Multi-Outcome Trajectories in Traumatic Brain Injury

by

Vladyslav Shein

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Public Health Sciences

Waterloo, Ontario, Canada, 2025

© Vladyslav Shein 2025

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Traumatic Brain Injury (TBI) presents a global health challenge, affecting millions of individuals annually, resulting in diverse outcome trajectories that complicate patient management. The heterogeneity in TBI outcomes, influenced by varied clinical presentations and injury responses, requires advanced analytical approaches. The analysis of trajectories using single metrics, such as the Glasgow Outcome Scale Extended Global Score (GOSE Score), falls short of capturing the multi-faceted nature of TBI progression, often overlooking the complexity of individual patient experiences.

This thesis reports on **two** studies. First, a systematic scoping review was conducted to synthesize the current research on trajectory analysis in TBI, followed by a modeling study. This work identifies **6** distinct multi-outcome trajectories in TBI patients by employing Latent Class Mixed Models (LCMM) and clustering approaches. Utilizing longitudinal data from the Transforming Research and Clinical Knowledge in Traumatic Brain Injury study, a prospective multicenter observational cohort study conducted at 18 level 1 trauma centers across the United States (TRACK-TBI), which includes **17** selected outcome measures collected at four time points post-injury, provides a comprehensive understanding of the heterogeneous progression of TBI. By addressing the limitations of single-outcome analyses, this research contributes to a better understanding of TBI progression that can lead to the optimization of TBI management and treatment. The future integration of these trajectories will facilitate the development of personalized treatment strategies, ultimately improving patients' recovery.

Acknowledgements

I would like to express my deepest gratitude to my scientific supervisor and advisor, *Dr. Abel Torres Espín*, for his invaluable guidance, mentorship, and unwavering support throughout this research. His broad understanding, patience, and insightful advice continually steered me in the right direction. His expertise and encouragement have been instrumental in shaping this work and inspiring my development as a researcher.

I am profoundly thankful to all *health.data DRIVEN Lab* members for their continuous support, collaboration, and friendship. The welcoming and cooperative atmosphere they created made this journey both productive and enjoyable. I am especially grateful for their help throughout the entire process — from refining this work and preparing the presentation to offering thoughtful questions, comments, and recommendations that strengthened the quality of the thesis.

Finally, but no less important, I would like to extend my sincere appreciation to my committee members, *Dr. Joel Dubin* and *Dr. Peter A. Hall*, for their guidance, insightful feedback, and continuous encouragement. Their contributions and perspectives have played an essential role in the successful completion of this research.

Table of Contents

Author's Declaration ii

Abstract iii

Acknowledgements iv

List of Figures viii

List of Tables ix

List of Abbreviations x

1 Introduction 1

1.1 Background of TBI and its Clinical Challenges 1

1.2 Limitations of Current Analytical Approaches 2

1.3 Multi-outcome Trajectory Methodology 3

1.4 Research Objectives 4

2 Scoping Review 5

2.1 Introduction 5

2.2 Methods 5

2.2.1 Review Framework 5

2.2.2 Research Questions 6

2.2.3 Data Extraction and Management . . . 8
2.3 Results . . . 8
2.3.1 Study Selection and Characteristics . . . 8
2.3.2 Study Designs . . . 11
2.3.3 Outcome Measures and Assessment Tools . . . 11
2.3.4 Identified Trajectories . . . 12
2.3.5 Factors Associated with Recovery . . . 13
2.3.6 Methodological Challenges and Adaptations . . . 15
2.4 Discussion . . . 17
2.4.1 Limitations and Future Research . . . 19
2.5 Conclusions . . . 20

3 Modeling Multi-Outcome Trajectories **22**
3.1 Methods . . . 23
3.1.1 Study Population and Data Source . . . 23
3.1.2 Outcome Selection and Domain Grouping . . . 29
3.1.3 Missing Data Analysis . . . 33
3.1.4 Single-outcome Modeling . . . 39
3.1.5 Clustering Framework . . . 46
3.2 Results . . . 49
3.2.1 Single-outcome Modeling . . . 49
3.2.2 Multi-outcome Trajectory Identification . . . 52
3.2.3 Trajectory Analysis . . . 56

4 An Interactive Analysis Tool **63**
4.1 The R Shiny Framework . . . 63
4.2 Application 1: Link Function Comparison . . . 64
4.2.1 UI Architecture . . . 64

4.2.2 Use Cases and Scenarios . . . 65
4.2.3 Main Components . . . 65
4.3 Application 2: Single Model Comparison . . . 66
4.3.1 UI and Layout . . . 66
4.3.2 Interactive Model Comparison Features . . . 68
4.3.3 Use Cases and Workflows . . . 68
4.3.4 Trajectory Plots . . . 70
4.4 Application 3: Multi-outcome Clustering . . . 71
4.4.1 UI Architecture . . . 71
4.4.2 Clustering Control Panel . . . 74
4.4.3 Visualization Components . . . 74
4.4.4 Use Cases and Analytical Workflows . . . 75
4.4.5 Additional Functional Features . . . 76
4.5 System Integration and Deployment . . . 77
4.6 Conclusion . . . 77

5 Discussion and Contribution **79**

6 Limitations and Future Considerations **82**

7 Conclusion **86**

References **88**

APPENDICES **97**

List of Figures

2.1 PRISMA Flow Diagram 9

3.1 Modeling Pipeline Chart 23
3.2 Link Functions Plots 43
3.3 Single-outcome (LCMMs) Trajectories 52
3.4 Cluster Dendrogram 53
3.5 Cluster Trajectories 54

4.1 Sidebar Panel 64
4.2 Sidebar Panel Model Summary 67
4.3 Multi-tab Interface 71
4.4 Multi-tab Interface 73

List of Tables

2.1 Study Selection Criteria 7
2.2 Distribution of Studies by TBI Severity Focus 10
2.3 Primary Outcome Measures Utilized Across Studies 11
2.4 Factors Associated with Recovery 14

3.1 Subjects Characteristics by ED GCS 25
3.2 Selected Outcome Variables by Domain 30
3.3 Data Availability by Outcome Variable for Included Subjects 34
3.4 Distribution of Subjects Attrition Reasons 35
3.5 Variables with Highest Missing Percentages 37
3.6 Distribution of Imputation Methods 38
3.7 Selected Link Functions 43
3.8 Cluster Metrics for Different K 48

1 List of Articles Included in the Scoping Review 97
2 Comparison of Link Functions for Selected Outcomes 98
3 Subject's Characteristics by Cluster 103
4 Demographic Comparison Between Dropout and Retained Groups 104
5 Fit Metrics for Fitted Models 105

List of Abbreviations

ACRM American Congress of Rehabilitation Medicine 24

ADM Hospital Admission 24

AIC Akaike Information Criterion 16, 41, 42, 49, 65, 68

BA Brief Assessment 24

BIC Bayesian Information Criterion 16, 41, 42, 49, 65, 68

BSI-18 Brief Symptom Inventory 18 30

BSI-18 Anxiety Score Brief Symptom Inventory-18 Anxiety Subscale Raw Score 30, 32, 34, 37, 43, 50, 54, 56, 58, 59

BSI-18 Depression Score Brief Symptom Inventory-18 Depression Subscale Raw Score 30, 32, 34, 37, 43, 50, 54, 56

BSI-18 GSI Brief Symptom Inventory 18 Global Severity Index 30

CA Comprehensive Assessment 24

CA+MRI Comprehensive Assessment plus Magnetic Resonance Imaging 24, 28

CENTER-TBI Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury, a prospective longitudinal observational study 10

CHI Calinski–Harabasz-index 46–48

CI Confidence Interval 53, 54, 70, 71

COBRIT Citicoline Brain Injury Treatment Trial, a Phase 3 randomized controlled study that investigated whether citicoline improved functional and cognitive outcomes after TBI 11, 22

CRUD Create, Read, Update, Delete 66

CSV Comma-separated Values 65, 67–69

CT Computed Tomography 28

DOC Disorders of Consciousness 14

DRS Score Expanded Disability Rating Scale Post-Acute Interview 30, 32–34, 43, 50, 54, 56–59, 61, 80

ED Emergency Department 23, 24

FC Friend Control 24

FRP Functional Reactive Programming 63

GBTM Group-Based Trajectory Modeling 5

GCS Glasgow Coma Scale 13, 14, 24, 57, 59, 61, 62

GDP Gross Domestic Product 2

GMM Growth Mixture Model 5, 18

GOSE Score Glasgow Outcome Scale Extended Global Score iii, 2, 3, 30, 32, 34, 37–39, 43, 49–51, 55–59, 61, 62, 80

GSEM Generalized Structural Equations Model 16

HTML HyperText Markup Language 63

ICU Intensive Care Unit 12, 15, 17, 24, 28, 35, 57, 59, 61, 62

IMPACT International Mission for Prognosis and Analysis of Clinical Trials in TBI 32

ISI Score Insomnia Severity Index 31, 32, 34, 37, 43, 50, 51, 55, 57

LCGA Latent Class Growth Analysis 5, 16, 18

LCMM Latent Class Mixed Models iii, viii, 5, 12, 16, 18, 19, 22, 39–41, 46, 49, 52, 63, 70, 77, 82, 84, 86

LOS Length of Stay 13–15, 18

MAE Mean Absolute Error 39

MAR Missing at Random 35, 38

MCAR Missing Completely at Random 38

MICE Multiple Imputation by Chained Equations 36, 38

MIMIC Medical Information Mart for Intensive Care, a comprehensive longitudinal study at the Beth Israel Deaconess Medical Center ICU 10

ML Machine Learning 5, 18

MNAR Missing Not at Random 34

MRI Magnetic Resonance Imaging 28

OC Orthopedic Control 24

PCL-5 Score PTSD Checklist for DSM-5 Mental health score 31, 32, 34, 37, 43, 50, 51, 55–58

PMM Predictive Mean Matching 38

PROMIS Participant Reported Outcome Measurement Information System 30

PROMIS Intensity PROMIS Pain Intensity Pain Raw score 31, 32, 34, 43, 50, 51, 55, 57–60

PTA Post-Traumatic Amnesia 13, 18, 61

PTSD Post-Traumatic Stress Disorder 56–59

QoLIBRIOS Score Quality of Life After Brain Injury Overall Scale 32, 34, 39, 43, 50, 51, 55, 57, 58, 60

RAVLT Score Rey Auditory Verbal Learning Test II Sum Score 31, 32, 34, 36, 37, 43, 50, 51, 55, 57–60

RF Random Forest 16, 38

RMSE Root Mean Square Error 39

RPQ Score Rivermead Post-Concussion Questionnaire Total Score 31, 32, 34, 43, 50, 51, 55, 58–60

RPQ-13 Score Rivermead Post-Concussion Questionnaire Subscale 13: Cognitive and Emotional Symptoms 31, 32, 34, 43, 50, 51, 55, 58, 59

RPQ-3 Score Rivermead Post-Concussion Questionnaire Subscale 3: Core Physical Symptoms 31, 32, 34, 43, 50, 51, 55, 58, 59

SC Silhouette Coefficient 46–48

SES Socioeconomic status 14

SWLS Score Satisfaction With Life Scale Total Score 32, 34, 43, 50, 51, 55, 57–60

TBI-CDEs NIH TBI Common Data Elements 23

TMT Trail Making Test 29, 59

TMTA Attention Score Trail Making Test A Scaled Score 31, 32, 34, 36, 37, 43, 50, 51, 55, 57, 58, 60

TMTB Executive Score Trail Making Test B Scaled Score 31, 32, 34, 36, 37, 43, 50, 51, 55, 57

TRACK-TBI Transforming Research and Clinical Knowledge in Traumatic Brain Injury study, a prospective multicenter observational cohort study conducted at 18 level 1 trauma centers across the United States iii, 4, 10, 22–24, 28, 29, 32, 33, 83, 84, 86

UI User Interface 63, 64, 74

WAIS Wechsler Adult Intelligence Scale 29

WAIS Coding WAIS-IV Coding Processing Speed Raw Score 31–34, 36, 37, 43, 50, 52, 56–60

WAIS Symbol WAIS-IV Symbol Search Processing Speed Raw Score 31, 32, 34, 36, 37, 43, 50, 52, 56–58

WCSS Within-Cluster Sum of Squares 46–48

Chapter 1

Introduction

1.1 Background of TBI and its Clinical Challenges

Traumatic brain injury or TBI is characterized by damage to the brain caused by an external mechanical force, which can lead to temporary or permanent impairment of cognitive, physical, and psychosocial functions [1]. TBI represents a significant public health challenge, affecting between **27** and **69** million individuals annually worldwide and contributing to significant personal, societal, and economic burdens [2, 3]. The heterogeneous nature of TBI, characterized by diverse mechanisms of injury, anatomical locations, and severity levels, presents unique challenges for understanding progression after injury and predicting long-term outcomes [4]. The consequences of TBI are profound, affecting not only the individuals who suffer from it but also their families, friends, and communities, with outcomes ranging from complete recovery to permanent disability or even death. The immediate aftermath often includes emergency medical care, intensive monitoring, and possibly surgical interventions, while long-term consequences can involve chronic health issues, such as neurological symptoms that can include motor dysfunction, cognitive and emotional impairments, and secondary complications like epilepsy, which require ongoing medical and psychological support [1].

The diverse etiology of TBI, including falls, motor vehicle accidents, sports injuries, and assaults, contributes to the varied clinical presentations observed in patients [4]. Falls are particularly prevalent among the elderly and young children, while motor vehicle accidents are a leading cause among adolescents and young adults [5]. Sports-related TBI is common in younger populations, particularly in contact sports [6]. Recent epidemiological studies reveal shifting patterns, with falls now becoming the most common cause of TBI in some

populations, surpassing traffic accidents and indicating evolving public health priorities for TBI prevention [7].

Physical and motor deficits following TBI include balance disorders, gait abnormalities, and movement difficulties. Cognitive deficits affect attention, memory, executive function, language skills, and orientation. Behavioral and emotional deficits involve changes in behavior and emotional regulation, which are common consequences of TBI and significantly affect patients' quality of life and social interactions [8]. This diversity in symptoms and severities makes it challenging to group TBI patients into homogeneous categories, which complicates the development of specific therapies and diagnostic approaches [9].

The economic burden of TBI is substantial, encompassing direct medical costs, rehabilitation expenses, and indirect costs related to loss of productivity and long-term care. In the US alone, nonfatal TBI-related medical expenditures in 2016 were estimated at **$40.6** billion, with **$10.1** billion from private insurance, **$22.5** billion from Medicare, and **$8** billion from Medicaid [10]. Broader estimates suggest that the total economic impact of TBI cases in 2012 may have reached **$758** billion, of which only **$26** billion were direct medical costs—highlighting the dominant share of indirect costs due to lost productivity and diminished quality of life. Globally, the annual economic cost of TBI has been estimated at approximately **$400** billion, representing around **0.5%** of the global Gross Domestic Product (GDP) [4].

This complexity necessitates a multidisciplinary and integrated approach to treatment and management, while also emphasizing the urgent need for more advanced analytical methods to better understand and predict recovery trajectories.

1.2 Limitations of Current Analytical Approaches

Traditional approaches to understanding TBI outcomes have relied heavily on single-outcome metrics and cross-sectional analyses, which fail to capture the dynamic and multidimensional nature of neurological recovery. A popular approach for understanding patient heterogeneity is to determine subgroups of individuals that present similar progression over time, or trajectory. For instance, Gardner *et al.* (2018) identified divergent six-month functional recovery trajectories among TBI patients using GOSE Score [5]. Similarly, Wickwire *et al.* (2022) identified multiple insomnia trajectories within the first year post-injury [1]. Although these efforts have been instrumental in discovering group trajectories, single-outcome metrics such as GOSE Score often fall short of providing a comprehensive understanding of TBI recovery. These metrics fail to account for the multidimensional nature of

recovery, which includes global function, neurological, and quality-of-life [5]. Indeed, studies have shown that different symptoms can lead to the same GOSE Score, masking the true variability in patient recovery [11], [12]. The reliance on single metrics can lead to incomplete assessments and hinder the development of effective management and treatment [1].

1.3 Multi-outcome Trajectory Methodology

Multi-outcome trajectory modeling represents a powerful analytical framework that addresses many of the limitations inherent in traditional single-metric approaches to TBI research. This methodology focuses on characterizing patterns of change over time across multiple outcome domains simultaneously, allowing for the identification of distinct recovery pathways and the factors that influence these complex trajectories.

A multi-outcome analysis of TBI recovery allows for a more holistic understanding of patient outcomes by considering various functional and symptomatic dimensions concurrently. This approach aligns with the principles of precision health, which emphasize the importance of tailored treatment strategies based on individual patient characteristics and recovery patterns. By identifying distinct trajectories through the integration of cognitive, neurological, and quality-of-life metrics, clinicians can better understand the specific needs of TBI patients and develop targeted interventions to improve outcomes. The fundamental principle underlying multi-outcome trajectory analysis is the recognition that recovery from TBI is not a uniform process, but rather consists of multiple distinct patterns that can be mathematically characterized and predicted across multiple outcomes simultaneously [13]. By employing advanced statistical techniques such as growth mixture models, researchers can identify subgroups of patients who follow similar recovery patterns across multiple outcome measures, capturing the complex interplay between different recovery domains [14].

