trying to determine if there was “clinically meaningful improvement” in pain over 3 months (Liew, 2022). Numerous different attributes of patient health were included in the model training. After testing models like K-Nearest Neighbors, Logistic Regression, Lasso Regression, and an Artificial Neural Network, the XGBoost proved to have the best overall accuracy for determining if there would be a difference in pain for a patient. Due to the parallels in this study and the research we are doing, it will be valuable for us to see if we can get similar results by implementing this algorithm to our dataset as well.

3.3.4 Model Hyperparameter Tuning

After implementing these models, to get the highest predictive power possible, we will be using grid search to tune the hyperparameters. It is cited in the literature that grid search can improve model performance significantly. It runs through various combinations of hyperparameter values for a particular model to determine the best fit for the dataset. One study, "Deep Learning and Machine Learning with Grid Search to Predict Later Occurrence of Breast Cancer Metastasis Using Clinical Data" found that grid search improved almost all of their machine learning models. In this study, researchers were trying to find the best model to be able to predict whether or not a patient will have a later occurrence of metastasis. The models that improved in accuracy with grid search were: XGBoost, Random Forest, and the Support Vector Machine. It will be useful for us to implement grid search into the fitting of our models so that we can determine the hyperparameters that allow us to be able to predict changes in pain for a particular participant with the best accuracy. It is important to note that the computational time required for grid search is significantly longer than using the vanilla approach to the machine learning model and not tuning the hyperparameters.

3.4 Evaluation

As discussed in the data preprocessing section, our target variable is the change in pain for the Roland and Morris Disability questionnaire. The evaluation of our models will look into understanding whether or not the change in pain is significant or not from the initial time point to either the 8-week time point or the 6-month time point. After training our models using Random Forest and XGBoost, we will evaluate the accuracy both in terms of regression and classification. When evaluating our models in terms of classification, essentially whether or not there was a significant change in pain, we will use metrics like accuracy, precision, recall, and F1 score. These metrics allow us to understand how well the model can make correct predictions on our dataset. For the evaluation of our regression models, we will use metrics such as Mean Squared Error, R-squared, Mean Average Error, and Root Mean Squared Error. These metrics will help us determine whether our models fit the dataset well and whether they can accurately make predictions for new data.

4. Results

4.1 Datasets

Our research utilizes two distinct datasets: the “sum” dataset and the “all” dataset. (add distinction). Each of these datasets has been subdivided into six subsets based on baseline data. These subsets are further organized according to their respective target values, identified at subsequent time points.

Specifically, the subsets are categorized based on data collected at three different time intervals: 8 weeks, 6 months, and 12 months. Within each time interval, the data is separated into two groups: the intervention group and the control group. This division results in six unique datasets for each primary dataset.

Furthermore, our study investigates two different target values, leading to the creation of two separate sets of six datasets for each primary dataset. These are labeled as "ptchange" and "rmchange," with the latter being derived from the rmscore variable. Consequently, this approach yields a total of 12 distinct dataframes per primary dataset, culminating in an aggregate of 24 dataframes currently under analysis. In the following section, we will review how we use these datasets to do feature importance analysis to determine attributes that most have an impact on a participant's ability to respond to treatment.

4.2 Understanding the Outcome Measures

In the discussion above, we highlighted two key outcomes we are focusing on for our machine learning analysis: rmchange and ptchange. Rmchange comes from the rmscore, which is a way to measure how much difficulty a participant has doing everyday activities. The score ranges from 0 to 24, where a higher score means more difficulty. Rmchange is the change in this score from the beginning of the study to another time point in the study. So, if rmchange is positive, it means the patient's ability to do daily activities has gotten better throughout the study.

Ptchange is about how participants feel their pain has changed and is not measured at the start of the study. It is based on the global impression of pain score, which ranges from 1 to 7. A score of 1 means the participant feels their pain has improved, a score of 4 means they don't see any change, and a score of 7 means they feel their pain has gotten worse. Here, a lower score is better because it means the participant feels their pain has decreased since the study began.

4.3 Feature Importance

The Random Forest and XGBoost Classifiers play a pivotal role in our analysis by identifying the most critical factors from our dataset that influence the effectiveness of mindfulness in alleviating symptoms of chronic pain for participants. This process allows us to predict which individuals are more likely to benefit from this form of treatment. Our datasets facilitate a comprehensive feature importance analysis, offering insights into which participant attributes significantly impact treatment success. It is important to note that the determinants of treatment efficacy may differ based on the pain scale under examination and the specific time point measured from the baseline. For instance, the influence of certain variables on the outcome measures—ptchange and rmchange—might not be consistent. Additionally, the impact of these variables can change when comparing short-term outcomes (from baseline to 8 weeks) to longer-term outcomes (from baseline to 6 months or 12 months).

We have initially focused on analyzing the significance of various features for ptchange and rmchange over an 8-week period. This analysis reveals that the variables crucial for the model's predictions vary slightly between targeting ptchange and rmchange as the outcome. The findings, presented through bar charts and tables, rank the top 10 features according to their importance to the model, with 1 being the most critical.

4.3.1 Random Forest Feature Importance for ptchange

When reviewing our results from running the Random Forest model it is important to note that the features are all measures from the baseline timepoint and that ptchange is measured at the 8-week timepoint. Our goal is to understand how measures taken at baseline affect a participant’s impression of pain after treatment is completed at 8 weeks. The bar chart in Figure 1 below shows the top 10 features with the most importance on the random forest model's prediction of ptchange at 8 weeks.

