# Greenspace Mental Health
# Group 3
# Therapeutic Alliance

## Draft Proposal

Kohsin Chen
Zerui Zhang
Zheng Zhang
Bingshen Yang

# Introduction

## Background and Literature Review

Therapeutic alliance, often described as the collaborative and affective bond between a therapist and a patient, plays a pivotal role in the efficacy of psychological treatments. Existing research underscores the importance of the Therapeutic Alliance in mental health treatment outcomes. A study by Ardito and Rabellino highlights that the Therapeutic Alliance is a consistent predictor of positive clinical outcomes across different psychotherapies, suggesting that the relationship quality can significantly influence recovery trajectories [1]. Additionally, research by Martin et al. discusses various methods and models used to assess and predict the quality of Therapeutic Alliance, including the use of machine learning techniques to analyze session transcripts and patient feedback to predict alliance ratings [2]. This research shows factors such as empathy, agreement on therapy goals, and mutual collaboration are integral to fostering a strong therapeutic relationship. These studies provide a foundation for understanding the complexities of Therapeutic Alliance and demonstrate the potential of using advanced analytics to enhance treatment outcomes in mental health care.

Further compounding the complexity of mental health treatment is the challenge of identifying at-risk patients. Soler et al. explore the development of early warning systems in healthcare settings, using predictive analytics to flag patients who show signs of critical deterioration [3]. These systems serve as a blueprint for applying similar predictive methodologies to mental health scenarios, where timely intervention can be crucial.

Greenspace Health[4] has developed a comprehensive Measurement Based Care (MBC) platform tailored for mental health professionals, offering critical insights into patient outcomes and enhancing care decisions. The platform’s assessments component allows patients to complete mental health condition-related questionnaires throughout their treatment.Therefore, we aim to leverage the advanced capabilities of the Greenspace Health MBC platform to build predictive models that not only assess and enhance the Therapeutic Alliance but also identify patients who are at risk of deteriorating mental health conditions.

### Problem Overview

The significance of a robust therapeutic alliance drives the need for intake workers and clinical supervisors to have detailed insights into the factors that enhance this relationship, thereby facilitating the assignment of the most suitable therapist at the onset of treatment. However, establishing and maintaining a robust therapeutic alliance can be challenging due to the myriad factors influencing its development. These factors can be intrinsic to the patient, inherent to the therapist, or a combination of both, influencing their interpersonal dynamics.
Compounding this challenge is the crucial task of flagging risky patients. Early identification of these individuals is vital for the timely adjustment of therapeutic strategies and interventions. Beyond just identifying at-risk patients, it is essential to determine which specific measurements can be implemented to improve treatment outcomes. Utilizing predictive analytics and tailored interventions not only supports the maintenance of the therapeutic alliance but also ensures that each patient receives the most effective care tailored to their specific needs.

### Motivation and Purpose

The primary motivation for this project is to enhance the effectiveness of mental health treatment by identifying the predictors of a strong Therapeutic Alliance and flagging patients at risk of poor outcomes. By developing models that can predict these elements, we aim to enable healthcare providers to allocate the right resources efficiently, personalize care, and ultimately improve patient outcomes. This project seeks to answer critical questions such as what factors contribute to a robust Therapeutic Alliance and how to efficiently identify and manage patients who are at risk of worsening conditions or disengagement.

## Aims and Objectives

### Research Questions

The primary research questions we aim to investigate in this project are:

- What factors predict or influence the Therapeutic Alliance between a patient and their therapist? This question seeks to identify variables (demographic, therapeutic history, assessment responses, etc.) that correlate with or predict higher Therapeutic Alliance scores.

- How can we identify and prioritize high-risk patients for therapists? This involves defining 'high-risk' in terms of worsening assessment scores and disengagement levels and determining the predictors for such risks. Additionally, we could explore cases of patients who started out as “risky” and then showed improvement. By analyzing what actions or interventions were implemented by their therapists during treatment, we can gain insights into effective strategies and recommend these practices to therapists when they encounter future high-risk patients.

**Methods**

The analytical approach will begin with an extensive Exploratory Data Analysis (EDA) to uncover trends, detect outliers, and establish correlations within the data. This initial exploration will help us understand the variables influencing the therapeutic alliance and patient risk levels. Python will be the primary programming language used throughout our project.

For predictive modeling, we plan to use a variety of techniques. Regression Analysis will help us predict Therapeutic Alliance scores, examining how continuous predictor variables like assessment scores influence these scores. Random Forests, an ensemble learning method, will be employed to classify patients into different risk categories based on their assessment scores and other behavioral data. Additionally, XGBoost will be utilized to enhance predictive accuracy by addressing data biases and variance, especially useful for our complex datasets.

Classification tasks will be handled using Logistic Regression to categorize patients as high or low risk based on their data profiles. For more complex, higher-dimensional data, Support Vector Machines (SVM) will be crucial for delineating the optimal boundary between different classes of therapeutic outcomes.

This project is inherently multi-faceted and open-ended, reflecting the complex nature of therapeutic interactions and patient care dynamics. Recognizing that the challenges and variables involved in predicting therapeutic alliance and identifying high-risk patients are extensive and varied, our approach is designed to be highly adaptable and iterative. Throughout the project, we will engage in an iterative process, where initial findings and model performances inform subsequent rounds of data analysis, feature selection, and model refinement. This iterative cycle will allow us to continuously improve our models and strategies based on real-time feedback and evolving data insights. By embracing this flexible and progressive

methodology, we aim to progressively hone our tools and techniques to better meet the project's goals and effectively contribute to the enhancement of patient care at Greenspace Health.