This approach offers several advantages over traditional analytical methods. It allows for the characterization of both typical and atypical recovery patterns across multiple outcomes, the identification of critical time points in the recovery process, and the development of personalized prediction models. Furthermore, multi-outcome trajectory analysis can accommodate missing data and irregular follow-up intervals, which are common challenges in longitudinal TBI research, while providing a more comprehensive and nuanced understanding of recovery patterns.

1.4 Research Objectives

Given the limitations of single-outcome analysis outlined above, this research aims to identify distinct multi-outcome functional trajectories in TBI patients through two specific objectives that build upon each other to provide a comprehensive understanding of trajectory-based recovery patterns.

The first objective involves conducting a scoping review to identify and analyze existing literature on trajectory analysis in TBI. This review gathers and synthesizes current knowledge on the topic, pinpoints methodological approaches previously used, and highlights gaps in existing research. The scoping review examines studies that employ various outcome metrics to evaluate distinct trajectories.

The second objective focuses on identifying distinct multi-outcome trajectories in TBI patients over time using a combination of probabilistic models and cluster analysis. This will be achieved through secondary analysis of longitudinal data from the TRACK-TBI [15].

To support both objectives, an interactive analysis tool has been developed. This tool is designed to enhance the decision-making process by enabling systematic comparison of different models and evaluation of clusters. It provides additional visualization features that allow for clearer interpretation of trajectories, offering a user-friendly interface that simplifies data exploration and model selection.

Through these objectives, this research seeks to contribute to the optimization of TBI management and treatment, ultimately improving patient outcomes and quality of life while supporting precision health initiatives in neurorehabilitation.

Chapter 2

Scoping Review

2.1 Introduction

In this scoping review, I synthesize the current evidence on *trajectory- and cluster-based approaches* to characterizing post-injury progression after TBI. I focus on methodological frameworks, the types, and shapes of outcome trajectories, and their clinical implications. The review emphasizes data-driven techniques—including latent class and mixture models (e.g., Latent Class Growth Analysis (LCGA), Growth Mixture Models (GMMs)/LCMMs), Group-Based Trajectory Modelings (GBTMs), and selected Machine Learning (ML)-based approaches—to capture heterogeneity in recovery. By consolidating recent findings, the review highlights emerging trends and delineates opportunities for more personalized, trajectory-informed TBI care [16].

2.2 Methods

2.2.1 Review Framework

This scoping review was conducted following the framework proposed by *Arksey and O'Malley*, which is widely recognized for its structured and systematic approach to reviewing broad research topics [17]. This methodology is particularly suitable for identifying key concepts, gaps, and research trends across diverse study designs. The process consisted of **four** main stages: establishment of relevant research questions and study planning, study selection through comprehensive database searching, identification of relevant

studies through systematic screening and application of inclusion and exclusion criteria, and summarization and reporting of findings. Given the breadth of designs in the included literature, this approach is well-suited to mapping concepts and methodological gaps across study types. To prevent conflating conceptually distinct approaches while keeping the synthesis concise, the subsequent sections primarily focus on studies that explicitly modeled recovery trajectories. However, several included studies employed related but non-trajectory approaches—such as longitudinal group comparisons—that nevertheless contribute relevant contextual insights. These works are referenced selectively where they support or complement findings from trajectory-based analyses.

2.2.2 Research Questions

Nine specific research questions were formulated to ensure a comprehensive analysis of the literature:

1. What are the aims of the studies examining TBI trajectories?
2. What populations are studied in trajectory-based TBI research?
3. What study designs are employed in this field?
4. What outcome metrics are used to assess trajectories?
5. What analytical methods are used for identifying trajectories?
6. How many distinct trajectories have been identified?
7. What are the main characteristics of these trajectories?
8. What are the primary challenges identified in trajectory-based TBI research?
9. How have methods been adapted to address heterogeneity in TBI recovery?

Search Strategy

The search strategy was designed to ensure comprehensive coverage of the literature relevant to TBI functional trajectories. PubMed was selected as the primary database for its extensive coverage of biomedical literature, supplemented by Google Scholar to capture broader sources and minimize the risk of overlooking relevant studies.

The following search strings were used:

- **PubMed query:**
  ((Traumatic Brain Injury) OR (TBI)) AND (("recovery trajectory") OR ("outcome trajectory") OR ("longitudinal cluster") OR ("temporal cluster")) AND (humans[Filter]) NOT (review[Filter])
- **Google Scholar query:**
  "Traumatic Brain Injury" OR TBI AND ("recovery trajectory" OR "outcome trajectory" OR "longitudinal cluster" OR "temporal cluster") AND humans -review

These queries were designed using Boolean operators to balance sensitivity and specificity, ensuring comprehensive yet focused literature retrieval. Recognizing the evolving nature of the literature, the search process extended over four months, during which articles emerging from citation searches and newly published works were manually included (including those alerted by PubMed). While acknowledging that broader search terms could potentially yield additional studies, the current search strategy and search strings align with the primary review's focus on trajectory-based approaches.

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were developed based on an analysis of previous literature on TBI-related topics and are presented in Table 2.1.

Table 2.1: Study Selection Criteria

Code	Inclusion Criteria	Exclusion Criteria
IC1/EC1	Articles published in English	Non-English publications
IC2/EC2	Peer-reviewed journal publications	Non-peer-reviewed literature
IC3/EC3	Diverse TBI populations	Non-human subjects
IC4/EC4	Longitudinal TBI recovery tracking	Non-longitudinal studies
- /EC5	-	Studies without ethical approval
- /EC6	-	Reviews without original data

2.2.3 Data Extraction and Management

Data extraction was systematically conducted using a structured form that captured key characteristics about the articles and critical information directly linked to the research questions. The extraction process was facilitated using Microsoft Excel, with separate sheets created for each research question to ensure clarity and systematic organization. Covidence was used to manage articles efficiently, facilitating streamlined data extraction, automatic removal of duplicates, and systematic organization throughout the review process. Unlike conventional systematic reviews, no formal quality assessment tool was applied to assess the included studies, as, to my knowledge, no standardized method exists for evaluating trajectory studies specifically. Therefore, studies were included based on their methodological clarity and alignment with the research questions.

2.3 Results

2.3.1 Study Selection and Characteristics

After conducting targeted searches and applying inclusion and exclusion criteria, a total of **16 articles** were selected for this review. Figure 2.1 illustrates the article selection process. All included sources were peer-reviewed journal publications published between **2016 and 2024**, indicating a growing interest in trajectory-based TBI research. The highest publication frequency was observed in **2019** and **2021**, with **four articles each**, followed by **2018** and **2022**, with **two articles each**.

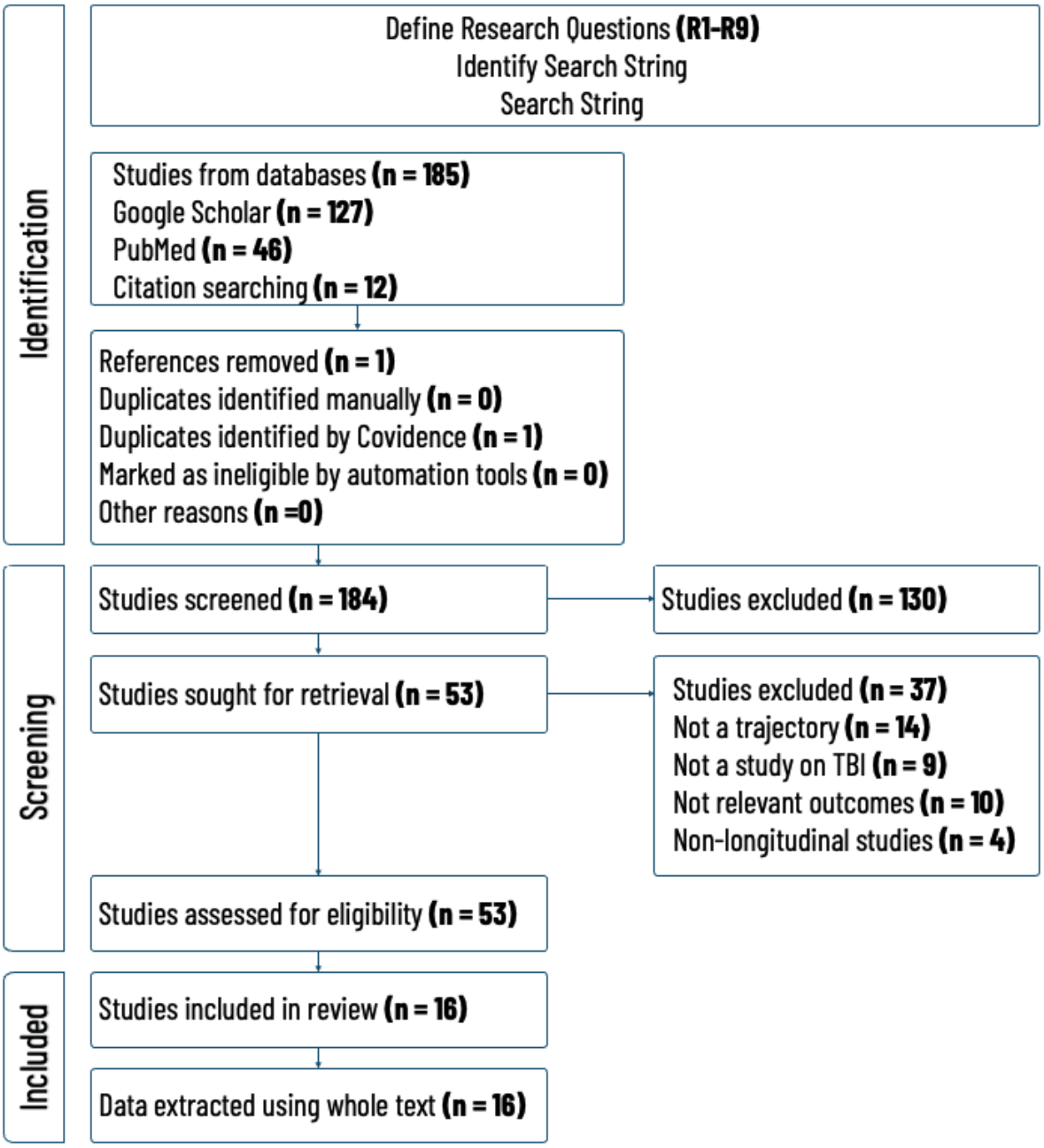


Figure 2.1: PRISMA Flow Diagram

Geographic Distribution. The included studies demonstrated international scope, with representation from multiple continents:

- **US:** 9 studies
- **EU:** 5 studies (Netherlands: 2, Italy: 1, Switzerland: 1, Norway: 1)
- **AU:** 1 study
- **NZ:** 1 study

Notably, three of the US-based studies utilized the TRACK-TBI [1, 18, 19], indicating the significant overlap between these research platforms and our methodological approach.

TBI Severity Distribution. The reviewed studies covered the complete spectrum of TBI severities:

Table 2.2: Distribution of Studies by TBI Severity Focus

TBI Severity	**Studies (n)**
Mild TBI only	7
Moderate/Severe TBI only	2
Mixed TBI severities	7

Sample Sizes and Population Characteristics. Sample sizes varied considerably across studies, ranging from **139** participants in a pediatric sports concussion study [20] to **17,470** participants in the large multicenter moderate-to-severe TBI study [21]. The total combined sample across all studies was **36,607** participants; however, the number reflects the sum of reported sample sizes and should not be interpreted as representing unique individuals, as several studies relied on overlapping data sources such as TRACK-TBI, Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury, a prospective longitudinal observational study (CENTER-TBI) or Medical Information Mart for Intensive Care, a comprehensive longitudinal study at the Beth Israel Deaconess Medical Center ICU (MIMIC).

2.3.2 Study Designs

All presented studies employed longitudinal designs to capture temporal changes in outcomes, with follow-up intervals ranging from acute assessments (less than one month post-injury) to extended periods up to **5 years** post-injury. Most studies are prospective observational cohorts with repeated follow-ups [1, 18, 22–27]. These follow patients from weeks to years post-injury to characterize recovery in symptoms, function, or sleep using standardized measures over time.

Several works analyze existing data retrospectively, such as McGeown *et al.* (2019) [28], Gardner *et al.* (2019) [5] using Citicoline Brain Injury Treatment Trial, a Phase 3 randomized controlled study that investigated whether citicoline improved functional and cognitive outcomes after TBI (COBRIT) data, and Kowalski *et al.* (2021) [21] analyzing TBIMS; while Puffer *et al.* (2019) [29] conducted a secondary analysis of a Phase 3 RCT. These designs repurpose prospectively collected clinical trial or registry data to discover outcomes without new enrollment. Multiple studies draw from large multicenter platforms: TRACK-TBI (Martinez *et al.* (2024) [18], Ghaderi *et al.* (2024) [19], Wickwire *et al.* (2022) [1]), UPFRONT (De Koning *et al.* (2018) [23]). These cohorts are typically prospective at the platform level, even when the specific paper's analysis is retrospective.

2.3.3 Outcome Measures and Assessment Tools

Studies employed diverse outcome measures to assess various domains of TBI recovery:

Table 2.3: Primary Outcome Measures Utilized Across Studies

Outcome Domain	**Studies (n)**	**Primary Instruments**
Global Functional	5	GOSE
Functional Independence	3	FIM
Neuropsychiatric Symptoms	6	PHQ-9, BSI-18, PCL-5, ISI
Biomarker Trajectories	1	GFAP, NF-L, UCH-L1
Quality of Life	4	WHOQOL-BREF, SF-12
Symptom-Specific Scales	7	NSI, HISC, fatigue scales

2.3.4 Identified Trajectories

Across the analyzed literature, the number of trajectories identified varied considerably. Most studies employed data-driven methods such as latent class modeling or mixture modeling [1, 5, 18, 25, 27]. Other studies adopted cluster-based analytical approaches [19, 30]. Despite methodological diversity, the overall findings reveal a consistent pattern in which three to four trajectory classes are most frequently observed, representing the majority of empirical models in mild to severe TBI cohorts.

Several studies identified two major trajectories, typically distinguishing between groups with rapid versus prolonged recovery. For example, DeKoning *et al.* (2018) [23] differentiated between a "rapid recovery" profile characterized by high functional outcomes and minimal symptoms, and a "secondary complaints" group exhibiting elevated anxiety, depression, and delayed recovery. Similarly, Orr *et al.* (2021) [20] reported exercise-tolerant versus exercise-intolerant trajectories, where the latter showed symptom exacerbation and protracted recovery. These binary models tend to represent the most clinically intuitive division of post-concussive outcomes.

Studies identifying three trajectories commonly represented graded recovery patterns. Lu *et al.* (2018)[25] observed stable low, elevated good, and stable good motor recovery groups, ranging from minimal functional improvement to rapid independence. In parallel, Ghaderi *et al.* (2024) [19] characterized three phenotypic patterns of acute TBI: a mild phenotype (α) with the shortest Intensive Care Unit (ICU) stay and best outcomes, a moderate phenotype (γ) with intermediate recovery, and a severe phenotype (β) associated with the highest mortality. More complex trajectory models, incorporating four or more classes, captured additional heterogeneity across symptom domains and recovery rates. For instance, Rakers *et al.* (2021) [22] identified four fatigue–coping patterns after mild TBI, including two favorable recovery trajectories and two persistent fatigue profiles with or without psychological distress. Likewise, Molteni *et al.* (2021) [27] reported four long-term functional trajectories in pediatric acquired brain injury, ranging from fast responders to non-responders, while Puffer *et al.* (2019) [29] and McDonald *et al.* (2022) [26] demonstrated similar four-class structures representing gradations from full recovery to severe disability or death. Cognitive recovery models such as that of Lu *et al.* (2018) [25] revealed stable, delayed, and rapid improvement classes, further underscoring domain-specific variability in recovery processes.

Gardner *et al.* (2019) [5] identified seven distinct six-month functional recovery trajectories using LCMM. The trajectories ranged from good recovery to moderate disability and severe disability or death, illustrating pronounced heterogeneity in post-injury outcomes.

Beyond these, only a few studies identified highly granular trajectory structures. Wick-

wire *et al.* (2022) [1] delineated five sleep recovery patterns, distinguishing between persistent, resolving, and emergent insomnia types, while Martinez *et al.* (2024) [18] described six symptom trajectories across depression, anxiety, pain, and sleep dimensions, ranging from stable low to worsening profiles. The most detailed clustering was reported by Bui *et al.* (2021) [31], who identified eleven distinct biomarker trajectories across four blood biomarkers, reflecting flat versus rising biomarker patterns associated with cognitive and neurological outcomes.