Participant Global Impression of Change (ptchange)

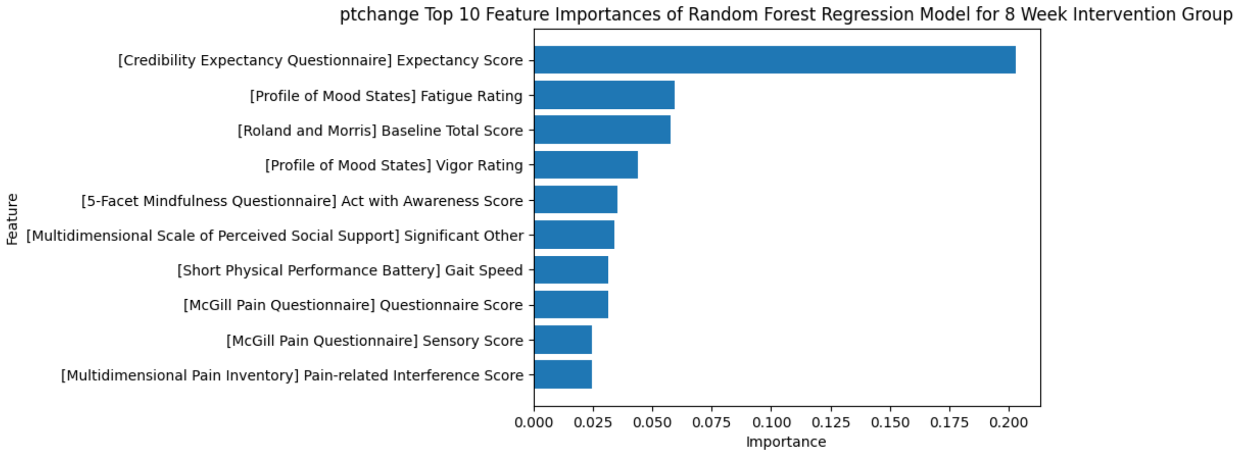

Figure 1 Bar Chart for Top 10 Feature Importances using Random Forest for predicting ptchange

The table below, highlights the top 10 features important to the Random Forest model in predicting ptchange. By understanding features important to our analysis we can get closer to predicting what attributes and scores of a participant at the baseline measurement will result in them either responding or not responding to treatment.

Participant Global Impression of Change (ptchange)		
Rank	**Description**	**Scoring**
1	[Credibility Expectancy Questionnaire] Expectancy Score	Ranks patients expectation of results from treatment
2	[Profile of Mood States] Fatigue Rating	Higher score indicates worse feelings of fatigue
3	[Roland and Morris Pain Scale]	Higher score indicated worse functional limitation
4	[Profile of Mood States] Vigor Rating	Higher score indicates worse feelings of vigor
5	[Fear Avoidance and Beliefs Questionnaire] Total Score	Higher score is worse fear avoidance behaviors
6	[Multidimensional Scale of Perceived Social Support] Significant Other	Higher score indicates higher perceived social support from significant other
7	[Short Physical Performance Battery] Gait Speed	Higher score is a faster walking speed
8	[McGill Pain Questionnaire] Questionnaire Score	Higher score is worse pain
9	[McGill Pain Questionnaire] Sensory Score	Higher score is worse pain
10	[Multidimensional Pain Inventory] Pain-related inference score	Higher score is worse pain

Table 3 Top 10 Feature Importances using Random Forest for predicting ptchange

The scatterplots below show the relationship between a particular feature from the chart above that had more of an importance on the model and ptchange. From the 10 features we only focus on a few scatterplots that show some sort of correlation and are within the top 5 five most important features.

Our first scatterplot, shown in Figure 2, indicates the relationship between a scale on a participant's vigor and ptchange. It should be noted that a higher score in the Profile of Mood States for Vigor indicates higher feelings of vigor. Since vigor is a more positive emotion, higher scores also mean the participant felt more energetic and enthusiastic. There appears to be a correlation here with participants with higher feelings of vigor also tending to have a lower ptchange score meaning they feel their pain improved after treatment.

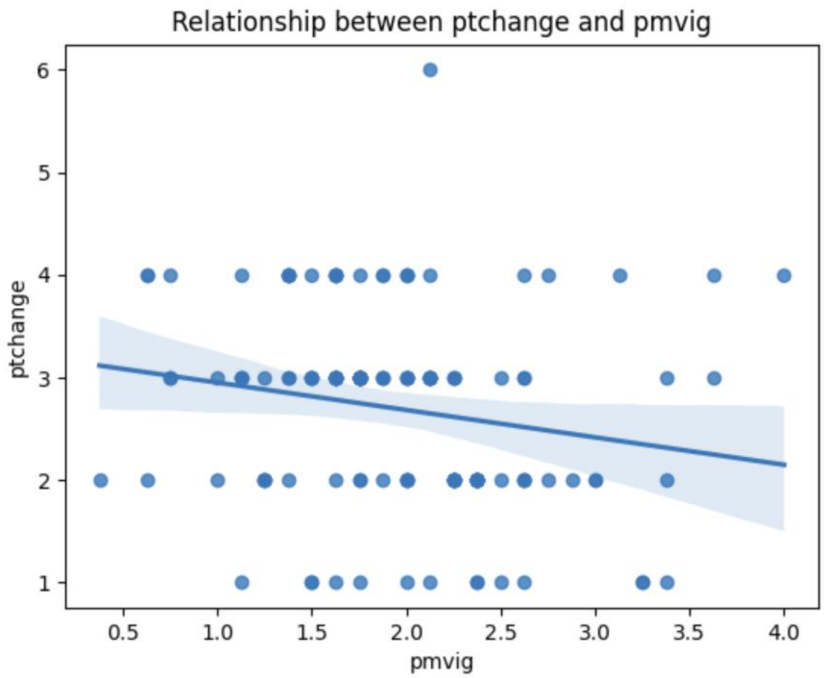

Figure 2 Higher Vigor Rating at baseline is associated with better participant global impression of change in pain at 8 weeks

The next plot, shown in Figure 3, indicates the relationship between a participant's rmscore and ptchange. It should be noted that a higher rmscore indicates worse overall functional limitation. There appears to be a correlation between these variables with participants with worse functional limitation at the baseline also tending to report a lower ptchange score after treatment meaning their pain improved after 8 weeks.