## Dataset

### Dataset Description

The dataset comprises a series of tables designed to capture the various aspects of therapeutic interactions, including generalized information about psychological assessments, patient responses and therapist attributes.
The BR-WAI (Bonding-Relationship Working Alliance Inventory) assessment measures the quality of the alliance between a patient and therapist. This is particularly interesting as it can directly correlate with therapeutic outcomes. Besides, data of therapists can enhance this analysis by linking patient outcomes to therapist effectiveness. Understanding which therapists achieve better results with specific types of patients can guide more strategic therapist-patient pairings. This data can indicate which therapists are potentially better equipped to handle high-risk patients.

In the context of assigning risk levels to patients, keeping track of which assessments are given to which patients over time. This allows for monitoring changes in patients' conditions, particularly in cases where patients start as risky and then improve. Understanding what therapists did to cause this improvement can provide valuable insights into effective therapeutic approaches. Similarly, monitoring how patients react to their treatments over time is also essential. This helps in detecting non-improvement or worsening conditions, which could indicate the need for intervention or a change in treatment approach. This information can then be used to recommend altered treatment plans or more focused care strategies for therapists dealing with similar "risky" patients in the future.

### Data Preparation

In preparing the dataset for analysis, the data cleaning process is essential to ensure the accuracy and usability of the information. The first step involves validity checks, with a focus on key fields such as scores and dates. This is crucial to ensure there are no inconsistencies in dates that could skew analysis results. After that, missing data handling depends on the attribute types, appropriate techniques like simple imputation or deletion are applied to manage these missing entries. Besides,

the process involves scanning and eliminating any duplicate records to maintain the uniqueness and integrity of the data.

Once the data is cleaned, the next procedure is data integration. This involves ensuring robust linkage across different tables through universal identifiers for patients and therapists, which is important for maintaining referential integrity while safeguarding personal information. Furthermore, aggregated views of the data from multiple tables should be created for analyses. These views compile relevant data from various tables into a single dataset, simplifying access and building models.

## Deliverables and Schedule/Timeline

The project will kick off on April 29th and include a mid-project presentation on May 27th, culminating in a final presentation on June 25th. We will adopt a weekly sprint model, launching a new sprint each Sunday beginning May 5. Our communication plan includes weekly check-ins with Mr. Nicholas Kwan every Thursday via Zoom, and meetings with team members every Wednesday and Saturday.

The final deliverable for enhancing the Therapeutic Alliance includes a predictive model report, a therapist recommendation system, and a guide of actionable insights for therapists. The predictive model report outlines significant factors that influence the relationship between a patient and their therapist, utilizing statistical analysis and machine learning to identify key predictors. Building on this, the recommendation system matches patients with therapists who are likely to form a strong therapeutic alliance with them, considering factors such as therapist specialties and communication styles. Also, the actionable insights guide compiles effective strategies from successful therapy interactions, offering therapists concrete suggestions on how to improve their therapeutic relationships. These strategies might include adjustments in therapy techniques or enhancements in communication, aimed at fostering a stronger connection between the therapist and patient. This set of tools is designed to optimize therapeutic outcomes by ensuring a good match and providing ongoing support for therapy professionals.

Below is a concise outline of the weekly objectives for the project:

Sprint 1: Exploring

| Task | Category | Assign to | Due by |
|---|---|---|---|
| Dataset exploring | Exploring | Bingshen | 2024-05-12 |
| Dataset loading | Exploring | Zheng | 2024-05-12 |
| Dataset cleaning | Exploring | Kohsin, Zerui | 2024-05-12 |

Sprint 2: Setting up the model

| Task | Category | Assign to | Due by |
|---|---|---|---|
| Model Type Discussion | Planning | Bingshen, Zheng, Kohsin, Zerui | 2024-05-15 |
| Model Setup and Initialization | Implementation | Bingshen, Zheng, Kohsin, Zerui | 2024-05-23 |
| Model Fine Tuning i | Optimization | Bingshen, Zheng, Kohsin, Zerui | 2024-05-26 |
| Mid-point presentation | Presentation | Bingshen, Zheng, Kohsin, Zerui | 2024-05-26 |

Sprint 3: Analyze the result

| Task | Category | Assign to | Due by |
|---|---|---|---|
| Model Comparison | Evaluation | Bingshen, Zheng, Kohsin, Zerui | 2024-05-31 |
| Key Features identification | Analysis | Bingshen, Zheng, Kohsin, Zerui | 2024-06-03 |
| Key Features Verification | Validation | Bingshen, Zheng, Kohsin, Zerui | 2024-06-03 |
| Model Prediction Accuracy Test | Testing | Bingshen, Zheng, Kohsin, Zerui | 2024-06-09 |

Sprint 4: Final report and presentation

| Task | Category | Assign to | Due by |
|---|---|---|---|
| Model Fine Tuning ii | Optimization | Bingshen, Zheng, Kohsin, Zerui | 2024-06-12 |
| Confirm the factors | Validation | Bingshen, Zheng, | 2024-06-15 |

| | | | |
|---|---|---|---|
| | | Kohsin, Zerui | |
| Final report | Documentation | Bingshen, Zheng, Kohsin, Zerui | 2024-06-27 |
| Final presentation | Presentation | Bingshen, Zheng, Kohsin, Zerui | 2024-06-27 |

## References


1. Ardito, R. B., & Rabellino, D. (2011). Therapeutic alliance and outcome of psychotherapy: historical excursus, measurements, and prospects for research. Frontiers in Psychology, 2, 270.
2. Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. Journal of Consulting and Clinical Psychology, 68(3), 438.
3. Soler, J. K., & Okkes, I. (2003). Early warning systems for identifying at-risk patients in hospital settings. Journal of Health Informatics, 15(2), 99-110.
4. Greenspace Health. (n.d.). Home. Retrieved May 4, 2024, from https://greenspacehealth.com/en-ca/