2.3.5 Factors Associated with Recovery

Trajectory membership was associated with an interplay of demographic, injury-related, and psychological factors, reflecting observed patterns rather than causal mechanisms. Injury severity indicators such as Glasgow Coma Scale (GCS), Length of Stay (LOS), Post-Traumatic Amnesia (PTA) and imaging findings were consistently associated with slower recovery, whereas mild cases and shorter coma duration were associated with faster or improving trajectories [21, 27]. Demographic patterns revealed that younger age generally aligned with faster recovery, while female sex was overrepresented in fatigue, insomnia, or prolonged symptom trajectories. Education and socioeconomic position showed consistent associations, with lower education and enlisted rank linked to worse outcomes [22, 23, 26]. Race-specific effects were also identified, such as Black participants showing higher probabilities of persistent severe insomnia [1].

Psychological and behavioral variables played a major role in mild TBI outcomes. Passive coping, baseline anxiety or depression, sleep disturbance, and pain burden were associated with unfavorable or persistent trajectories, while active coping supported recovery [23, 24]. Mechanistic and exposure factors—such as repeated concussions, contact sports, or vestibulo-ocular symptom profiles—also stratified recovery patterns [20, 28, 31]. A summary of factors associated with trajectory membership is presented in Table 2.4.

Although the identified factors are described as associated with recovery trajectories, the analyses summarized here are observational and do not establish causality. Nevertheless, given consistent findings across multiple studies, some of these associations may reflect underlying causal mechanisms that warrant investigation in future longitudinal or interventional research.

Table 2.4: Factors Associated with Recovery

Category	Representative Findings
Demographic	
Age	Younger age was associated with faster or more complete recovery; older patients showed slower improvement, higher mortality, and persistent disability [5, 19, 21, 23, 26, 27, 29].
Sex	Female sex was overrepresented in unfavorable trajectories including fatigue, insomnia, and prolonged symptom recovery; males showed better functional change [1, 21, 22, 28, 31].
Education and Socioeconomic status (SES)	Lower education, enlisted military rank, or unemployment were linked to poorer recovery and greater disability; higher education associated with better outcomes [1, 22, 23, 26].
Race / Ethnicity	Black race associated with greater risk for persistent severe insomnia and lower odds of mild symptom resolution [1].
Injury Severity	
Initial severity	Lower GCS and longer LOS were associated with poorer functional trajectories; mild cases clustered into improving [18, 21, 25, 31].
Coma duration	Longer coma associated with slow or non-responder groups; shorter coma linked to faster recovery and higher functional scores [21, 27].
Etiology	Anoxic etiology linked to slower functional progress compared to traumatic causes [27].
Injury Mechanism	
Injury mechanism	High-velocity mechanisms (e.g., motor vehicle collisions) linked to persistent Disorders of Consciousness (DOC); low-velocity (e.g., falls) associated with better recovery [21].
Sports exposure	Repeated concussions, contact-sport injuries, and sport-type differences (contact vs. non-contact) distinguished biomarker clusters and recovery groups [31].
Symptom cluster	Vestibulo-ocular or cervicogenic clusters were associated with slower symptom resolution compared with physiological patterns [28].
Exercise response	Exercise tolerance was associated with faster recovery ($\sim$10 days); intolerance with prolonged recovery [20].

(continued) Summary of Factors Influencing Trajectory Membership Across Studies

Category	Representative Findings
Psychological and Behavioral	
Coping style	Passive or avoidant coping strongly associated with fatigue and emotional distress; active coping supported recovery [22–24].
Emotional distress	Baseline anxiety and depression were associated with poor recovery and higher risk for secondary complaints or persistent symptoms [1, 22, 23, 28].
Sleep and fatigue	Sleep disturbances before and after injury, together with fatigue and pain, were associated with persistent or worsening trajectories [1, 18, 22].
Psychiatric history	Pre-existing psychiatric disorders increased risk of severe insomnia and maladaptive recovery [1].
Initial symptom load	Higher baseline symptom scores were associated with prolonged recovery in sport-related mTBI [28].
Treatment	
LOS (ICU/rehab)	Longer hospital or rehabilitation stays were linked to more severe injuries and poorer functional outcomes [21, 25, 27].
Neurosurgical intervention	Decompressive craniotomy/neurosurgery was associated with slower functional gains and longer inpatient rehabilitation [27].
Duty Status	Return-to-duty evaluations and chronic post-injury time associated with higher symptom exaggeration scores [30].

2.3.6 Methodological Challenges and Adaptations

Studies identified a set of methodological challenges in trajectory modeling, reflecting both sample and data limitations as well as the intrinsic heterogeneity of the injury. Despite these barriers, many studies implemented a range of strategies to optimize modeling accuracy, handle missing data, and improve the generalizability of findings.

Sample Size and Representation Limitations. Small or imbalanced cohorts were among the most frequently reported challenges. Several analyses relied on relatively re-

stricted samples, such as military cadets or collegiate athletes [23, 31], or on single-site cohorts in rehabilitation settings [25, 27]. This limited the ability to generalize findings to the broader TBI population. To mitigate these constraints, investigators adopted data-driven approaches capable of identifying latent subgroups within modest sample sizes, such as LCGA [22, 26], LCMM [1, 5], and generalized structural equation modeling (SEM) [27]. Others used computationally efficient algorithms such as Locality-Sensitive Hashing combined with feature extraction via autoencoders to cluster multi-dimensional biomarker data [31].

Missing and Incomplete Data. Missing data were common in longitudinal research, often resulting from participant attrition, inconsistent follow-up, or incomplete records. Studies addressed these issues through a combination of imputation and modeling techniques. Random Forest (RF) imputation with conditional specification was used to maintain temporal structure in biomarker data [31]. Generalized Structural Equations Model (GSEM) was applied in growth and clustering models to account for incomplete follow-ups [25]. Algorithmic approaches such as the SLAC-Time model further mitigated missing values in time-series data [19]. Mixed linear models effectively handled missing repeated measures by allowing individual variation in trajectories [24], while latent class mixed models accommodated incomplete observations without excluding participants entirely [1, 5].

Methodological Biases and Data Quality. Several studies acknowledged potential biases arising from their design, such as selection bias (e.g., inclusion only of inpatients with rehabilitation referral) or measurement bias in self-reported data [21, 22]. Data quality concerns were also noted, including floor and ceiling effects in FIM score and missing longitudinal endpoints [25, 27]. Additional sources of bias were linked to incomplete follow-up or exclusion of cases with missing data [29]. To counter these limitations, researchers adopted multi-site data validation and cross-dataset harmonization [19], standardized clinical assessment protocols [21]. Some studies implemented cross-validation procedures and model comparison metrics (Akaike Information Criterion (AIC)/ Bayesian Information Criterion (BIC)) to verify trajectory stability and avoid overfitting [19, 27].

Demographic and Clinical Heterogeneity. While many studies emphasized the need for diverse populations, limited demographic variability remained a significant limitation [1, 23, 31]. To partially address this, stratified analyses were conducted by sex, education level, or psychiatric comorbidity [1, 22, 23]. In some cases, studies explicitly modeled demographic and clinical variables as covariates within latent class or regression frameworks [21, 26, 27]. However, authors consistently noted that such statistical adjustments cannot

fully compensate for structural underrepresentation of key subgroups, such as individuals from lower socioeconomic or ethnically diverse backgrounds [29, 31].

Computational and Multi-Domain Approaches. Beyond addressing traditional statistical issues, several studies introduced methodological innovations to better capture the complexity of TBI trajectories. These included joint time-series and phenotype modeling (SLAC-Time) to manage heterogeneous ICU datasets with missing temporal data [19]; integration of cognitive, emotional, and functional domains into multi-dimensional trajectory frameworks [22, 27]; and predictive or growth-based modeling frameworks designed to forecast individual recovery patterns [5, 25, 26]. Collectively, these methods demonstrate a gradual methodological shift from group-level analyses to individualized, precision-based modeling of TBI outcomes.

2.4 Discussion

This scoping review reveals several critical insights into the current state of TBI trajectory research and its implications for clinical practice. The convergence of findings across sixteen studies demonstrates that TBI progression follows multiple, distinct trajectories that can be reliably identified and characterized using specialized analytical approaches. All reviewed studies consistently identified multiple distinct outcome trajectories within seemingly homogeneous TBI populations. Even within specific severity categories, researchers consistently identified **3-11** distinct subgroups with markedly different recovery patterns. This finding fundamentally challenges traditional approaches that rely on average group outcomes and supports the growing recognition that TBI should be conceptualized as a heterogeneous collection of conditions rather than a single diagnostic entity [32].

To synthesize heterogeneous designs without overstatement, I present findings in two tracks. First, explicit trajectory-identification models consistently yielded discrete longitudinal classes (most commonly three to four), underscoring non-average recovery patterns across nominally similar TBI cohorts. Second, supporting and non-trajectory analyses contextualized or predicted outcomes without estimating time-parameterized classes, offering scalable validation and explanatory contrasts.

Three key findings emerge from this analysis. **First**, trajectory identification enables precision medicine approaches through early risk stratification using biomarker patterns [31], symptom profiles [1, 22], and imaging findings [29]. **Second**, methodological innovations including latent class growth analysis [26], growth mixture modeling [18], and machine learning clustering successfully address TBI heterogeneity despite persistent challenges in

sample size, missing data, and demographic diversity [21, 25]. **Third**, trajectory-specific interventions show promise for optimizing treatment intensity and resource allocation compared to severity-based approaches alone [24, 33].

Early Identification of High-Risk Trajectories. Across the reviewed studies, several early indicators of poor recovery were consistently observed. Biomarker analyses showed that early increases in GFAP and NF-L levels were associated with LOS, PTA [31]. Similarly, early symptom severity predicted longer recovery times—high initial fatigue or emotional distress were linked to persistent fatigue and distress trajectories over six months [22], while specific insomnia patterns predicted long-term sleep problems [1]. Demographic and clinical factors further shaped outcomes. Older age and pre-existing conditions were linked to poorer recovery and functional outcomes [5, 21, 25–27, 29], while female sex and psychiatric history were associated with higher fatigue and insomnia risk [1, 20, 22, 28].

These findings suggest that systematic screening protocols implemented in the acute phase could identify patients requiring intensive and specialized interventions to prevent progression to poor outcome trajectories. Indeed, a recent initiative by thought leaders in clinical TBI and research are advocating for new characterization schemes of acute TBI that incorporate clinical, biomarker, imaging, and modifier factors (CBI-M model) [32].

Analytical Approaches. Across the reviewed studies, the analytical strategies for trajectory identification have evolved from group comparisons toward more probabilistic and data-driven methods. Latent class–based approaches such as LCGA, GMM, and LCMM were the most widely applied to capture unobserved heterogeneity in recovery. However, they differ in flexibility and assumptions: LCGA treats individuals within a class as homogeneous, whereas GMM and LCMM incorporate random effects to capture within-class variation [34]. This allows LCMM to balance between population-level classification and individual-level variability, making it suitable for complex TBI data where inter-individual differences are substantial.

ML–based clustering techniques have also been used to complement these probabilistic approaches. For example, Bui *et al.* (2021) [31] applied Locality-Sensitive Hashing combined with autoencoder-based feature extraction to classify biomarker trajectories, while Ghaderi *et al.* (2024) [19] developed SLAC-Time to identify dynamic phenotypes without relying on imputation. These methods are well suited for high-dimensional datasets and can uncover non-linear patterns not captured by traditional models. However, their interpretability and clinical translation remain limited because clusters often lack direct biological meaning or clear correspondence to functional outcomes [34].

An emerging methodological direction involves multi-domain or multi-outcome modeling, in which cognitive, functional, emotional, and psychosocial domains are analyzed jointly rather than independently [22, 27]. These models recognize that recovery trajectories across domains are interrelated—for instance, fatigue and emotional distress often co-vary with cognitive slowing or sleep disruption. Integrating multiple outcomes into a unified latent framework provides a more ecologically valid representation of TBI recovery, aligning with the biopsychosocial model of brain injury. Importantly, this approach moves beyond single-dimensional outcomes toward identifying patient profiles that reflect complex real-world functioning.

Addressing Methodological Challenges. Although trajectory-based models have advanced substantially, several methodological challenges persist across the reviewed literature. One key issue is the variability in study design and data quality. Many investigations were based on relatively small or demographically narrow cohorts or single-center rehabilitation populations [25, 31], which limits generalizability. Missing longitudinal data remain another major concern; while some studies used random forest or multiple imputation techniques [19, 31], these rely on assumptions about missingness that are difficult to verify in clinical settings. Mixed-effects and LCMM frameworks mitigate this to some extent by accommodating irregular follow-up intervals [1, 24], yet the problem of attrition bias remains unresolved.

Another limitation lies in the heterogeneity of outcome measures. Different studies assessed recovery through diverse endpoints—ranging from functional scales to self-reported fatigue or sleep quality—making synthesis and meta-analysis difficult. The absence of standardized outcome batteries reduces comparability across models and impedes validation of derived trajectory classes. Finally, despite progress in analytic sophistication, external validation and clinical translation of trajectory findings remain limited.

Taken together, these observations highlight both the promise and the current limitations of trajectory modeling in TBI research. Statistical innovation has clearly improved the capacity to identify distinct recovery subgroups, but future progress will depend on harmonized data collection, multi-domain outcome integration, and prospective validation across diverse populations. Such advances will be essential for transforming trajectory research from descriptive insight into clinically actionable prediction.

2.4.1 Limitations and Future Research

Across the reviewed studies, several limitations and corresponding directions for future research are evident, reflecting the evolving methodological landscape of trajectory modeling.

A key limitation is the considerable **methodological heterogeneity** across studies, with diverse statistical and computational approaches—ranging from latent class and mixture models to unsupervised clustering, hindering direct comparison of findings. The absence of **standardized outcome measures** further limits cross-study replication, underscoring the need for unified longitudinal frameworks that capture cognitive, emotional, and functional domains [18, 22, 27]. Another recurring limitation concerns **population specificity and representativeness**: many studies relied on homogeneous samples or single-center rehabilitation cohorts [23, 25, 31], restricting generalizability to broader and more diverse TBI populations; expanding to under-represented minority.

Limited **follow-up duration** also limits understanding of chronic and late-emerging patterns, as cohorts often span months to one year [1, 22]. Moreover, while several studies proposed data-driven or machine-learning approaches to characterize heterogeneous recovery, their clinical function remains limited by the lack of external validation and translational frameworks [19]. Future work might prioritize developing and validating **predictive models** that support early identification and trajectory-informed decision-making [18, 21, 26, 27].

1. **Methodological Heterogeneity:** Diverse analytical approaches limit direct comparison across studies
2. **Population Specificity:** Many findings may not generalize beyond study-specific populations
3. **Follow-up Duration:** Limited long-term follow-up in many studies restricts understanding of chronic trajectories
4. **Outcome Measures:** Lack of standardization in outcome assessment limits synthesis

2.5 Conclusions

This scoping review employing trajectory-based approaches to TBI research reveals several fundamental insights that have significant implications for future research directions. The consistent identification of **3-11** distinct recovery trajectories across diverse populations reveals that average group responses inadequately capture the complexity inherent in TBI recovery [5].

The research community must prioritize three critical areas: standardization of trajectory assessment protocols to enable cross-study comparisons [30], multi-site collaborations for adequate sample sizes across trajectory types [20, 35], and investigation of biological mechanisms underlying distinct recovery patterns [31]. Implementation science research is essential for translating trajectory findings into routine clinical practice [4].

This evidence supports transforming TBI care from uniform approaches to precision medicine models that recognize recovery heterogeneity [16]. Success depends on developing practical clinical tools and protocols implementable across diverse healthcare settings to benefit the estimated 27-69 million individuals experiencing TBI annually worldwide [2, 3].

Chapter 3

Modeling Multi-Outcome Trajectories

Building upon the insights from the scoping review, this chapter focuses on the application of statistical and machine learning modeling techniques to identify multi-outcome trajectories using longitudinal data from the TRACK-TBI study.

Recent advances in trajectory modeling have demonstrated the clinical utility of identifying distinct recovery patterns following TBI. For instance, Gardner *et al.* (2019) [13] used LCMMs to identify seven distinct functional recovery trajectories over 6 months in the COBRIT, revealing dramatically divergent patterns ranging from stable good recovery to death. After **5** years moving forward in this direction, Martinez *et al.* (2024) [18] extended trajectory modeling to neuropsychiatric symptoms, identifying **three** distinct patterns (Stable-Low, Worsening, and Improving) across six symptom dimensions (Depression, Anxiety, Fear, Sleep, Physical, and **6** trajectories for Pain domain over 12 months post-TBI in the [15] cohort.

While these studies represent important growth in understanding recovery patterns, they have examined trajectories within single domains — either functional outcomes [13] or neuropsychiatric symptoms [18] — without simultaneously modeling trajectories across multiple outcomes representing multiple domains. This limitation restricts our understanding of how different aspects of recovery (cognitive, functional, and quality of life measures) may be interconnected and co-evolve over time.