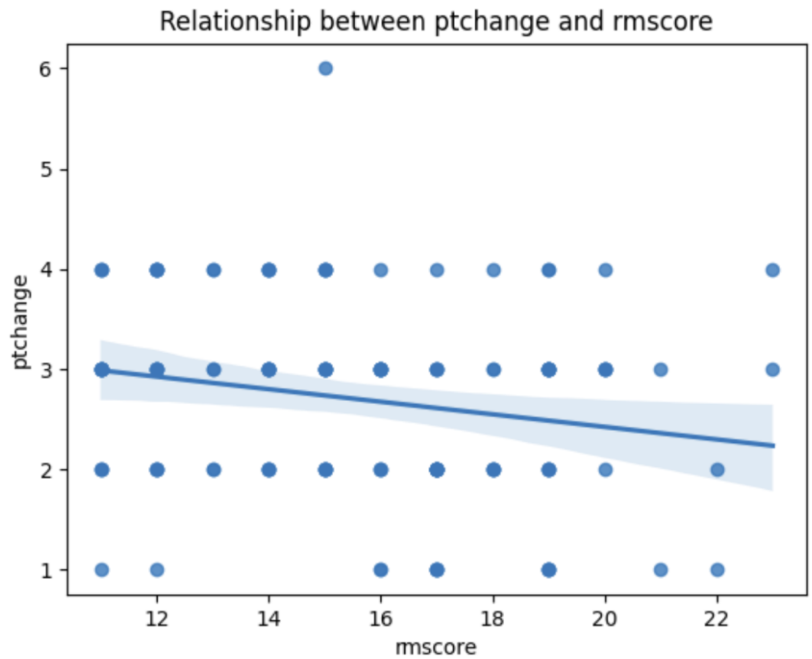

Figure 3 Higher Roland and Morris Total Score at baseline is associated with better participant global impression of change in pain at 8 weeks

4.3.2 Random Forest Feature Importance for rmchange

The bar chart in Figure 4 below shows the top 10 features with the most importance on the random forest model's prediction of rmchange at 8 weeks. When reviewing our results from running the Random Forest model it is important to note that the features are all measures from the baseline timepoint and that rmchange is measured as the difference in rmscore from the 8-week timepoint to the baseline timepoint. Our goal is to understand how measures taken at baseline affect a participant's impression of pain after treatment is completed at 8 weeks.

Roland and Morris Pain Scale Change (rmchange)

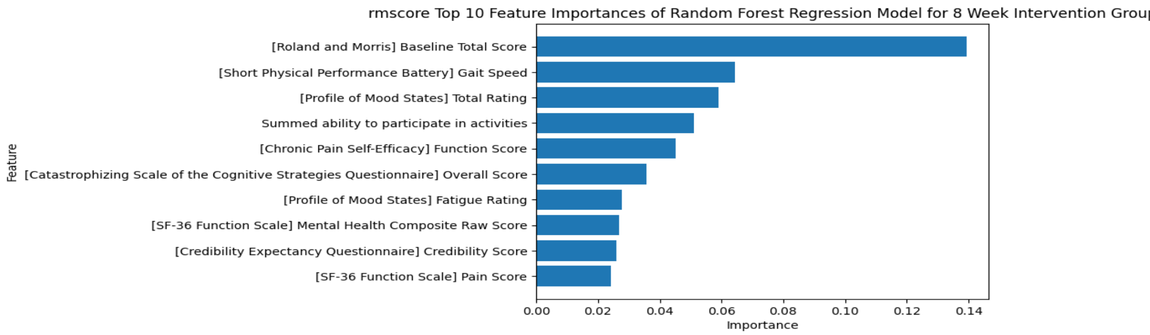

Figure 4 Bar Chart for Top 10 Feature Importances using Random Forest for predicting rmchange

The table below highlights the top 10 features important to the Random Forest model in predicting rmchange.

Roland and Morris Pain Scale Change (rmchange)		
Rank	**Description**	**Scoring**
1	[Roland and Morris Pain Scale] Total Score	Higher score indicated worse functional limitation
2	[Short Physical Performance Battery] Gait Speed	Higher score is the faster speed for walking
3	[Profile of Mood States] Total Rating	Higher score indicates overall less stable mood
4	Summed ability to participate in activities	
5	[Chronic Pain Self-Efficacy] Function Score	
6	[Catastrophizing Scale of the Cognitive Strategies Questionnaire] Overall Score	Higher scores reflect higher levels of catastrophic thoughts
7	[Profile of Mood States] Fatigue Rating	Higher score indicates worse feelings of fatigue
8	[SF-36 Function Scale] Mental Health Composite Raw Score	Higher score indicates better overall mental health
9	[Credibility Expectancy Questionnaire] Credibility Score	
10	[SF-36 Function Scale] Pain Score	Lower score is worse pain

Table 4 Top 10 Feature Importances using Random Forest for predicting rmchange

The plot below for rmchange, shown in Figure 5, indicates the relationship between a participant's rmscore and rmchange. It should be noted that a higher rmscore indicates worse functional limitation. There appears to be a correlation between these variables with participants with worse overall functional limitation at the baseline measurement also tending to result in a higher rmchange meaning that their pain seemed to improve.

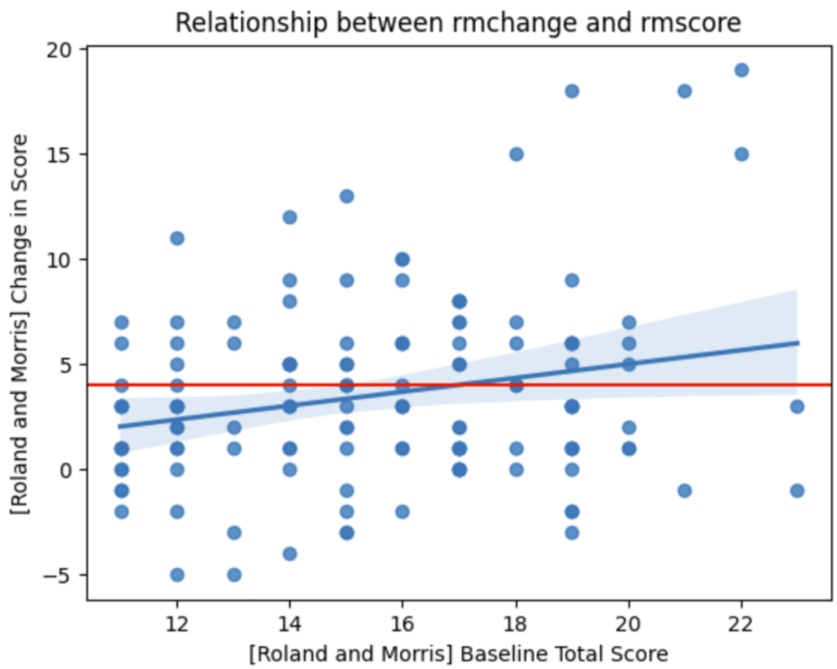

Figure 5 A higher Roland and Morris Total Score at baseline is associated with less functional limitation at 8 weeks

4.3.3 XGBoost Feature Importance for ptchange

When reviewing our results from running the XGBoost model it is important to note that the features are all measures from the baseline timepoint and that ptchange is measured at the 8-week timepoint. Our goal is to understand how measures taken at baseline affect a participant's impression of pain after treatment is completed at 8 weeks. The bar chart in Figure 6 below shows the top 10 features with the most importance on the XGBoost model's prediction of ptchange at 8 weeks.

Participant Global Impression of Change (ptchange)

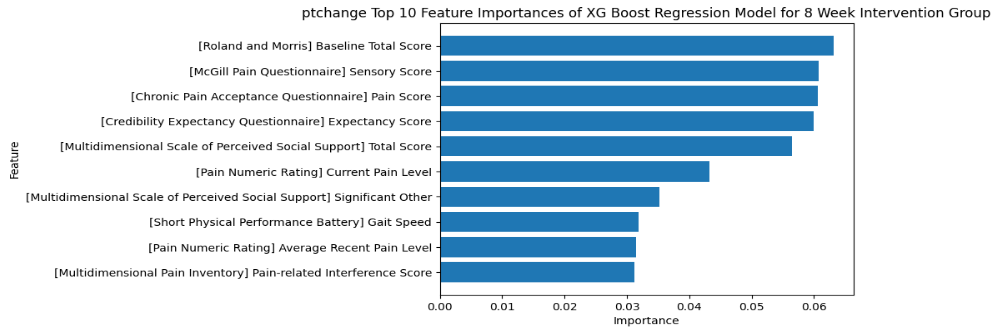

Figure 6 Bar Chart for Top 10 Feature Importances using XGBoost for predicting ptchange

The table highlights the top 10 features important to the XGBoost model in predicting ptchange.

Participant Global Impression of Change (ptchange)		
Rank	**Description**	**Scoring**
1	[Roland and Morris Pain Scale] Baseline Total Score	Higher score indicated worse functional limitation
2	[McGill Pain Questionnaire] Sensory Score	Higher score is worse sensory issues
3	[Chronic Pain Self-Efficacy] Pain Score	
4	[Credibility Expectancy Questionnaire] Expectancy Score	Ranks patient's expectation of results from treatment
5	[Multidimensional Scale of Perceived Social Support] Total Score	Higher score indicates higher overall perceived social support
6	[Pain Numeric Rating] Current Pain Level	Higher score indicates worse pain at the exact time of the question being asked
7	[Multidimensional Scale of Perceived Social Support] Significant Other	Higher score indicates higher overall perceived social support from significant other
8	[Short Physical Performance Battery] Gait Speed	Higher score is the faster speed for walking
9	[Pain Numeric Rating] Average Recent Pain Level	Higher score indicates worse pain on average recently
10	[Multidimensional Pain Inventory] Pain-related inference score	Higher score is worse pain

Table 5 Top 10 Feature Importances using XGBoost for predicting ptchange

4.3.4 XGBoost Feature Importance for rmchange

The bar chart, in Figure 7, below shows the top 10 features with the most importance on the XGBoost model's prediction of rmchange at 8 weeks. When reviewing our results from running the XGBoost model it is important to note that the features are all measures from the baseline timepoint and that rmchange is measured as the difference in rmscore from the 8-week timepoint to the baseline timepoint. Our goal is to understand how measures taken at baseline affect a participant's impression of pain after treatment is completed at 8 weeks.

Roland and Morris Pain Scale Change (rmchange)

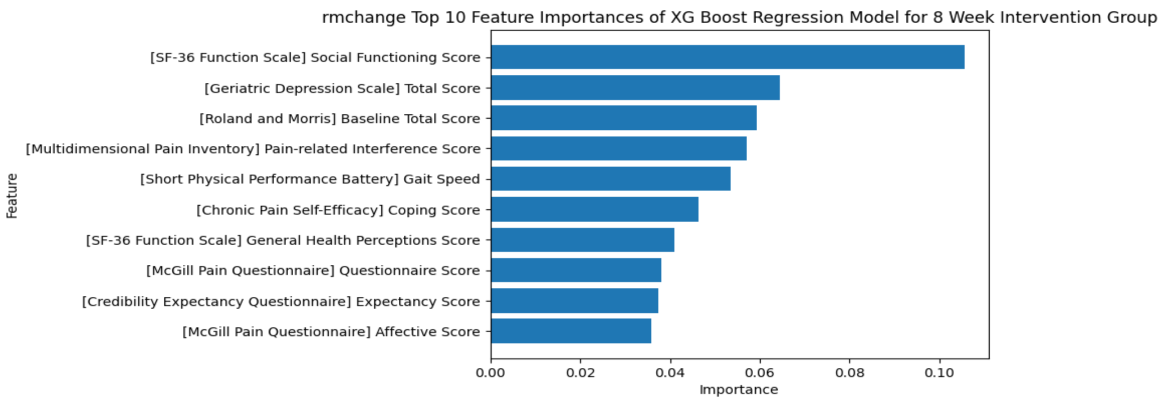

Figure 7 Bar Chart for Top 10 Feature Importances using XGBoost for predicting rmchange

The table below highlights the top 10 features important to the XGBoost model in predicting rmchange.