3.1 Methods

In this section, an overview of the method is presented, including the approach, rationale, and step-by-step procedure. Figure 3.1 illustrates our main modeling pipeline, from the initial exploratory analysis of the TRACK-TBI to the determination of trajectories.

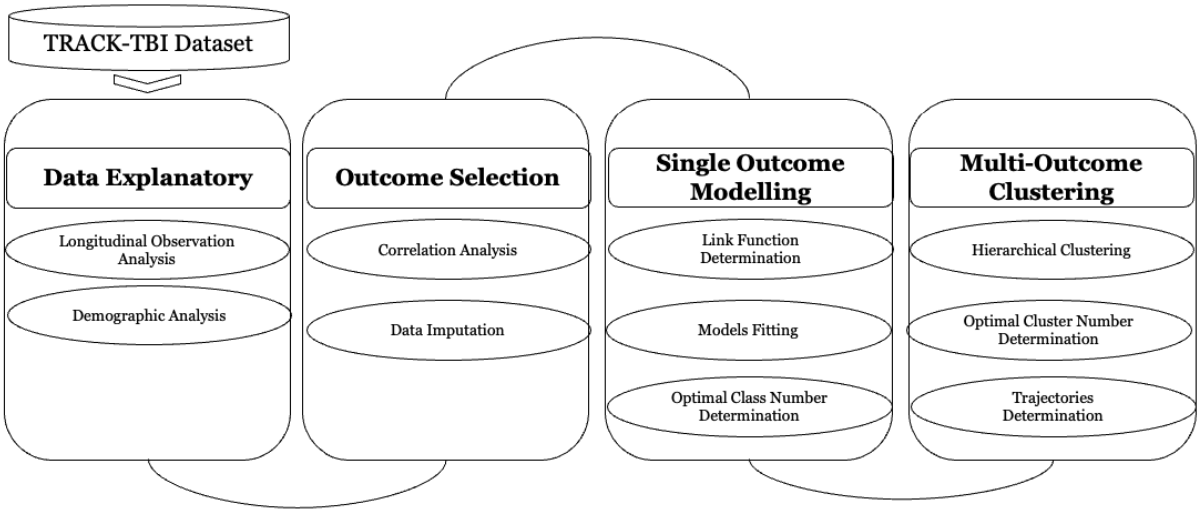


Figure 3.1: Modeling Pipeline Chart

3.1.1 Study Population and Data Source

This thesis utilized data from the TRACK-TBI, a prospective, multicenter observational cohort study conducted at *18 Level 1 trauma centers* across the United States [15]. The TRACK-TBI study was designed to create a comprehensive TBI Information Commons that integrates demographics, socioeconomics, clinical, imaging, and outcomes from over two thousand patients from acute hospitalization to the first year after injury.

The TRACK-TBI study enrolled participants between **February 26, 2014**, and **August 8, 2018**, with institutional review board approval from each participating site and written informed consent obtained from participants or their legally authorized representatives. The study represents one of the largest longitudinal prospective observational cohorts of acute TBI in North America to date, with standardized data collection protocols implementing the NIH TBI Common Data Elements (TBI-CDEs) framework. Participants were enrolled if they presented to the Emergency Department (ED) with a history of acute TBI

according to American Congress of Rehabilitation Medicine (ACRM) criteria, defined as a traumatically induced physiological disruption of brain function.

Samples used. Our working extract comprised **2,851** TRACK-TBI records. For baseline characterization, presented in Table 3.1, we report the **Core (Comprehensive Assessment (CA)) cohort** without controls with **complete ED GCS**, **n=2,293**. For longitudinal trajectory modeling, we applied predefined completeness criteria ($\geq$2 observations for $\geq$1 outcome across 4 visits), yielding the **analytic sample** n = 2,337 (with completed or not ED GCS). The counts, therefore, differ across sections by design; we state the relevant n in each analysis.

Study Cohorts and Data Collection Framework

The TRACK-TBI study employed a stratified enrollment design based on clinical care pathways rather than traditional severity classifications, facilitating comparative effectiveness research analyses:

1. **ED Group:** Patients evaluated and discharged from the ED.
2. **Hospital Admission (ADM) Group:** Patients admitted to the hospital but not to the ICU.
3. **ICU Group:** Patients admitted to the ICU.

Within each clinical group, participants were further stratified into assessment cohorts, which defined the extent of data collection: the Brief Assessment (BA) Cohort, CA Cohort, and Comprehensive Assessment plus Magnetic Resonance Imaging (CA+MRI) Cohort. An overview of the main cohort characteristics at study baseline is presented in Table 3.1. The data in this table include all subjects from the CA cohort, excluding the Friend Control (FC) and Orthopedic Control (OC) groups. Characteristics were calculated in *R* and sorted by GCS scores: mild (13–15), moderate (9–12), and severe (3–8). Subjects with missing or invalid GCS scores were excluded.

Table 3.1: Subjects Characteristics by ED GCS

Characteristic	Total (n = 2293)	Mild TBI (13–15) (n = 1924)	Moderate TBI (9–12) (n = 113)	Severe TBI (3–8) (n = 256)
Age (Mean ± SD)	40.13 ± 18.16 (n=2287)	40.43 ± 18.24 (n=1919)	37.36 ± 17.59 (n=113)	39.13 ± 17.75 (n=255)
Age - Missing (%)	6 (0.3%)	5 (0.3%)	0 (0%)	1 (0.4%)
Sex - Male (%)	1572 (68.6%)	1280 (66.5%)	91 (80.5%)	201 (78.5%)
Sex - Female (%)	721 (31.4%)	644 (33.5%)	22 (19.5%)	55 (21.5%)
Sex - Missing (%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Race - White (%)	1743 (76%)	1462 (76%)	85 (75.2%)	196 (76.6%)
Race - Black (%)	377 (16.4%)	318 (16.5%)	20 (17.7%)	39 (15.2%)
Race - Mixed Race (%)	41 (1.8%)	37 (1.9%)	2 (1.8%)	2 (0.8%)
Race - Asian (%)	83 (3.6%)	68 (3.5%)	5 (4.4%)	10 (3.9%)
Race - Indian (%)	7 (0.3%)	6 (0.3%)	0 (0%)	1 (0.4%)
Race - Race Unknown (%)	5 (0.2%)	4 (0.2%)	0 (0%)	1 (0.4%)
Race - Missing (%)	31 (1.4%)	24 (1.2%)	1 (0.9%)	6 (2.3%)
Ethnicity - Hispanic or Latino (%)	503 (21.9%)	413 (21.5%)	36 (31.9%)	54 (21.1%)
Ethnicity - Non Hispanic or Latino (%)	1758 (76.7%)	1487 (77.3%)	77 (68.1%)	194 (75.8%)
Ethnicity - Unknown (%)	9 (0.4%)	5 (0.3%)	0 (0%)	4 (1.6%)
Ethnicity - Missing (%)	23 (1%)	19 (1%)	0 (0%)	4 (1.6%)
Education (Mean ± SD)	13.1 ± 3.25 (n=2123)	13.24 ± 3.22 (n=1804)	12.24 ± 3.83 (n=102)	12.34 ± 3.07 (n=217)
Education - Missing (%)	170 (7.4%)	120 (6.2%)	11 (9.7%)	39 (15.2%)

Continued on next page

Table 3.1 – *Continued from previous page*

Characteristic	Total (n = 2293)	Mild TBI (13–15) (n = 1924)	Moderate TBI (9–12) (n = 113)	Severe TBI (3–8) (n = 256)
Psychiatric History (%)	487 (21.2%)	408 (21.2%)	19 (16.8%)	60 (23.4%)
Psychiatric History - No (%)	1805 (78.7%)	1515 (78.7%)	94 (83.2%)	196 (76.6%)
Psychiatric History - Missing (%)	1 (0%)	1 (0.1%)	0 (0%)	0 (0%)
TBI History - No (%)	1635 (71.3%)	1380 (71.7%)	77 (68.1%)	178 (69.5%)
TBI History - Yes ED Visit (%)	260 (11.3%)	230 (12%)	14 (12.4%)	16 (6.2%)
TBI History - Yes Hospital Admit (%)	164 (7.2%)	145 (7.5%)	6 (5.3%)	13 (5.1%)
TBI History - Missing (%)	234 (10.2%)	169 (8.8%)	16 (14.2%)	49 (19.1%)
Cause of Injury - Act of mass violence (%)	1 (0%)	1 (0.1%)	0 (0%)	0 (0%)
Cause of Injury - Incidental fall (%)	634 (27.6%)	536 (27.9%)	35 (31%)	63 (24.6%)
Cause of Injury - Other (%)	92 (4%)	81 (4.2%)	4 (3.5%)	7 (2.7%)
Cause of Injury - Other non-intentional injury (%)	125 (5.5%)	110 (5.7%)	2 (1.8%)	13 (5.1%)
Cause of Injury - Road traffic incident (%)	1279 (55.8%)	1077 (56%)	52 (46%)	150 (58.6%)
Cause of Injury - Suicide attempt (%)	2 (0.1%)	0 (0%)	0 (0%)	2 (0.8%)

Continued on next page

Table 3.1 – *Continued from previous page*

Characteristic	Total (n = 2293)	Mild TBI (13–15) (n = 1924)	Moderate TBI (9–12) (n = 113)	Severe TBI (3–8) (n = 256)
Cause of Injury - Violence/assault (%)	147 (6.4%)	113 (5.9%)	17 (15%)	17 (6.6%)
Cause of Injury - Missing (%)	13 (0.6%)	6 (0.3%)	3 (2.7%)	4 (1.6%)
Highest Level of Care - ED Discharge (%)	477 (20.8%)	473 (24.6%)	4 (3.5%)	0 (0%)
Highest Level of Care - Hospital admit no ICU (%)	844 (36.8%)	833 (43.3%)	9 (8%)	2 (0.8%)
Highest Level of Care - Hospital admit with ICU (%)	972 (42.4%)	618 (32.1%)	100 (88.5%)	254 (99.2%)
Highest Level of Care - Missing (%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Loss of Consciousness - No (%)	271 (11.8%)	257 (13.4%)	11 (9.7%)	3 (1.2%)
Loss of Consciousness - Suspected (%)	126 (5.5%)	103 (5.4%)	10 (8.8%)	13 (5.1%)
Loss of Consciousness - Unknown (%)	103 (4.5%)	85 (4.4%)	10 (8.8%)	8 (3.1%)
Loss of Consciousness - Yes (%)	1775 (77.4%)	1469 (76.4%)	82 (72.6%)	224 (87.5%)
Loss of Consciousness - Missing (%)	18 (0.8%)	10 (0.5%)	0 (0%)	8 (3.1%)
Posttraumatic Amnesia - No (%)	311 (13.6%)	295 (15.3%)	5 (4.4%)	11 (4.3%)

Continued on next page

Table 3.1 – *Continued from previous page*

Characteristic	Total (n = 2293)	Mild TBI (13–15) (n = 1924)	Moderate TBI (9–12) (n = 113)	Severe TBI (3–8) (n = 256)
Posttraumatic Amnesia - Suspected (%)	139 (6.1%)	68 (3.5%)	18 (15.9%)	53 (20.7%)
Posttraumatic Amnesia - Unknown (%)	272 (11.9%)	176 (9.1%)	20 (17.7%)	76 (29.7%)
Posttraumatic Amnesia - Yes (%)	1553 (67.7%)	1375 (71.5%)	70 (61.9%)	108 (42.2%)
Posttraumatic Amnesia - Missing (%)	18 (0.8%)	10 (0.5%)	0 (0%)	8 (3.1%)
GCS at ED Admission (Mean ± SD)	13.4 ± 3.32 (n=2293)	14.73 ± 0.53 (n=1924)	10.49 ± 1.13 (n=113)	4.71 ± 1.85 (n=256)

Data Collection Timeline

Data collection in the TRACK-TBI study followed a standardized timeline across multiple domains. During the acute phase, baseline demographics, medical history, and injury characteristics were recorded. Clinical brain Computed Tomography (CT) scans and any clinically indicated Magnetic Resonance Imaging (MRI) scans were obtained, and blood specimens were collected. Daily clinical monitoring was conducted for admitted patients, and high-resolution ICU monitoring data were recorded when applicable. After baseline, follow-up assessments were conducted at the following time points:

- **2 weeks:** Research MRI (CA+MRI cohort), comprehensive outcomes, blood draw
- **3 months:** Telephone-based outcome assessment
- **6 months:** Research MRI (CA+MRI cohort), comprehensive outcomes, blood draw
- **12 months:** Comprehensive outcome assessment

Data Access

Access to the de-identified TRACK-TBI dataset used in this study was granted through a formal collaboration request and a Data Transfer Agreement between the University of California, San Francisco, host of the TRACK-TBI database, and the health.data DRIVEN Lab directed by *Dr. Abel Torres Espín*. This agreement authorized the use of original or derivative data for research collected as part of the TRACK-TBI study.

3.1.2 Outcome Selection and Domain Grouping

For the purpose of this work, the selection of outcome variables requires a systematic multi-stage process designed to ensure comprehensive domain coverage and clinical relevance [36]. Specifically, we prioritized outcomes that were relatively well represented across the majority of time points, thereby maintaining sufficient data availability while reducing the extent of imputation required. Within each domain, we aimed to identify a limited set of representative outcomes that capture the central construct of that domain. This strategy serves two key purposes: first, it ensures balanced coverage of all relevant dimensions without over-representing any single domain; and second, it improves computational efficiency by minimizing redundancy among highly correlated outcomes, which is particularly important for the model fitting.

Selection Methodology

Outcome selection for multi-outcome trajectory analysis begins with findings from the literature review, which identified over **30** candidate outcome measures commonly employed in TBI trajectory research across **7** main domains (Domain Global Outcome, Domain Psychological Health, Domain Cognition, Domain Post-Concussive/TBI-Related Symptoms, Domain Sleep, Domain Pain, Domain Participation and Quality of Life). For measures not directly available in our dataset, we systematically identified appropriate similar measures to maintain conceptual and domain consistency. The selection process then proceeded through several refinement stages. First, we evaluated data availability and completeness across key time points (2 weeks, 3 months, 6 months, and 12 months after TBI), excluding outcomes with excessive missing data or inconsistent measurement that would compromise trajectory modeling. For instance, outcomes with more than *30%* overall missingness, or more than *30%* missingness at any individual time point, were generally excluded. However, exceptions were made for certain central outcomes (e.g. Trail Making Test (TMT) and Wechsler Adult Intelligence Scale (WAIS)), which were retained despite higher levels

of missingness given their role in representing special domain mechanisms. In these cases, missing values were planned to be imputed in the next step.

Second, we addressed redundancy through pairwise correlation analysis across all candidate outcome variables to identify highly correlated measures. For example, Participant Reported Outcome Measurement Information System (PROMIS) Pain Interference and Pain Intensity, two patient-reported outcomes on pain perception, showed a correlation of $r = 0.89$, which is expected since the pain items in PROMIS overlap in the construct they measure. In another example, the Brief Symptom Inventory 18 (BSI-18) sub-scales (anxiety and depression) and Brief Symptom Inventory 18 Global Severity Index (BSI-18 GSI) had $r > 0.99$. This was also expected as the individual items are summed to form the Global Severity Index. Therefore, only the sub-scales were maintained for analysis to capture different symptoms. Across all correlated outcome variables, we retained those variables that offered the optimal combination of conceptual clarity while reducing redundancies.

This process yielded a final selection of **17** outcome measures spanning six main domains of TBI outcomes and symptoms, as presented in Table 3.2. These measures provide comprehensive coverage while maintaining analytical integrity through sufficient data density across 4 time points. The next section provides an overview of these outcomes and their respective domains.

Table 3.2: Selected Outcome Variables by Domain

Domain	Outcome Variable	Name	Range
Global Outcome	GOSEScore	GOSE Score	1–8
	DRSPIScore	Expanded Disability Rating Scale Post-Acute Interview (DRS Score)	0–36
Psychological Health	BSI18AnxScoreRaw	Brief Symptom Inventory-18 Anxiety Subscale Raw Score (BSI-18 Anxiety Score)	0–24
	BSI18DeprScoreRaw	Brief Symptom Inventory-18 Depression Subscale Raw Score (BSI-18 Depression Score)	0–24

Continued on next page

Table 3.2 continued from previous page

Domain	Outcome Variable	Name	Range
	PCL5TotalScore	PTSD Checklist for DSM-5 Mental health score (PCL-5 Score)	0–80
Cognition	RAVLTSum	Rey Auditory Verbal Learning Test II Sum Score (RAVLT Score)	0–75
	TMTAScaledScore	Trail Making Test A Scaled Score (TMTA Attention Score)	1–15
	TMTBScaledScore	Trail Making Test B Scaled Score (TMTB Executive Score)	1–15
	WAISCoding	WAIS-IV Coding Processing Speed Raw Score (WAIS Coding)	0–75
	WAISSymbol	WAIS-IV Symbol Search Processing Speed Raw Score (WAIS Symbol)	0–75
Post-Concussive/TBI-Related Symptoms	RPQTotalScore	Rivermead Post-Concussion Questionnaire Total Score (RPQ Score)	0–64
	RPQ3Score	Rivermead Post-Concussion Questionnaire Subscale 3: Core Physical Symptoms (RPQ-3 Score)	0–12
	RPQ13Score	Rivermead Post-Concussion Questionnaire Subscale 13: Cognitive and Emotional Symptoms (RPQ-13 Score)	0–52
Sleep	ISITotalScore	Insomnia Severity Index (ISI Score)	0–28
Pain	PROMISIntensity	PROMIS Pain Intensity Pain Raw score (PROMIS Intensity)	0–10

Continued on next page

Table 3.2 continued from previous page

Domain	Outcome Variable	Name	Range
Participation and Quality of Life	SWLSTotalScore	Satisfaction With Life Scale Total Score (SWLS Score)	5–35
	QoLIBRIOSTotalScore	Quality of Life After Brain Injury Overall Scale (QoLIBRIOS Score)	1–5

Selected Outcomes and Groups

The selected outcomes were organized into seven distinct recovery domains, each addressing critical aspects of the patient's recovery.