Roland and Morris Pain Scale Change (rmchange)		
Rank	**Description**	**Scoring**
1	[SF-36 Function Scale] Total Score	
2	[Geriatric Depression Scale] Total Score	A score of 5 or higher suggests depression
3	[Roland and Morris Pain Scale] Baseline Total Score	Higher score indicated worse functional limitation
4	[Multidimensional Pain Inventory] Pain Related Inference Score	Higher score indicates higher overall perceived social support
5	[Short Physical Performance Battery] Gait Speed	Higher score is the faster speed for walking
6	[Chronic Pain Self-Efficacy] Coping Score	
7	[SF-36 Function Scale] General Health Perceptions Score	Higher score indicates better general health perceptions
8	[McGill Pain Questionnaire] Questionnaire Score	Higher score is worse pain score
9	[Credibility Expectancy Questionnaire] Expectancy Score	Ranks patients expectation of results from treatment
10	[McGill Pain Questionnaire] Affective Score	

Table 6 Top 10 Feature Importances using XGBoost for predicting rmchange

4.3.5 Decision Tree Analysis

The previous analysis focused on the one-dimensional analysis of the features involved in our dataset. Although this is helpful in terms of being able to understand if there are correlations between variables it is even more useful to be able to look at how numerous variables together can affect the outcome of a model. One way we can do this is by running a Decision Tree classifier. Instead of having many different trees that are built in Random Forest and XGBoost, the Decision Tree will only build one tree that classifies all data in the model. For our analysis, we built two decision trees: one that is for regression and one for classification. The decision tree for classification has the goal of predicting that either rmchange is below a score of 4, or greater than or equal to 4, this tree is shown in Figure 8. When rmchange is greater than or equal to 4 it means there was a medically significant improvement in physical functioning. The decision tree for regression attempts to predict the rmchange based on other features in the dataset; this tree is shown in Figure 10. It should be noted that the decision trees shown are slightly condensed from the original tree that was outputted because we pruned the tree to remove leaf nodes of the tree that were repetitive or insignificant to our data.

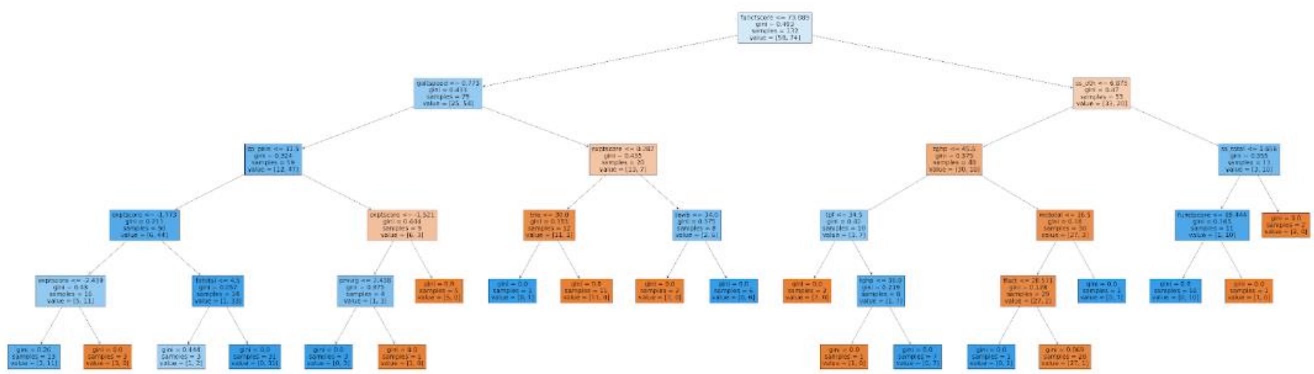

Figure 8 Decision Tree for classification analysis on rmchange

In Figure 8 above its important to note that all leaf nodes in orange represent that rmchange is greater than or equal to 4 and blue represents that rmchange was less than 4. As we recurse down a branch of this decision tree each branch contains a number of samples that have been classified due to splits within that particular branch. Of these samples, we determine a number of them in which the result is considered medically significant(rmchange is greater than or equal to 4) and a number that are considered medically insignificant(rmchange is less than 4).

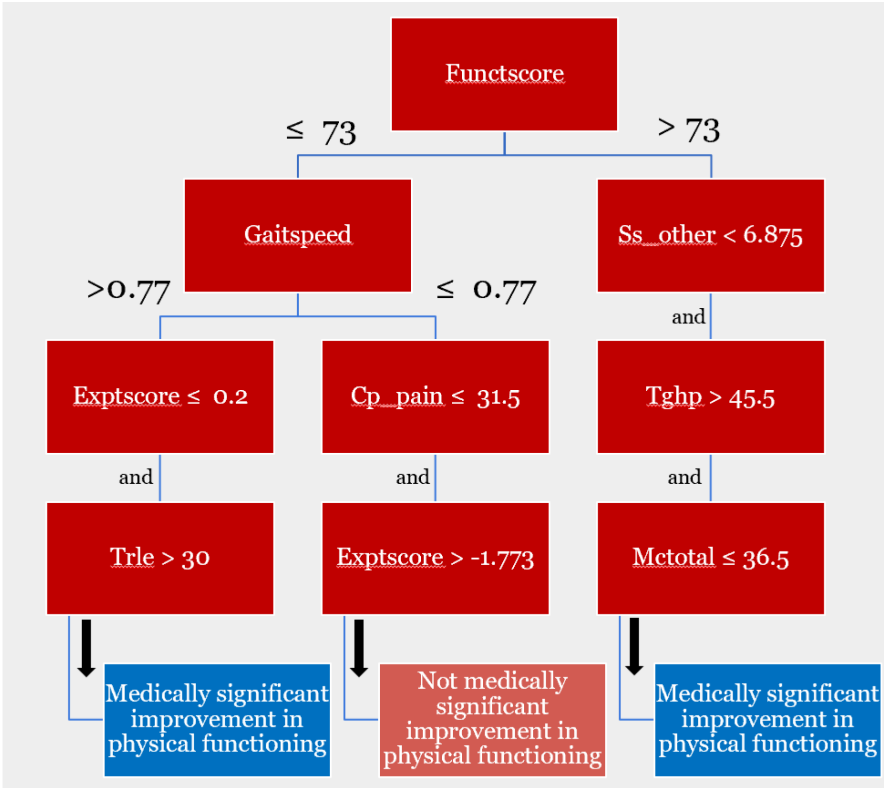

Figure 9 condensed Decision Tree for classification analysis on rmchange

In the following section we will review three branches which have the highest number of samples. In Figure 9, we focus on those three branches so that the nodes and results are more clear. The branches all result in an end node that represents either an outcome of having a medically significant improvement in physical functioning or not medically significant. All of these branches for classification begin with using functscore which comes from the Chronic Pain Self-Efficacy Scale.