Domain Global Outcome employs the GOSE Score, a widely used measure recommended by the International Mission for Prognosis and Analysis of Clinical Trials in TBI (IMPACT) consortium, alongside the DRS Score for comprehensive assessment of functional independence. Together, these two outcomes form a unified functional domain, which is utilized in the TRACK-TBI for domain grouping purposes.

Domain Psychological Health addresses the substantial psychological burden through three complementary measures: BSI-18 Anxiety Score and BSI-18 Depression Score for anxiety and depression symptoms, and PCL-5 Score for post-traumatic stress disorders.

Domain Cognition encompasses five measures targeting the most vulnerable cognitive domains: RAVLT Score for verbal memory, TMTA Attention Score, WAIS Coding, and WAIS Symbol for processing speed, and TMTB Executive Score for executive functioning and cognitive flexibility.

Domain Pain and Domain Sleep utilize single comprehensive measures: ISI Score addresses sleep disturbances, while PROMIS Intensity captures pain experiences impacting recovery trajectories.

Domain Post-Concussive/TBI-Related Symptoms employs three correlated yet clinically distinct measures from the Rivermead questionnaire: RPQ-3 Score for immediate physical symptoms, RPQ-13 Score for cognitive-emotional symptoms, and RPQ Score for comprehensive symptom burden.

Domain Participation and Quality of Life combines generic (SWLS Score) and TBI-specific (QoLIBRIOS Score) assessments to capture both general life satisfaction and

injury-specific quality of life dimensions.

3.1.3 Missing Data Analysis

Despite careful selection of outcome measures emphasizing those with higher data availability, missing data and patient attrition remain inherent challenges in longitudinal studies, especially with TBI population [37]. Properly addressing these issues is vital to ensure valid trajectory modeling. Accordingly, we examined patterns of missingness and attrition mechanisms and implemented strategies to handle incomplete observations.

Subjects Inclusion and Exclusion

To support our modeling approach, which requires repeated assessments across multiple time points, we established minimum data completeness criteria. Patients were included in the final analytical sample if they had at least **two** non-missing observations for at least **one** outcome variable across the **four** follow-ups (2 weeks, 3 months, 6 months, and 12 months post-injury).

Of the initial **2,851** records in our extract, **2,337** patients (82.0%) met the longitudinal completeness criterion and entered the analytic sample (**514** (18.0%) were excluded).

Table 3.3 presents the data availability by outcome variable for patients meeting inclusion criteria. The number of patients with sufficient data varied across outcome domains, ranging from **1,542** patients for WAIS Coding to **2,154** patients for DRS Score. Cognitive assessments generally had lower data availability (**1,542-1,567** patients) compared to questionnaire-based measures (**2,023-2,057** patients), which may reflect the more intensive administration requirements for neuropsychological testing.

Attrition Patterns

To understand the mechanisms underlying patient attrition and assess potential bias in missing data patterns, we conducted an analysis of dropout reasons using TRACK-TBI administrative data. Attrition information was extracted from multiple sources including death records, consent withdrawal documentation, and follow-up contact logs.

A classification algorithm was developed to categorize attrition reasons based on text mining of administrative records and structured data fields. Keywords were identified for different categories including administrative issues, health-related problems, financial

Table 3.3: Data Availability by Outcome Variable for Included Subjects

Outcome	**N ≥ 2 Obs**
DRS Score	2,154
RPQ-13 Score	2,057
RPQ-3 Score	2,057
RPQ Score	2,057
QoLIBRIOS Score	2,052
BSI-18 Anxiety Score	2,049
BSI-18 Depression Score	2,049
ISI Score	2,042
PROMIS Intensity	2,042
SWLS Score	2,039
PCL-5 Score	2,023
GOSE Score	1,979
RAVLT Score	1,567
TMTA Attention Score	1,559
TMTB Executive Score	1,547
WAIS Symbol	1,544
WAIS Coding	1,542

concerns, participant burden, relocation, loss of interest, and family decisions. Table 3.4 presents the distribution of attrition reasons across the study population.

The majority of patients (*85.5%*; denominator = initial extract $n = 2{,}851$) had undetermined status, indicating they remained in the study through their intended follow-up period or the drop-out reason could not be confirmed. Among patients with documented attrition, death was the most common reason (**176** patients, *5.3%*), followed by voluntary withdrawal due to participant burden or time constraints (**62** patients, *1.9%*).

The **176** deceased patients represent a significant clinical subgroup requiring special consideration in trajectory modeling. Death following TBI typically occurs among patients with more severe injuries and represents a competing risk that could bias trajectory estimates if not appropriately handled.

For trajectory modeling, deceased patients contributed data up to the point of death, with subsequent time points treated as missing. This approach assumes that the mechanisms leading to death are related to TBI severity and baseline characteristics rather than unobserved factors affecting trajectory outcomes, supporting the Missing Not at Random

Table 3.4: Distribution of Subjects Attrition Reasons

Attrition Reason	**Number of Patients**
Deceased	176
Withdrawal (Participant Burden/Time)	62
Withdrawal (Reason Unspecified)	60
Withdrawal (Loss of Interest)	51
Withdrawal (Health-Related)	36
Lost to Follow-up	35
Withdrawal (Family Decision)	21
Withdrawal (Logistics/Relocation)	14
Administrative Reason	13
Withdrawal (Financial)	5

(MNAR) assumption for this subgroup.

The pattern of missing data across different outcome domains and patient characteristics suggests a combination of Missing at Random (MAR) mechanisms. Cognitive assessments showed higher rates of missing data, likely reflecting the increased burden of in-person testing and potential cognitive impairment. Questionnaire-based measures had more complete data, suggesting that the administration method influences missing data patterns.

To assess potential bias from missing data, we compared baseline characteristics between patients included and excluded from the trajectory analysis. Excluded patients had similar demographic profiles but slightly higher rates of older patients who admitted to ICU, suggesting that missing data may be related to injury severity or functional status. The demographic table presented in Appendix Table 7.

Implications for Trajectory modeling

Even in the presence of missing data, trajectory modeling approaches can exploit correlations among outcomes to enhance estimation efficiency and reduce bias under MAR assumptions [38]. The relatively high retention rate (*82.0%*) and diverse reasons for attrition suggest that missing data patterns are largely explicable by observable factors including injury severity, functional status, and study participation burden.

The predominance of undetermined (no attrition) status (*85.5%*) indicates that most patients remained engaged throughout their intended follow-up period, supporting the

validity of estimated trajectories for the study population. However, the differential missing data rates across outcome domains necessitate careful interpretation of domain-specific trajectory patterns and cross-domain comparisons.

Missing Data Patterns

From the previous analysis we already know the dataset comprised **2,337** subjects assessed across four time points. This design generated a theoretical maximum of **9,348** observations per outcome variable. However, the analysis of missing data revealed complex patterns of missingness that necessitated specific handling strategies.

Initial data preprocessing involved transformation from long format (*SubjectID, TimePoint, Variables*) to wide format, where each outcome variable was expanded across time points with systematic naming conventions to facilitate the Multiple Imputation by Chained Equations (MICE) algorithm's requirement for matrix-based imputation.

Preliminary analysis identified five variables with complete missingness (*100%*) at the 3-month follow-up: RAVLT Score, TMTA Attention Score, TMTB Executive Score, WAIS Coding, and WAIS Symbol. These variables represented systematic per-study data collection missingness at this specific time point and were excluded from the imputation matrix to prevent algorithmic instability. In other words, including variables with entirely missing values could cause errors or unreliable results in the imputation procedure.

Table 3.5 presents the missing data analysis for the **20** variables with the highest missingness percentages.

Analysis revealed **two** primary missing data mechanisms:

1. **Progressive Attrition.** Missing data increased with follow-up duration, with 12-month assessments showing *28-44%* missingness compared to *27-33%* at earlier time points.

2. **Mortality-Related Censoring. 176** subjects deceased during follow-up (excluding these who were initially identified as died deceased by the algorithm), creating structured missingness for all post-mortality time points.

Imputation Framework

Previous work has established recommendations for handling missing data for TBI population. For instance, Nielson *et al.* (2021) [38] suggest using multiple imputation for data

Table 3.5: Variables with Highest Missing Percentages

Variable	Missing Count	Missing %
TMTA Attention Score 12 Month Follow Up	1027	44.0
WAIS Coding 12 Month Follow Up	1027	44.0
WAIS Symbol 12 Month Follow Up	1026	43.9
TMTB Executive Score 12 Month Follow Up	1022	43.7
RAVLT Score 12 Month Follow Up	1018	43.6
WAIS Coding 6 Month Follow Up	794	34.0
WAIS Symbol 6 Month Follow Up	793	33.9
TMTB Executive Score 6 Month Follow Up	787	33.7
TMTA Attention Score 6 Month Follow Up	781	33.4
RAVLT Score 6 Month Follow Up	769	32.9
WAIS Coding 2 Week Follow Up	768	32.9
WAIS Symbol 2 Week Follow Up	762	32.6
TMTB Executive Score 2 Week Follow Up	755	32.3
TMTA Attention Score 2 Week Follow Up	735	31.4
RAVLT Score 2 Week Follow Up	720	30.8
GOSE Score 12 Month Follow Up	676	28.9
PCL-5 Score 12 Month Follow Up	673	28.8
ISI Score 12 Month Follow Up	653	27.9
BSI-18 Anxiety Score 12 Month Follow Up	651	27.9
BSI-18 Depression Score 12 Month Follow Up	651	27.9

that are MAR or Missing Completely at Random (MCAR). Similarly, To *et al.* (2025) [39] demonstrated that incorporating contemporaneous outcomes into the imputation model substantially improves accuracy of imputations for the GOSE Score, underscoring the importance of using both longitudinal and cross-sectional information. Given these, we implemented a multiple imputation framework using MICE with RF or Predictive Mean Matching (PMM) approaches depending on the type of the variable to be imputed as explained below.

We selected imputation methods according to variable characteristics. PMM was used for continuous outcomes with limited variance to preserve distributional properties, while RF was applied to high-variance continuous outcomes. This automated procedure classified **45** variables (*71.4%*) for RF imputation and **18** variables (*28.6%*) for PMM, balancing flexibility across different measurement types as shown in Table 3.6.

Table 3.6: Distribution of Imputation Methods

Method	**N Variables**	**Percentage**	**Variable Types**
RF	45	71.4%	High-variance
PMM	18	28.6%	Low-variance
Total	**63**	**100%**	**All variables**

Imputation was conducted at the participant-time point level, treating each subject's outcomes at each assessment as independent observations, consistent with the participant-to-time point approach suggested by To *et al.* (2025) [39]. Each variable was assigned an imputation role in a predefined plan: *target* variables (to be imputed), *predictors* (used for imputation but not imputed themselves), *predictor-for-all* variables (time-invariant covariates such as demographics included in every model), and *not used*.

A predictor matrix was designed to integrate multiple layers of information, considering empirical demonstration by To *et al.* (2025) [39] that similar outcomes can enhance imputation quality. The predictor categories were defined:

- **Longitudinal predictors:** the same variable measured at different timepoints, included if its correlation with the target exceeded $r \geq 0.30$;
- **Cross-sectional predictors:** other outcomes assessed at the same timepoint, again filtered by correlation threshold;

To avoid overfitting, the number of predictors per target variable was capped at ten, prioritizing those with the strongest correlations. The MICE procedure was run with five imputations ($m = 5$) and twenty iterations per chain, producing five completed datasets which we applied to validation in the next step.

Validation Framework

We implemented a validation approach in which *20%* of observed values were randomly masked and subsequently imputed. For variables with at least **20** observed data points, validation metrics were calculated across all five imputations and averaged. To assess imputation performance, we employed three complementary metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Pearson correlation.

The results demonstrate varying imputation accuracy across different outcomes and time points. For example, QoLIBRIOS Score demonstrated exceptional imputation accuracy with RMSE values between *0.896-1.072* and correlations of *0.490-0.674*. This high performance likely reflects the stability and internal consistency of quality of life self-reports.

On the other hand, GOSE Score showed strong imputation performance (RMSE: *1.296-1.920*, correlations: *0.657-0.867*), indicating that functional disability patterns are well-captured by the predictor variables. The completed imputation process generated **five** multiple-imputed datasets, each comprising **2,337** subjects, and a total of **9,348** observations per dataset. For the modeling, **one** of these completed datasets was randomly selected for analysis. The final dataset was then converted back to long format for longitudinal trajectory modeling, ensuring complete data coverage for all the next modeling stage.

3.1.4 Single-outcome Modeling

LCMM Foundation

The LCMMs is a form of finite mixture model that provides a statistical framework for heterogeneous longitudinal data by addressing two key sources of variability: **1)** between-subject heterogeneity via latent subgroups with distinct trajectory patterns, and **2)** within-subject variability via individual-specific random effects. This modeling approach enables the discovery of meaningful subpopulations while accounting for individual deviations from class-specific trajectories. The LCMMs was implemented using the `lcmm` package (v.2.2.1) in R, following the framework established by Proust-Lima *et al.* (2017) [14].

To explain LCMMs, we first introduce linear mixed models and latent process mixed models briefly, and then describe how LCMMs extends this framework.

Let Y_{ij} denote the observation for subject i at time t_{ij}, with $i = 1, \ldots, N$ and $j = 1, \ldots, n_i$. Following Laired and Ware (1982) [40], the linear mixed model (LMM) can be defined as:

$$Y_{ij} = \boldsymbol{X}_{ij}^{\top}\boldsymbol{\beta} + \boldsymbol{Z}_{ij}(t_{ij})^{\top}\boldsymbol{u}_i + w_{ij}(t_{ij}) + \varepsilon_{ij},$$

where $\boldsymbol{X}_{ij}$ and $\boldsymbol{Z}_{ij}$ are covariate vectors for fixed and random effects (which typically includes functions of time) at time t_{ij}, respectively. The random effects $\boldsymbol{u}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{B})$ capture subject-specific deviations, while $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ are independent measurement errors. The additional term $w_{ij}(t_{ij})$ represents the value of a subject-specific stochastic zero mean Gaussian process evaluated at time t_{ij}.

The LMM applies to longitudinal outcomes that are continuous and assume Gaussian random deviations. To study non-Gaussian longitudinal outcomes, we adopt the latent process mixed models approach [14]. Rooted in the latent variable framework, this methodology separates the structural model—describing the underlying quantity of interest (a latent process) as a function of time and covariates—from the measurement model, which links this latent quantity to the observations. A latent process $\Lambda_i(t)$ is defined in continuous time according to a standard linear mixed model

$$\Lambda_i(t) = \boldsymbol{X}_i(t)^{\top}\boldsymbol{\beta} + \boldsymbol{Z}_i(t)^{\top}\boldsymbol{u}_i + w_i(t),$$

where $\boldsymbol{X}_{ij}$ and $\boldsymbol{Z}_{ij}$ and $w_i(t)$ are as defined in the equation of linear mixed models. At time t_{ij}, the observed data are then linked to the latent process via a measurement model $Y_{ij} = H(\Lambda_i(t_{ij}) + \varepsilon_{ij}; \eta)$, where H is the parameterized link function.