The branch with the highest number of samples in terms of containing samples that are medically insignificant has the following splits in order from top of the tree to bottom of the tree:

functscore ≤ 73.889	gaitspeed≤ 0.773	cp_pain ≤ 31.5	exptscore > -1.773

With these splits it resulted in 33 medically insignificant samples and 1 medically significant outlier. From here we split on the participants walking speed(gaitspeed). In this case, it represents that when a patient’s walking speed is slower at the beginning of treatment they tend to not have a significant improvement in their physical functioning after treatment.

On the other hand, the following branch had a high number of samples that were medically significant:

functscore ≤73.889	gaitspeed > 0.773	exptscore ≤0.287	trle > 30

With these splits, it resulted in exactly 11 medically significant samples. This means that with a faster walking speed the participant tends to have a medically significant improvement in their physical functioning after treatment. This aligns with the previsous analyis for the split before where we found that slower walking speed resulted in not significant improvement in physical functioning.

Another branch with a high number of medically significant samples has the following splits:

functscore > 73.889	ss_oth ≤ 6.875	tghp > 45.5	mctotal ≤ 36.5

This split had 27 medically significant samples and 2 medically insignificant samples. It has an even higher number of samples included in comparison to the previous branch but it does also have a few medically insignificant samples due to it being combined with another branch in the tree. In this branch we focus in on looking at the participant's social support(ss_oth). It appears as if with less social support at the beginning of treatment from a patient's significant other, it resulted in a medically significant improvement in physical functioning after 8 weeks of the mindfulness treatment. Another split in this branch is the mctotal which is the McGill Pain Questionnaire Total Score where a higher score is worse pain. In this branch, with less pain before treatment, the patient had more improvement in their physical functioning after treatment.

After seeing the three branches of the decision tree with a high number of samples, it is clear that gaitspeed is an important attribute in predicting medical significance or insignificance as it showed up in two of the branches. In addition to doing the classification analysis we also did a decision tree for regression analysis where instead of predicting whether rmchange was above or below our threshold value, we were attempting to predict the actual value for rmchange. The problem with this type of analysis is that is more difficult to accurately predict values for rmchange and have a large number of samples classified in one branch. In Figure 10 below we visualize the decision tree model.

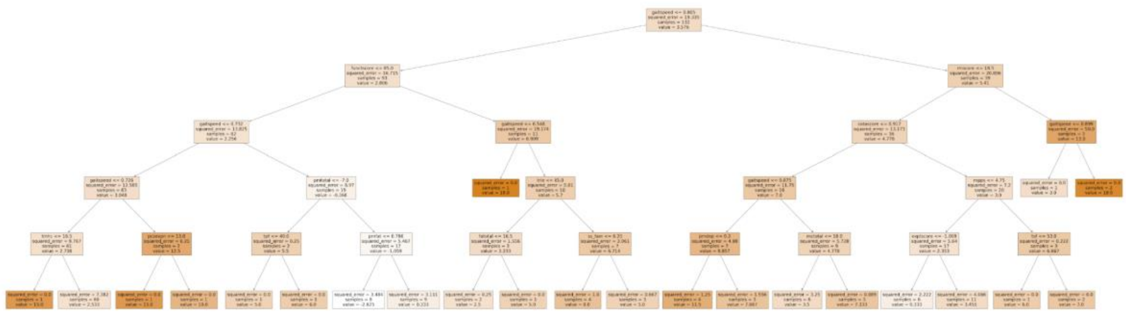

Figure 10 Decision Tree for regression analysis on rmchange

To understand the results of the decision tree above which does regression analysis on the outcome measure rmchange, we will review two branches of the tree with a high number of samples, a significant rmchange and a low overall squared error. Both branches that are shown below, result in an rmchange which is above 4(medically significant) meaning that either of these ways of splitting the data will result in samples that are labeled as being a significant positive change in the rmscore. The splits of these two branches are shown in Figure 11 below.

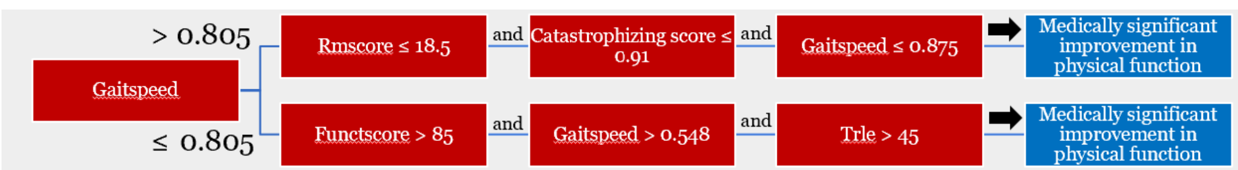

Figure 11 branches from Decision Tree for regression analysis on rmchange with high number of samples

The above two branches use gaitspeed as the root node for the split. Additionally, rmscore is included in the analysis where if rmscore is lower, less functional limitation at baseline, it resulted in a medically significant improvement in physical functioning after treatment.

In the analysis of the decision tree from both classification and regression there are a few common trends. The first being that gaitspeed(walking speed), came up numerous times and that in both the classification and regression analysis. When a participant had a faster walking speed before treatment, they had a medically significant improvement in their physical functioning after 8 weeks of treatment. The other common trend in both the analyses was that with less pain before treatment or with less functional limitation before treatment, the patient had more significant improvment in their functioning after treatment.

4.3.6 XGBoost Analysis

In addition to the Decision Tree analysis we also ran the XGBoost model on the same data for both classification and regression. XGBoost creates many small trees and runs each datapoint through an individual tree to a final leaf node, and adds the result of the leaf node to the overall sum. For classification, a result of 0 is considered medically significant, and 1 is medically insignificant. A negative leaf value leans towards a medically significant classification, and a positive value leans towards a medically insignificant classification. Every individual node has the gain (contribution of the split to the model) and the cover (how many datapoints were affected by a given split). We are interested in looking at branches with later nodes with high gain and cover, as well as a strong negative leaf value(indicating a medically significant classification).

For regression, we had a few branches which appeared to have both a high gain and cover as well as medically significant improvement in physical functioning. These branches are shown in the table below where each row represents a branch in the tree.

gaitspeed<0.733494699	rmscore≥22				
gaitspeed<0.733494699	rmscore<22	copescore≥65	rmscore≥16	tewb<45	
ss_oth<6.75	cp_pain ≥22	tghp<59	tpf<49	tpa≥30	gaitspeed≥0.469664425

Table 7 Branches of XGBoost with medically significant results for regression

The first two branches split on gaitspeed, with a slower walking speed resulting in a medically significant improvement in physical functioning. There is also a split on rmscore, where more functional limitation at baseline resulting in a medically significant improvement(for this first branch. In the second branch, we keep gaitspeed as being slower, but we focus on rmscore being between the values of 16 and 22. The final important branch of the tree has the first split on social support from significant other being less that 6.75 indicating the participant had less social support at baseline. Additionally this branch ended with a split on gaitspeed being higher(faster walking speed). This branch aligns more with our results from the decision tree with a faster walking speed at baseline and lower social support from significant other resulting in a medically significant improvement in physical functioning. The first two branches have almost the opposite results as the decision tree where these ones indicating that gaitspeed being lower(slower walking speed) resulting in medically significant improvement.

The table below highlights two branches from the classification of rmchange for the XGBoost model that are significant due to the fact that they have overall high gain and cover as well as resulted in a medically significant improvement in physical functioning.

gaitspeed≥0.7783162	mppi<5	ss_oth<7	gaitspeed<1.15622389	tpf≥36	pmcon<1.85714281
gaitspeed≥0.7783162	mcsens≥7	mcsens<19	tpf≥31	rmscore≥13	

Table 8 Branches of XGBoost with medically significant results for classification

The results for classification were more similar to the results we got in the decision tree analysis in that the first split was for gaitspeed being greater than 0.78 meaning that with the participant walking faster, they had more significant improvement in physical functioning. Additionally, the first branch included the social support attribute indicating that less social support from significant other at baseline also resulted in medically significant results. In the second branch, rmscore being greater than 13, meaning more functional limitation, resulted in medically significant results. While our analysis from XGBoost is not exactly the same as the Decision tree, there are some valuable results that we can draw from this analysis in an effort to understand what attributes of a patient’s health we can review to know if a patient will respond well to mindfulness treatment.

5. Conclusion

In conclusion, our study has allowed us to understand the intricate relationships between various characteristics and the efficacy of mindfulness interventions in alleviating symptoms of chronic pain. By analyzing our datasets, each segmented into subsets based on time intervals and target values, we have gathered significant insights into the predictive power of various features concerning treatment outcomes. Our utilization of machine learning techniques, including Decision Trees, Random Forests, XGBoost, has allowed us to pinpoint attributes that most significantly influence participant responses to treatment.

Our analyses across different models have consistently highlighted the importance of baseline measures of pain, functional limitation, mood states, social support, and walking speed in predicting the effectiveness of mindfulness interventions. For instance, lower pain at baseline and faster walking speed have been associated with better responses to treatment. Additionally, certain baseline characteristics, such as higher social support and worse functional limitations, tend to predict less improvement. While we were able to pinpoint some attributes of a patient's health at baseline that can help in predicting their response to treatment, some of these results contradicted each other in the analysis for various models. For example, in the Decision Tree model, we were able to determine that a faster walking speed at baseline resulted in medically significant results but two of the branches for XGBoost for regression had the opposite results with a slower walking speed at baseline resulting in a significant change in physical functioning.

Future work should be done to train, test, and tune these machine learning models to reach stronger conclusions on whether these predictors truly indicate that a patient will or will not have an improvement in their physical functioning after treatment. To get stronger accuracy on the models, more hyperparameter tuning can be done to determine the parameter values that will result in the best built-out tree. Although we did some work to prune the trees and condense the branches, it would be worthwhile to continue in this work to try and have even smaller trees to be able to focus on various aspects of a patient's health that are key indicators of them responding to treatment.

Our work, however, highlights the importance of considering multiple factors in predicting the success of pain management strategies, offering a more holistic understanding of patient care. In our models, when we considered numerous attributes at the same time we were able to more accurately classify whether a patient would respond well to treatment. With continued research and model testing, this type of modeling has the potential to confidently predict whether a patient will have a medically significant response to mindfulness for treating their chronic pain.

References

Bae, J.-M. (2014). Clinical decision analysis using decision tree. Epidemiology and Health. https://doi.org/10.4178/epih/e2014025

Breitborde, N. J. K., Srihari, V. H., Pollard, J. M., Addington, D. N., & Woods, S. W. (2010). Mediators and moderators in early intervention research. *Early Intervention in Psychiatry*, *4*(2), 143–152. https://doi.org/10.1111/j.1751-7893.2010.00177.x

Chiesa, A., & Serretti, A. (2011). Mindfulness-Based Interventions for Chronic Pain: A Systematic Review of the Evidence. *The Journal of Alternative and Complementary Medicine*, *17*(1), 83–93. https://doi.org/10.1089/acm.2009.0546

Cohen, S. P., Vase, L., & Hooten, W. M. (2021). Chronic pain: an update on burden, best practices, and new advances. *Lancet (London, England)*, *397*(10289), 2082–2097. https://doi.org/10.1016/S0140-6736(21)00393-7

Correll, D. J. (2011). Chapter 22 - The Measurement of Pain: Objectifying the Subjective. In S. D. Waldman (Ed.), *Pain Management (Second Edition)* (pp. 191–201). W.B. Saunders. https://doi.org/10.1016/B978-1-4377-0721-2.00022-2