Latent Class Linear Mixed Models. The linear mixed model assumes that the population of N subjects are homogeneous. To account for heterogeneity in the population, the LMM can be extended to a latent class mixed model (LCMMs), in which we assume that the population is composed of G latent classes of subjects, each characterized by its own mean trajectory profile. The LCMMs introduces unobserved heterogeneity through a latent categorical variable $C_i \in \{1, \ldots, G\}$, indicating the class membership of subject i. Class membership probabilities can be modeled as,

$$P(C_i = g \mid \boldsymbol{X}_{ci}) = \frac{\exp(\boldsymbol{X}_{ci}^{\top}\boldsymbol{\xi}_g)}{\sum_{l=1}^{G}\exp(\boldsymbol{X}_{ci}^{\top}\boldsymbol{\xi}_l)}.$$

Latent Class Mixed Models. The extension of latent class linear mixed models also applies to the latent process mixed models. Conditional on $C_i = g$, subject i follows a class-specific latent process mixed model where

$$\Lambda_i(t) \mid_{C_i=g} = \boldsymbol{X}_{ci}(t)^{\top}\boldsymbol{\beta}_g + \boldsymbol{Z}_i(t)^{\top}\boldsymbol{u}_{i,g} + w_i(t_{ij}),$$

with $\boldsymbol{u}_{i,g} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{D}_g)$ and $w_i(t)$ a class-specific Gaussian process. Thus, LCMMs could capture both subject-specific variability (via random effects and stochastic processes) and population-level heterogeneity (via latent classes). According to Proust-Lima *et al.* (2017) [14], the parameters of the LCMMs can be estimated via maximum likelihood using an extended Marquardt algorithm. We utilized a simplified model structure of the LCMMs, where time is the only covariate included.

Modeling Framework

Our modeling approach followed a systematic three-phase process, broadly consistent with the eight-step framework for latent class trajectory modeling described by Lennon *et al.* (2018) [41], as well as with further methodological guidance on mixture model specification and class enumeration provided by van der Nest *et al.* (2020) [42]. In the first phase, the appropriate link function was specified and evaluated to ensure an adequate representation of the outcome's functional form (details are provided in Section 3.1.4). The second phase involved systematic model building, including class enumeration, to identify the number of latent trajectories that best captured population heterogeneity details are provided in Section 3.1.4. Finally, optimal model selection was conducted based on multiple statistical fit indices (e.g., AIC, BIC, etc.), in line with established recommendations for mixture modeling and latent trajectory analysis [42] (details are provided in Section 3.2.1). Each model specified the outcome as a function of numeric time points, with both fixed and mixture formulas incorporating time. Model estimation was performed using the `lcmm` package in `R`, with a maximum of **500** iterations and a default convergence tolerance of 1×10^{-7}, applying parallel processing across six cores within the ARM architecture.

Link Function

Within the latent process mixed model framework proposed by Proust-Lima *et al.* (2017) [14], the **link function** defines the measurement component of the model—that is, the transformation that connects the latent continuous process to the observed outcome scale. Formally, for subject i at time t_{ij}, the observation Y_{ij} is assumed to arise from a monotonic transformation of the underlying latent process $\Lambda_i(t_{ij})$ with an additive measurement error ε_{ij}:

$$Y_{ij} = H\big(\Lambda_i(t_{ij}) + \varepsilon_{ij}; \boldsymbol{\eta}\big),$$

where $H(\cdot; \boldsymbol{\eta})$ is a parametric or semi-parametric *monotonic* function parameterized by $\boldsymbol{\eta}$.

The role of the link function is to adapt the model to the measurement properties of the outcome. For normally distributed continuous outcomes, a linear link corresponds to

the traditional linear mixed model. When outcomes exhibit non-Gaussian distributions, bounded scales, or ordinal structures, alternative link specifications (e.g., splines, beta, or threshold functions) allow for flexible, non-linear mappings between the latent and observed scales, improving fit and interpretability.

For each of the selected outcome variables, we systematically evaluated multiple link function specifications to identify the optimal functional form that best transforms non-linearity:

- **Linear link:** Standard linear relationship between time and outcome, assuming constant rate of change
- **Beta link:** Appropriate for bounded continuous outcomes with natural [0,1] constraints after scaling
- **Quantile splines (2-6 knots):** Flexible piecewise polynomials with knots placed at quantiles of the time distribution
- **Equidistant splines (2-6 knots):** Piecewise polynomials with equally spaced knots.
- **Threshold link:** Piecewise-constant function based on a latent continuous variable, where ordered categories are defined by estimated thresholds

Link function selection followed a hierarchical approach, prioritizing:

1. **Model Convergence:** Only models that achieved stable convergence with reliable parameter estimates were retained. For example, the threshold link function was excluded early in the process, as it required more than the defined number of iterations without reaching stable coverage.
2. **Information Criteria:** Primary selection based on AIC, Log-likelihood and BIC. Information criteria are presented in Appendix Table 7.
3. **Parsimony:** When AIC / BIC values were similar, simpler models were preferred to avoid over-fitting
4. **Graphical Assessment:** Visual inspection of fitted curves and residual plots to ensure appropriate capture of temporal patterns. All link fit plots are presented in Figure 3.2

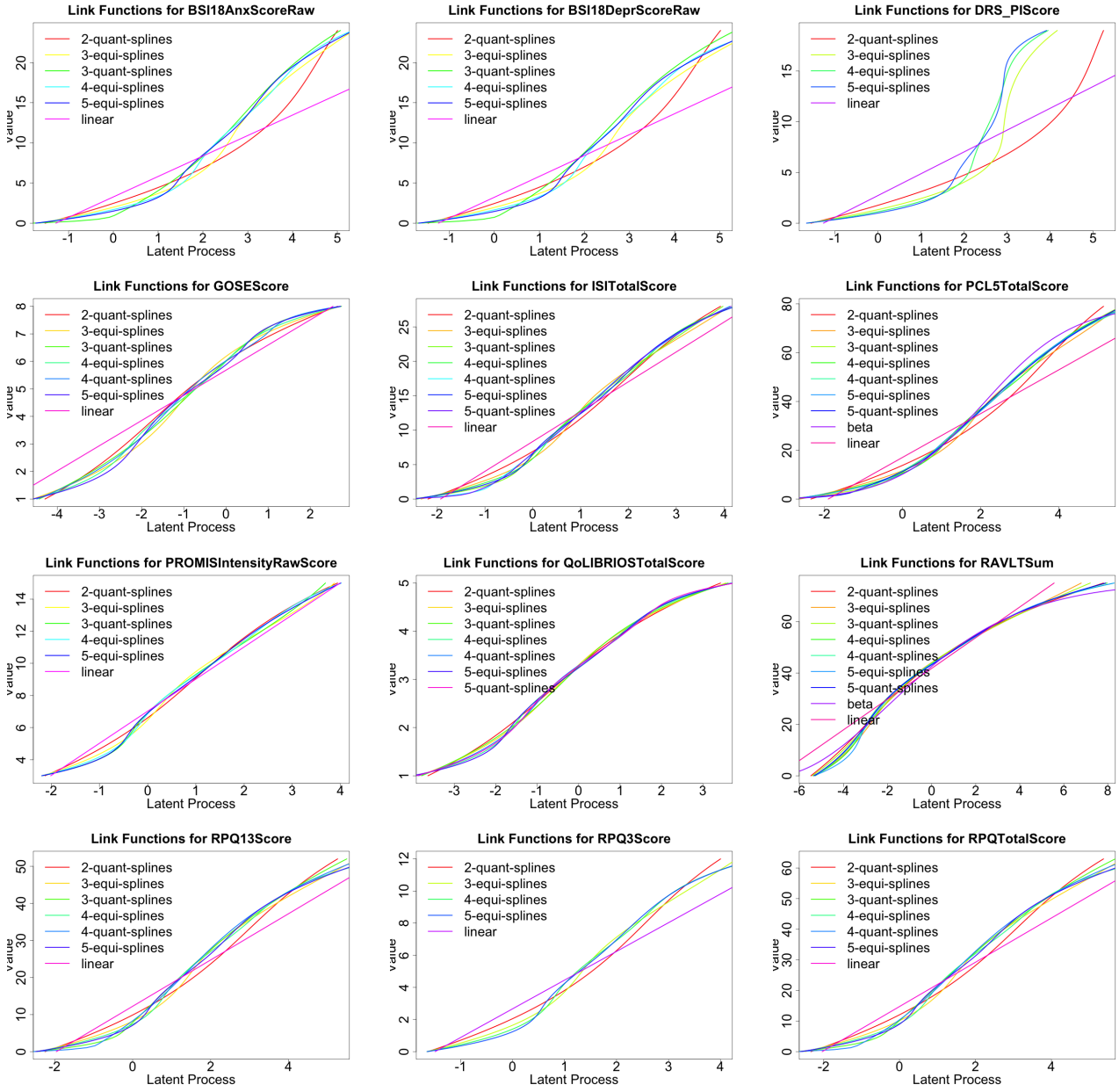
5. **Computational efficiency:** Iteration counts were also considered. In cases where models demonstrated similar statistical performance and comparable fit diagnostics, functions such as the beta link were deprioritized due to their substantially higher computational demands, which offered no practical advantage over more efficient alternatives.

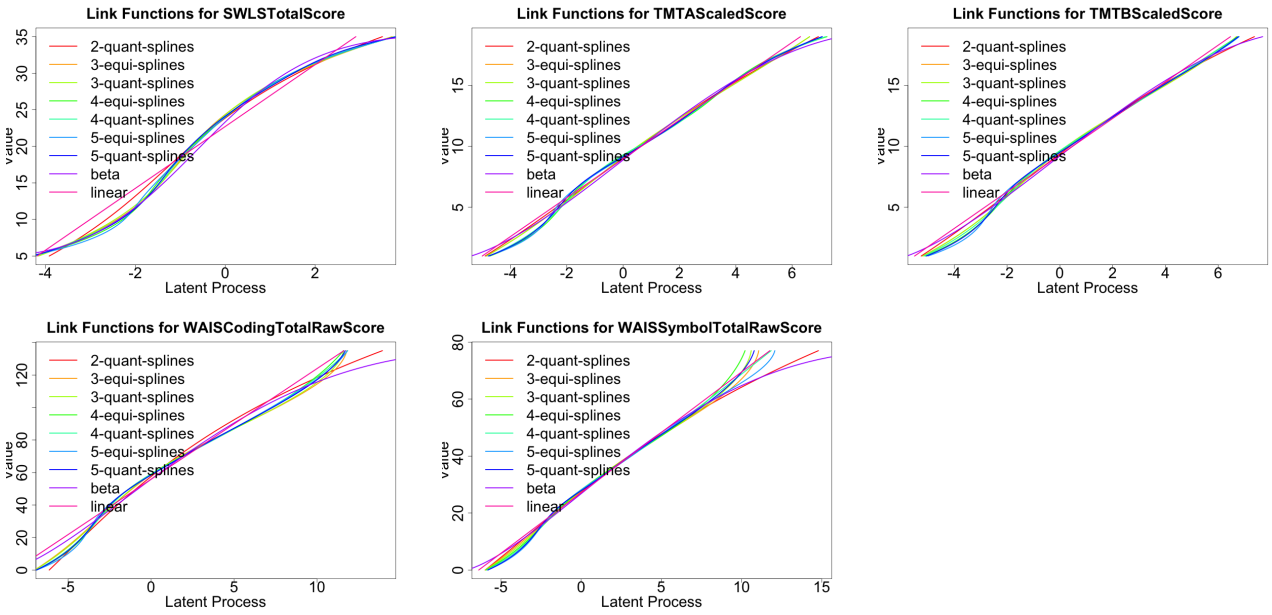
The link function selection process thus underscored the heterogeneity of recovery dynamics across different outcomes. These illustrate the importance of tailoring functional forms to the characteristics of each outcome and provide a methodological foundation for the subsequent modeling stage. Selected link functions are presented in Table 3.7.

Table 3.7: Selected Link Functions

Outcome Variable	**Link Function Type**
BSI-18 Anxiety Score	3-quantile splines
BSI-18 Depression Score	3-quantile splines
DRS Score	5-equidistant splines
GOSE Score	4-quantile splines
ISI Score	5-quantile splines
PCL-5 Score	5-quantile splines
PROMIS Intensity	3-quantile splines
QoLIBRIOS Score	5-quantile splines
RAVLT Score	4-quantile splines
RPQ-13 Score	4-quantile splines
RPQ-3 Score	5-equidistant splines
RPQ Score	4-quantile splines
SWLS Score	5-equidistant splines
TMTA Attention Score	5-equidistant splines
TMTB Executive Score	5-equidistant splines
WAIS Coding	5-equidistant splines
WAIS Symbol	5-equidistant splines

Figure 3.2: Link Functions Plots


Link Functions for BSI18AnxScoreRaw
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
5-equi-splines
linear
value
Latent Process
Link Functions for BSI18DeprScoreRaw
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
5-equi-splines
linear
value
Latent Process
Link Functions for DRS_PIScore
2-quant-splines
3-equi-splines
4-equi-splines
5-equi-splines
linear
value
Latent Process
Link Functions for GOSEScore
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
4-quant-splines
5-equi-splines
linear
value
Latent Process
Link Functions for ISITotalScore
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
4-quant-splines
5-equi-splines
5-quant-splines
linear
value
Latent Process
Link Functions for PCL5TotalScore
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
4-quant-splines
5-equi-splines
5-quant-splines
beta
linear
value
Latent Process
Link Functions for PROMISIntensityRawScore
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
5-equi-splines
linear
value
Latent Process
Link Functions for QoLIBRIOSTotalScore
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
4-quant-splines
5-equi-splines
5-quant-splines
value
Latent Process
Link Functions for RAVLTSum
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
4-quant-splines
5-equi-splines
5-quant-splines
beta
linear
value
Latent Process
Link Functions for RPQ13Score
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
4-quant-splines
5-equi-splines
linear
value
Latent Process
Link Functions for RPQ3Score
2-quant-splines
3-equi-splines
4-equi-splines
5-equi-splines
linear
value
Latent Process
Link Functions for RPQTotalScore
2-quant-splines
3-equi-splines
3-quant-splines
4-equi-splines
4-quant-splines
5-equi-splines
linear
value
Latent Process



Class Enumeration

Following link function selection, we employed a systematic approach to determine the optimal number of classes (enumeration) for each outcome using a structured model-building process:

Algorithm 1 Model Building Process

1: **Phase 1:** Fit 1-class model (init model)
2: **Phase 2:** Sequentially fit 2–10 class models
3: **for** each class number $g \in \{2, 3, \ldots, 10\}$ **do**
4: Initialize with 1-class model parameters
5: Fit model with g classes using optimized parameters
6: Assess convergence and model diagnostics
7: Store model statistics (AIC, BIC, entropy, class proportions)
8: **end for**
9: **Phase 3:** Select optimal model using information criteria

3.1.5 Clustering Framework

Following the application of LCMMs per outcome, we implemented a hierarchical clustering approach to identify multi-outcome trajectories. This method utilizes the posterior probabilities from the LCMMs to group patients based on their complex trajectories.

Hierarchical clustering using the Ward.D2 algorithm, which is particularly well-suited for high-dimensional data. Hierarchical clustering creates a dendrogram that visually represents the division of data into clusters at different levels of aggregation, allowing for flexible interpretation of patient groupings [43]. The Ward.D2 algorithm minimizes the within-cluster sum of squares at each merging step, ensuring compact and homogeneous clusters [44]. This method is optimal for probability data due to its interpretability, provided by the clear dendrogram, and its flexibility in choosing the number of clusters, allowing the clustering to be adapted to the specifics of the data.

Ward's Criterion: Minimizes within-cluster sum of squares at each merge step:

$$\Delta(A, B) = \frac{|A||B|}{|A| + |B|} ||m_A - m_B||^2$$

where $|A|$ and $|B|$ are cluster sizes and m_A, m_B are cluster centroids.

Before clustering we extracted all posterior probabilities from the selected LCMMs which resulted in a high-dimensional probability vector for creating the probability space for the clustering approach. Euclidean distance was selected as the distance metric, which is standard practice for biomedical data and works optimally with the Ward.D2 algorithm.

Selecting the optimal number of clusters is critical for meaningful trajectory interpretation. We implemented a multi-criteria evaluation framework to ensure robust model selection, systematically evaluating cluster numbers from $k = 2$ to $k = 10$ using multiple complementary criteria to ensure comprehensive assessment.

The primary statistical measure employed in our analysis was the Within-Cluster Sum of Squares (WCSS), which directly reflects cluster compactness; lower values indicate reduced intra-cluster dispersion and provide insight into the density of the formed clusters [44]. In addition to WCSS, we also considered the Silhouette Coefficient (SC), which measures within-cluster cohesion versus between-cluster separation, with values closer to 1 indicating well-formed clusters [45]. While informative, this metric served primarily as a supporting measure, as it is most suitable for elliptical-shaped clusters and may not fully capture the heterogeneous nature of our data. The Calinski–Harabasz-index (CHI) was also evaluated to assess the ratio of between-cluster dispersion to within-cluster dispersion, helping to confirm the stability of the chosen clustering configuration.

Beyond evaluation metrics, we incorporated clinical meaningfulness criteria to ensure practical applicability of our clustering solution. The coefficient of variation in cluster sizes was calculated to ensure balanced groupings across different patient subpopulations. We also monitored minimum and maximum cluster sizes to guarantee sufficient sample sizes for meaningful clinical interpretation and to prevent the formation of clusters too small for reliable statistical analysis or clinical conclusions.