Frota, N. T., Fidelis-de-Paula-Gomes, C. A., Pontes-Silva, A., Pinheiro, J. S., de Jesus, S. F. C., Apahaza, G. H. S., da Silva Souza, C., Avila, M. A., & Dibai-Filho, A. V. (2022). 15-item Roland-Morris Disability Questionnaire (RMDQ-15): structural and criterion validity on patients with chronic low back pain. *BMC Musculoskeletal Disorders*, *23*(1), 978. https://doi.org/10.1186/s12891-022-05953-y

Geneen, L. J., Moore, R. A., Clarke, C., Martin, D., Colvin, L. A., & Smith, B. H. (2017). Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. *Cochrane Database of Systematic Reviews*, *4*. https://doi.org/10.1002/14651858.CD011279.pub3

Goldin, P. R., & Gross, J. J. (2010). Effects of Mindfulness-Based Stress Reduction (MBSR) on Emotion Regulation in Social Anxiety Disorder. *Emotion (Washington, D.C.)*, *10*(1), 83–91. https://doi.org/10.1037/a0018441

Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care & Research*, *63*(S11). https://doi.org/10.1002/acr.20543

Hays, R. D., Spritzer, K. L., Schalet, B. D., & Cella, D. (2018). PROMIS®-29 v2.0 Profile Physical and

Mental Health Summary Scores. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *27*(7), 1885–1891. https://doi.org/10.1007/s11136-018-1842-3

Hilton, L., Hempel, S., Ewing, B. A., Apaydin, E., Xenakis, L., Newberry, S., Colaiaco, B., Maher, A. R., Shanman, R. M., Sorbero, M. E., & Maglione, M. A. (2017). Mindfulness Meditation for Chronic Pain: Systematic Review and Meta-analysis. *Annals of Behavioral Medicine*, *51*(2), 199–213. https://doi.org/10.1007/s12160-016-9844-2

Jiang, X., & Xu, C. (2022). Deep Learning and Machine Learning with Grid Search to Predict Later Occurrence of Breast Cancer Metastasis Using Clinical Data. *Journal of Clinical Medicine*, *11*(19), 5772. https://doi.org/10.3390/jcm11195772

Kernbach, J. M., & Staartjes, V. E. (2022). Foundations of Machine Learning-Based Clinical Prediction Modeling: Part II-Generalization and Overfitting. *Acta Neurochirurgica. Supplement, 134*, 15–21. https://doi.org/10.1007/978-3-030-85292-4_3

Krebs, E. E., Lorenz, K. A., Bair, M. J., Damush, T. M., Wu, J., Sutherland, J. M., Asch, S. M., & Kroenke, K. (2009). Development and Initial Validation of the PEG, a Three-item Scale Assessing Pain Intensity and Interference. *Journal of General Internal Medicine*, *24*(6), 733–738. https://doi.org/10.1007/s11606-009-0981-1

La Cour, P., & Petersen, M. (2015). Effects of Mindfulness Meditation on Chronic Pain: A Randomized Controlled Trial. *Pain Medicine*, *16*(4), 641–652. https://doi.org/10.1111/pme.12605

Liew, B. X. W., Kovacs, F. M., Rügamer, D., & Royuela, A. (2022). Machine learning versus logistic regression for prognostic modelling in individuals with non-specific neck pain. *European Spine Journal*, *31*(8), 2082–2091. https://doi.org/10.1007/s00586-022-07188-w

Lim, J.-A., Choi, S.-H., Lee, W. J., Jang, J. H., Moon, J. Y., Kim, Y. C., & Kang, D.-H. (2018). Cognitive-behavioral therapy for patients with chronic pain. *Medicine*, *97*(23), e10867. https://doi.org/10.1097/MD.0000000000010867

Lötsch, J., & Ultsch, A. (2018). Machine learning in pain research. *Pain*, *159*(4), 623–630. https://doi.org/10.1097/j.pain.0000000000001118

Melzack, R. (2005). The McGill pain questionnaire: from description to measurement. *Anesthesiology*, *103*(1), 199–202. https://doi.org/10.1097/00000542-200507000-00028

Morone, N. E. (2019). Not Just Mind Over Matter: Reviewing With Patients How Mindfulness Relieves Chronic Low Back Pain. *Journal of Evidence-Based Integrative Medicine*, *24*, 2515690X1983849. https://doi.org/10.1177/2515690X19838490

Morone, N. E., Greco, C. M., Moore, C. G., Rollman, B. L., Lane, B., Morrow, L. A., Glynn, N. W., &

Weiner, D. K. (2016). A Mind-Body Program for Older Adults With Chronic Low Back Pain: A Randomized Clinical Trial. *JAMA Internal Medicine*, *176*(3), 329–337. https://doi.org/10.1001/jamainternmed.2015.8033

Nephew, B. C., Incollingo Rodriguez, A. C., Melican, V., Polcari, J. J., Nippert, K. E., Rashkovskii, M., Linnell, L.-B., Hu, R., Ruiz, C., King, J. A., & Gardiner, P. (2022). Depression Predicts Chronic Pain Interference in Racially Diverse, Income-Disadvantaged Patients. *Pain Medicine*, *23*(7), 1239–1248. https://doi.org/10.1093/pm/pnab342

Sarwar, A., Agu, E. O., Polcari, J., Ciroli, J., Nephew, B., & King, J. (2022). PainRhythms: Machine learning prediction of chronic pain from circadian dysregulation using actigraph data — a preliminary study. *Smart Health*, *26*, 100344. https://doi.org/10.1016/j.smhl.2022.100344

Schuman-Olivier, Z., Trombka, M., Lovas, D. A., Brewer, J. A., Vago, D. R., Gawande, R., Dunne, J. P., Lazar, S. W., Loucks, E. B., & Fulwiler, C. (2020). Mindfulness and Behavior Change. *Harvard Review of Psychiatry*, *28*(6), 371–394. https://doi.org/10.1097/HRP.0000000000000277

Wolfe, F., Clauw, D. J., Fitzcharles, M., Goldenberg, D. L., Katz, R. S., Mease, P., Russell, A. S., Russell, I. J., Winfield, J. B., & Yunus, M. B. (2010). The American College of Rheumatology Preliminary Diagnostic Criteria for Fibromyalgia and Measurement of Symptom Severity. *Arthritis Care & Research*, *62*(5), 600–610. https://doi.org/10.1002/acr.20140