The final component of our evaluation framework involved a comprehensive visual inspection and structural analysis. Direct trajectory visualization allowed us to examine patterns across all outcome domains, ensuring that identified clusters represented distinct multi-outcome trajectories. Dendrogram analysis provided insights into the hierarchical structure and helped identify natural cluster boundaries that aligned with the underlying data structure. This visualization-based approach was particularly crucial in our decision-making process, as changes in trajectory patterns across different cluster solutions provided valuable insights into the optimal partitioning strategy.

Clustering Solution

The **3**-cluster solution demonstrated strong statistical performance. It achieved the highest CHI (**1,559.59**), indicating optimal cluster separation, alongside a reasonable SC (**0.413**), reflecting adequate cluster cohesion for clinical data. The WCSS (**35.21**) confirmed compactness, with cluster sizes ranging from **172** to **1,709** participants and a coefficient of variation (**1.054**), indicating moderate size variability. Collectively, these metrics identify three primary recovery archetypes representing key trajectories.

Despite its statistical robustness, the **3**-cluster solution lacked sufficient granularity to capture heterogeneous, multi-domain recovery patterns. Subjects within the same cluster often exhibited divergent trajectories across outcome domains, limiting clinical interpretability and the ability to inform personalized interventions. The solution effectively delineates broad recovery patterns but fails to resolve domain-specific differences crucial for individualized care.

To address these limitations, clustering solutions from **2** to **10** clusters were evaluated. Solutions with **7** or more clusters showed declining quality (SC $\approx$**0.25**) despite adequate minimum sizes. The **6**-cluster solution offered the best balance: SC = **0.298**, CHI = **1,079.26**, WCSS = **24.81**, and well-balanced cluster sizes (CV = **1.002**, range **30–1,082** subjects). This solution resolved cross-domain differentiation challenges while avoiding instability seen in larger cluster solutions.

The **6**-cluster solution improved multi-domain discrimination, capturing coherent phenotypes rather than domain-specific artifacts. Trajectory visualizations confirmed consis-

tent, interpretable patterns across recovery dimensions. Dendrogram analysis supported a hierarchical structure, showing that the additional clusters represent meaningful subdivisions rather than arbitrary splits.

While the SC declined from **0.413** (3 cluster) to **0.298** (6 cluster), this moderate reduction reflects the natural complexity of clinical recovery and an acceptable trade-off for enhanced granularity. Convergent evidence from dendrograms, silhouette plots, and multi-domain trajectories confirms that the **6**-cluster solution robustly identifies clinically meaningful, statistically valid recovery archetypes suitable for personalized care and prognostic assessment.

The **6**-cluster solution balances statistical validity and clinical utility. It recognizes the underlying structure of three fundamental recovery archetypes while providing the finer granularity necessary to capture multi-dimensional post-acute TBI recovery patterns. This nuanced classification supports individualized treatment planning and targeted intervention strategies more effectively than the simpler **3**-cluster model.

The integration of multiple evaluation criteria ensured that our final cluster solution balanced statistical metrics with interpretability, resulting in meaningful patient subgroups. These main numerical metrics are presented in Table 3.8. For all K, cluster counts sum to the **analytic** $n = 2{,}337$; "Min/Max Size" in Table 3.8 refers to the smallest and largest cluster at that K.

Table 3.8: Cluster Metrics for Different K

K	**SC**	**CHI**	**WCSS**	**Min Size**	**Max Size**	**CV Sizes**
2	0.567	1125.97	55.49	172	2165	1.206
3	0.413	1559.59	35.21	172	1709	1.054
4	0.279	1355.56	29.98	172	1082	0.652
5	0.294	1198.33	26.92	30	1082	0.896
6	0.298	1079.26	24.81	30	1082	1.002
7	0.238	1009.58	22.85	30	837	0.848
8	0.242	928.48	21.70	30	837	0.868
9	0.214	860.56	20.79	30	837	0.917
10	0.217	807.04	19.96	30	837	0.991

3.2 Results

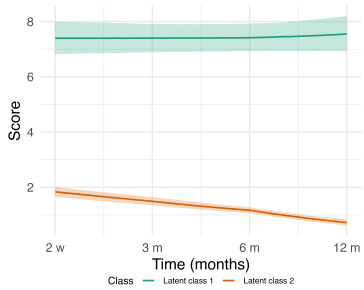
Following successful imputation, we applied a two-step process for developing multi-outcome trajectories. First, we used LCMMs to identify distinct trajectories for each outcome independently. We followed this in a second step of hierarchical clustering across LCMMs posterior probabilities for all outcomes and subjects. This section presents the results from single-outcome modeling and clustering analysis.

3.2.1 Single-outcome Modeling

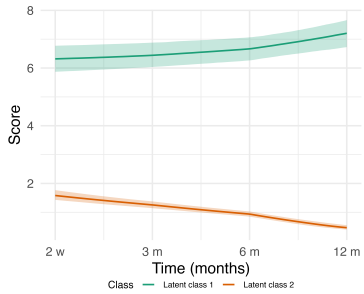
Model selection utilized a goodness-of-fit balance, meaning that we evaluated both information criteria of model fit and interpretability of the resulting classes. Specifically, we examined changes in AIC and BIC across models with increasing class numbers to assess relative improvements in fit. Also, class distributions have been considered to avoid spurious or very small classes; although no strict threshold was imposed, we aimed to avoid spurious micro-classes by requiring $\geq$**30 participants** ($\approx$1.3% of the analytic sample, n = 2,337). Mean posterior probabilities were inspected as indicators of classification quality. Finally, we compared the estimated class trajectories visually, ensuring that additional classes represented meaningful and that they remained consistent with trajectories previously described in the literature (e.g., for GOSE Score identified by Gardner *et al.* (2019) [5]). Therefore, we identified optimal class numbers that balance statistical fit with trajectory interpretability (Table 3.2.1). Information criteria along with class proportions distributions for the tested models are presented in Appendix Table 7.

Outcome	Optimal Ng
BSI-18 Anxiety Score	2
BSI-18 Depression Score	2
PCL-5 Score	3
RAVLT Score	3
TMTB Executive Score	3
WAIS Symbol	3
DRS Score	3
RPQ Score	3
ISI Score	4
TMTA Attention Score	6
WAIS Coding	3
PROMIS Intensity	3
RPQ-3 Score	2
RPQ-13 Score	2
SWLS Score	6
GOSE Score	6
QoLIBRIOS Score	3

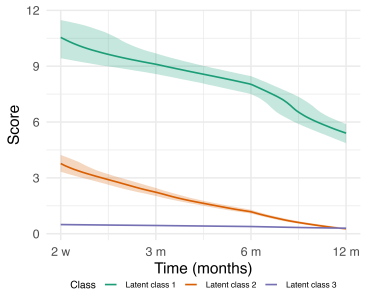
All **17** selected models achieved convergence and generated meaningful trajectory patterns. The models successfully classified all subjects, with class assignments ranging from **2** to **6** latent classes depending on the outcome variable. Each subject was assigned to the most probable latent class based on posterior probabilities. Comprehensive trajectory plots were generated for the selected models (separately presented in Figure 3.3), revealing distinct recovery patterns. Detailed explanation of plots construction provided in Section 4.3.4



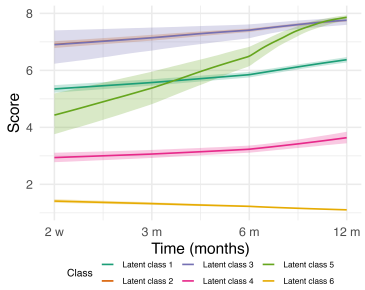
BSI-18 Anxiety Score



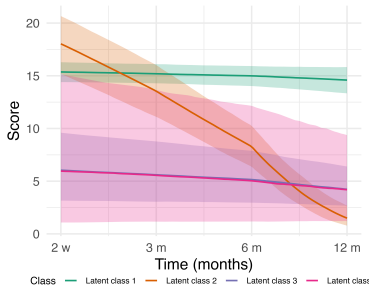
BSI-18 Depression Score



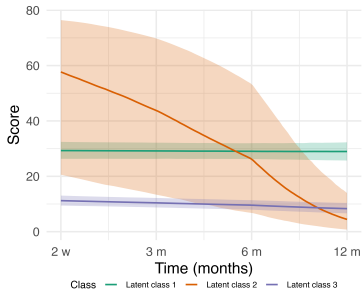
DRS Score



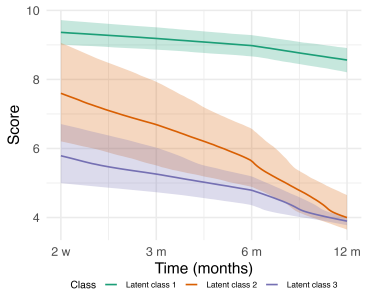
GOSE Score



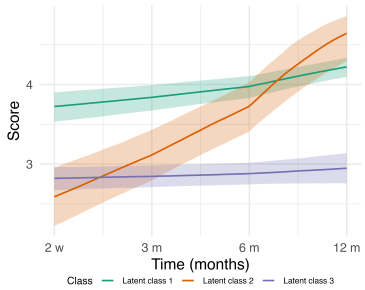
ISI Score



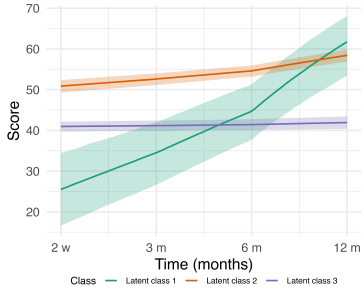
PCL-5 Score



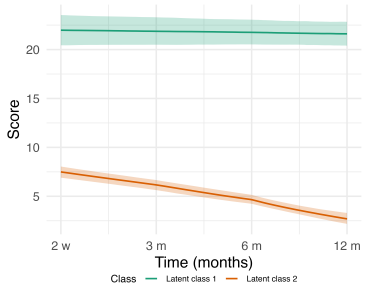
PROMIS Intensity



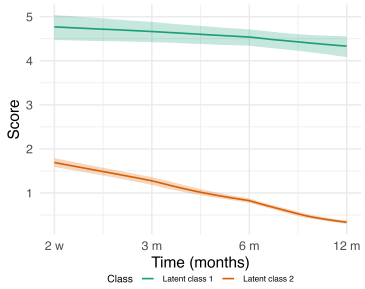
QoLIBRIOS Score



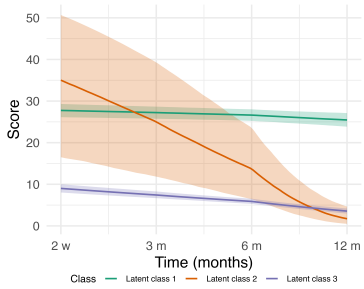
RAVLT Score



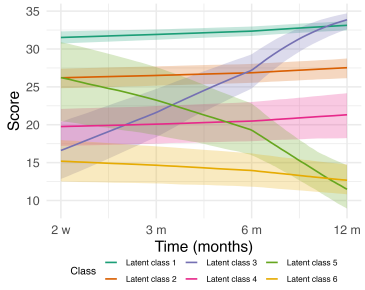
RPQ-13 Score



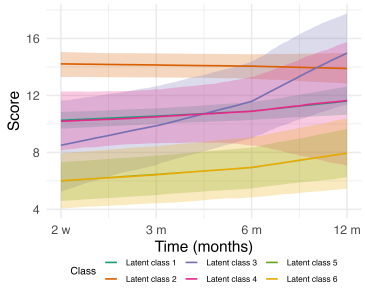
RPQ-3 Score



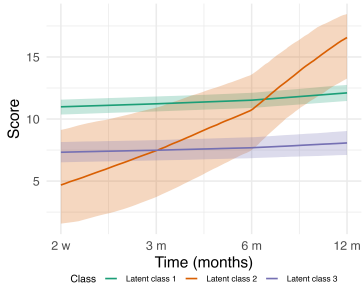
RPQ Score



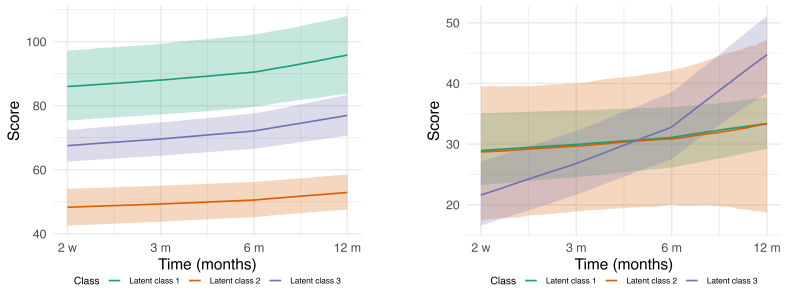
SWLS Score



TMTA Attention Score



TMTB Executive Score



WAIS Coding WAIS Symbol

Figure 3.3: Single-outcome (LCMMs) Trajectories

3.2.2 Multi-outcome Trajectory Identification

This section presents the trajectory analysis that identified six distinct recovery clusters, as illustrated by the dendrogram in Figure 3.4. Baseline demographic characteristics are summarized in Appendix Table 3.

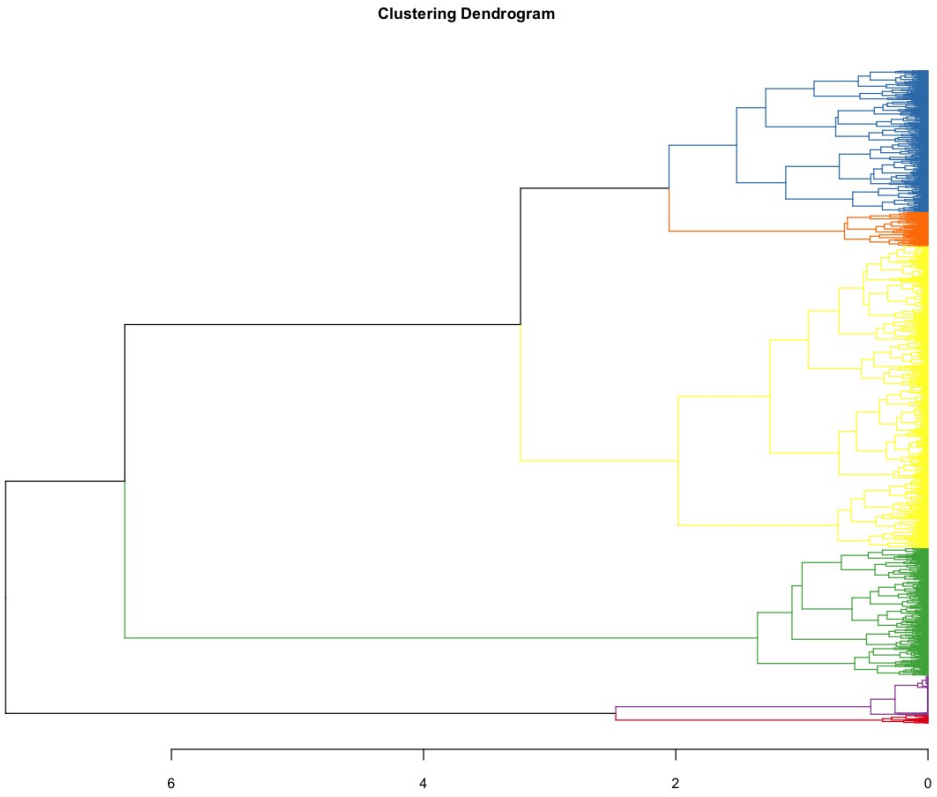


Figure 3.4: Cluster Dendrogram

The associated trajectories (Figure 3.5) illustrate the mean recovery patterns for each cluster with Confidence Intervals (CIs). In each figure, colored lines represent the mean fitted trajectories for the clusters, and shaded areas denote the CIs. The x-axis shows the follow-up time points, while the y-axis represents the corresponding outcome scores. For each outcome, the y-axis scale was adjusted to the observed data range to improve visual interpretation.

The trajectories were derived by aggregating individual outcome values within each cluster, based on the clustering assignment. These computations were performed in `R` using the `ggplot2` package, where summary statistics were used to calculate class-specific means and **95%** CIs.

CIs were estimated assuming approximate normality of the class-level mean distribution. For each time point and class, the **95%** CIs limits were computed using the `ggplot2`

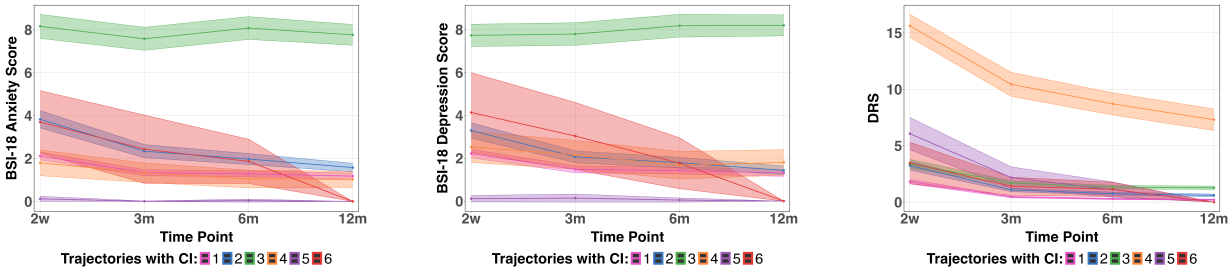
summary stats functions and defined as:

$$\text{Mean} \pm 1.96 \times \frac{\text{SD}}{\sqrt{n}},$$

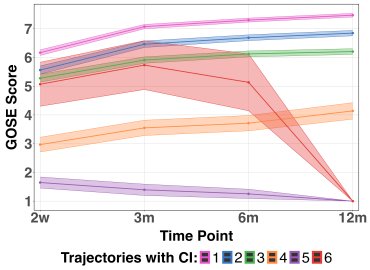
where *SD* denotes the standard deviation of the outcome values within the class at that time point, and n is the number of subjects.

It is also important to note that the estimated CIs at each time point are influenced by both the number of available observations and the variability of the outcome values within each cluster. When data are sparse or unevenly distributed across time, and particularly when the standard deviation of individual observations around the cluster mean is high, the resulting CIs tend to become noticeably wider. This widening accurately reflects increased uncertainty in the estimated cluster mean at those time points. For example, in Cluster **6**, the CIs are notably wide at the initial time points, reflecting high between-subject variability and a relatively small number of participants. Over time, these intervals gradually narrow, primarily because only a few observations remain available. The small number of clusters and the diminishing variability thus produce an apparently "tight" interval that should not be interpreted as high precision, but rather as a consequence of data limitation. Conversely, Cluster **5** illustrates a different pattern: the intervals remain consistently narrow across most time points and nearly disappear toward the end, owing to the very small cluster size (**n=30**) and highly similar outcome values among its members (declining to a minimal value). These examples highlight that the apparent variations in the width of the CIs are also driven by structural aspects of the data—such as sample size, missingness, and within-class variability.

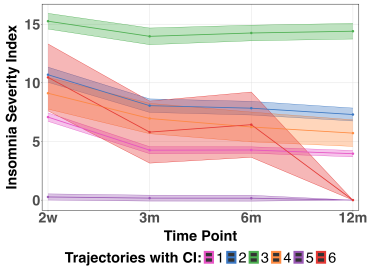
Figure 3.5: Cluster Trajectories



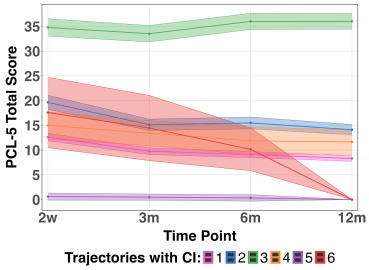
BSI-18 Anxiety Score BSI-18 Depression Score DRS Score



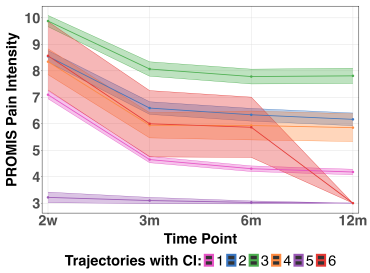
GOSE Score



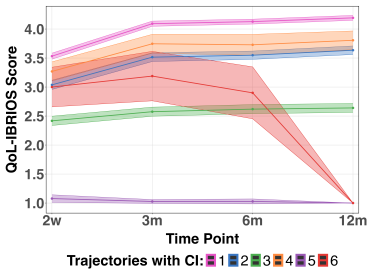
ISI Score



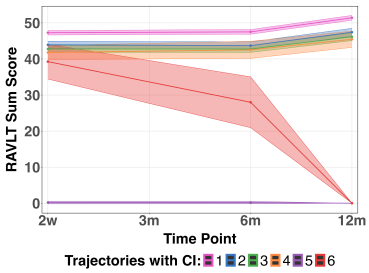
PCL-5 Score



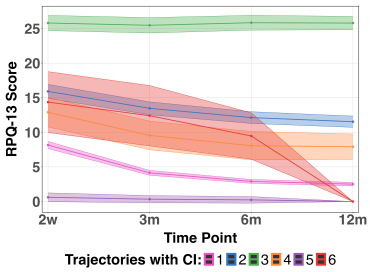
PROMIS Intensity



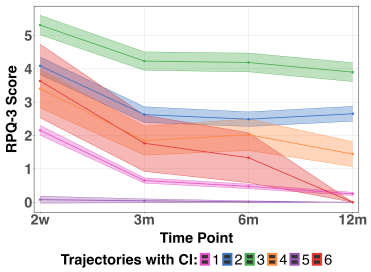
QoLIBRIOS Score



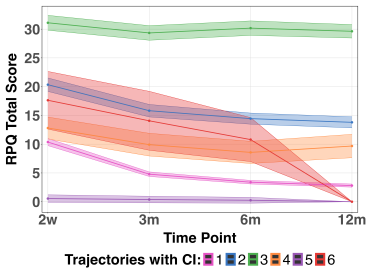
RAVLT Score



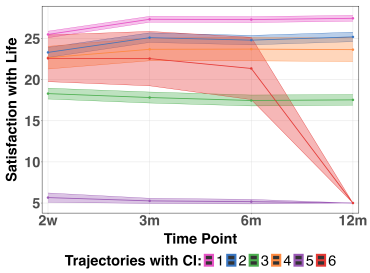
RPQ-13 Score



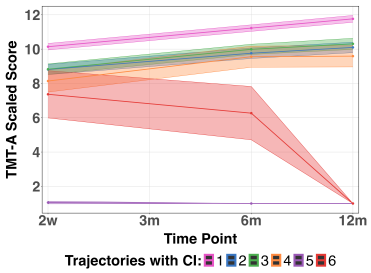
RPQ-3 Score



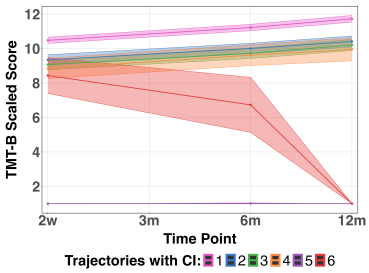
RPQ Score



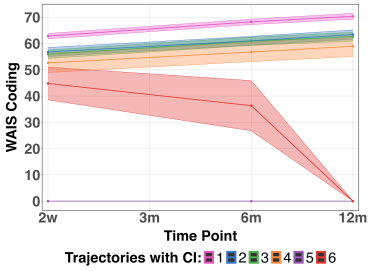
SWLS Score



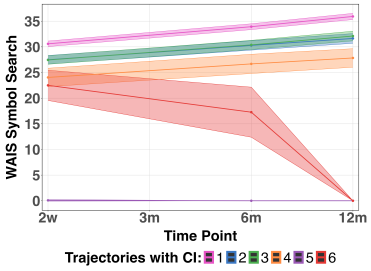
TMTA Attention Score



TMTB Executive Score



WAIS Coding



WAIS Symbol

3.2.3 Trajectory Analysis

Cluster 1: Good Recovery (Pink, n=1,082, 46.3%)

This largest cluster, representing nearly half the examined cohort, demonstrates the most favorable recovery trajectories across virtually all outcomes. These patients, characterized by their younger age (38.83 ± 17.18 years), higher education levels (14.09 ± 2.85 years), and predominantly mild TBI severity (86.9%), establish a benchmark for successful recovery following TBI. Their demographic profile—71.6% male, second-lowest psychiatric history prevalence at 18.2%, and minimal prior TBI history—suggests a population with solid baseline resilience factors that facilitate recovery.

The global recovery measures paint a promising picture of steady improvement. The GOSE Score scores progress from approximately 6.2 at two weeks to 7.5 by twelve months, indicating movement from lower moderate disability toward good recovery. This improvement is mirrored in the DRS Score, where scores decline from 1.8 to near-zero, suggesting minimal to no residual disability. What makes this cluster particularly interesting is not just the endpoint achieved, but the consistency of improvement—the narrow confidence intervals throughout the trajectory indicate that this recovery pattern is remarkably reliable within this population.

The neuropsychiatric profile of Cluster 1 reveals equally encouraging patterns. Depression and anxiety scores, already starting at relatively low levels (BSI-18 Depression Score ∼2.5, BSI-18 Anxiety Score ∼2.5), show continued improvement throughout the year, eventually converging to minimal symptomatology. The Post-Traumatic Stress Disorder (PTSD) symptom trajectory, primarily measured by PCL-5 Score scores, follows a similar pattern of gradual decline from 14 to 8, suggesting effective psychological adaptation to the trauma. Interestingly, the three-month time-point appears to represent a critical intersection across multiple symptom domains, where a peak is observed before steady improvement begins. This pattern might reflect the natural TBI recovery; the first few weeks

after a trauma have a dominant influence that characterizes the progress of recovery.

Cognitive recovery in this cluster is particularly robust. The neurocognitive measures—TMTA Attention Score and TMTB Executive Score, WAIS Coding and WAIS Symbol—all maintain high performance levels throughout the observation period, with scores consistently in the upper ranges of their respective scales. The RAVLT Score scores, maintaining values between 47 and 51 with a lightly increasing pattern, indicate preserved verbal learning and memory capacity. This cognitive resilience likely contributes to the cluster's overall positive outcomes, as intact executive function and processing speed facilitate adaptation.

Quality of life measures further reinforce the positive trajectory of this cluster. SWLS Score scores improve from 25 to 27, approaching the upper range of the scale, while QoLIB-RIOS Score scores show similar upward trends. PROMIS Intensity intensity remains manageable throughout, declining from moderate levels (7) to mild (4) by twelve months, with a local peak at 3 months (5) followed by a decline. Sleep disturbance, measured by the ISI Score, follows a gradual improvement trajectory from mild clinical insomnia to sub-threshold levels. These interconnected improvements in pain, sleep, and overall life satisfaction suggest a positive cycle of recovery where improvements in one interconnected domain facilitate progress in others.

Cluster 2 represents a transitional population with more complex presentations than the optimal recovery group in Cluster 1. These patients, slightly older at 43.03 ± 16.82 years and with Cluster 3 has the highest female representation (36.6% – 188), demonstrate recovery trajectories that, while positive, plateau at moderate levels across multiple domains. The higher incidence of falls as injury mechanism (30.3%) and increased ICU admission rate (42.2%) suggest a population with either more severe initial injuries (GCS at Admission 13.84 ± 2.69) or additional vulnerability factors that complicate recovery.

Global functional outcomes for this cluster show a pattern of initial improvement followed by stabilization. GOSE Score scores progress from approximately 5.5 to 6.8, indicating movement from upper severe disability to lower moderate disability, but not achieving the good recovery range seen in Cluster 1. The DRS Score shows a similar pattern, declining from 3.3 to 1.

The neuropsychiatric profile reveals persistent moderate symptomatology. Depression scores remain relatively stable in the low-moderate range (2-3), while anxiety shows minimal improvement. The PCL-5 Score trajectory, declining from 19 to 14, indicates some improvement in PTSD symptoms. This pattern suggests that while these patients achieve functional recovery, the psychological impact of their injury remains substantial. The stability of these symptoms after the initial recovery period might indicate the need for specific psychological interventions.

Cognitive performance in Cluster 2 maintains moderate levels throughout the observation period. TMTA Attention Score scores hover around 9.5-10.5, WAIS Coding improves from 57 to 63, and WAIS Symbol remains stable around 30-32 points. While these scores indicate functional cognitive capacity, they fall short of the much higher performance seen in Cluster 1. The RAVLT Score scores, stable around 44-47, suggest adequate but not optimal verbal memory function. This cognitive profile—functional but not optimal—may contribute to the plateau effect seen in other domains.

The post-concussion symptom burden remains substantial in this cluster. RPQ Score total scores show only modest improvement from 20 to 13-14, with both physical (RPQ-3 Score) and cognitive-emotional (RPQ-13 Score) symptoms persisting at moderate levels. This sustained symptom burden likely contributes to the moderate quality of life outcomes observed, with SWLS Score scores remaining relatively stable around 24-25 and QoLIB-RIOS Score scores showing only gradual improvement. PROMIS Intensity intensity follows a similar pattern, declining from 8.5 to 6.5 but remaining in the moderate range, while insomnia symptoms persist at clinically significant levels throughout the year.

Cluster 3: Persistent High Symptom Recovery (Green, n=456, 19.5%)

Cluster 3 presents a striking contrast to the recovery trajectories of the first and second clusters, characterized by persistent symptoms with minimal improvement over time. This cluster's demographic profile reveals significant vulnerability factors: the lowest education level across 3 larger clusters (12.78 ± 2.37 years), dramatically elevated psychiatric history (39.7%— almost twice the overall average), the highest rate of prior TBI (33.8%), and increased exposure to violence (9.9% of injuries). These pre-existing vulnerabilities appear to fundamentally alter the recovery trajectory.

The most distinctive feature of this cluster is the remarkable stability—or more accurately, the absence of improvement—across almost all outcomes. Global function measures show minimal change, with GOSE Score scores remaining around 5.3-6.1 and DRS Score scores showing only slight improvement from 3.3 to 1.2. This pattern suggests that whatever recovery mechanisms operate in other clusters are either absent or ineffective in this population. The flat trajectories can be interpreted as this is not simply variability in recovery but rather a consistent pattern of non-improvement.

The psychiatric symptoms in Cluster 3 are particularly concerning. BSI-18 Anxiety Score anxiety scores remain elevated around 7-8 throughout all time-points. Depression scores follow an identical pattern, maintaining moderate levels (7.5-8.5) without improvement. The PCL-5 Score scores, stable at 35-36, indicate severe, persistent PTSD symptomatology. This psychiatric symptom constellation are high, stable, and suggests that the interaction between pre-existing psychiatric vulnerability and acute TBI creates a state resistant to recovery.

Post-concussion symptoms in this cluster show the most dramatic deviation from observed recovery patterns. The RPQ-13 Score scores remain completely flat at approximately 27 points across all time-points—a finding that is remarkable for its consistency. Similarly, RPQ-3 Score physical symptoms start highest among all clusters (5.3) and show minimal improvement (4 at twelve months). The total RPQ Score scores maintain around 31 throughout the study period. This stability in post-concussion symptoms, particularly when contrasted with the improvement seen in other clusters, suggests a fundamentally different underlying process, which requires further investigation.

Cognitive performance, while not at the floor levels seen in some other clusters, remains consistently below average. TMT scores around 8.5-10.5, WAIS Coding at 57-64, and RAVLT Score scores of 42-45 indicate functional but suboptimal cognitive performance. The stability of these scores suggests that cognitive deficits are not worsened but are also not significantly improved with time and assumed rehabilitation. Quality of life outcomes reflect this broad pattern of persistent dysfunction, with SWLS Score scores remaining low (around 18) and showing no improvement over twelve months. PROMIS Intensity intensity remains high (10 declining minimally to 8), and insomnia severity persists at clinical levels (15 throughout).

Cluster 4: Initial Severe with Partial Recovery (Orange, n=122, 5.2%)

Cluster 4 represents patients with severe initial injuries, as evidenced by the extremely low proportion of mild TBI (20.5%), high rate of severe TBI (35.2%), and near-universal ICU admission (91%). These patients, with a mean GCS of 8.44 $\pm$ 4.76 at emergency department presentation, represent moderate-to-severe TBI requiring intensive medical management. The predominance of road traffic incidents (68%) as injury mechanisms suggests high-energy trauma contributing to injury severity.

Despite the severity of the initial injury, this cluster shows more dynamic recovery trajectories than might be expected. Global function measures reveal gradual but meaningful improvement, with GOSE Score scores progressing from 3.0 to 4.2—still within the severe disability range but showing measurable gains. The DRS Score scores, starting highest at 15.5, show the slowest improvement among all clusters but do demonstrate a decline to 7.2 by twelve months. This pattern suggests that while recovery is possible even after severe injury, it proceeds more slowly and requires longer observation periods to fully characterize, which we also discuss in the study limitations.

The neuropsychiatric recovery in Cluster 4 shows interesting patterns of improvement despite severe initial injury. Anxiety and depression scores, while starting at moderate levels similar to Cluster 1, show overlapping improvement trajectories, eventually reaching low levels (BSI-18 Anxiety Score). PTSD symptoms follow a similar improving pattern,

declining from 17 to 12. This psychological recovery, despite severe neurological injury, might reflect the absence of the pre-existing psychiatric vulnerabilities seen in Cluster 3, suggesting that baseline mental health status may be important for psychiatric outcomes.

Cognitive outcomes in this cluster reveal the expected impact of severe TBI. While starting lower than other clusters (TMTA Attention Score around 8, WAIS Coding at 52), gradual improvement is observed across cognitive measures. The trajectory may suggest ongoing neuroplasticity and recovery even in the context of significant initial brain injury. However, twelve-month outcomes still fall short of normal performance levels. The RAVLT Score scores, improving from 42 to 45, suggest that even memory function—often severely affected in serious TBI—shows capacity for improvement.

Post-concussion symptoms in Cluster 4 follow an intermediate pattern. RPQ Score show moderate improvement, falling between the good recovery of Cluster 1 and the non-improvement of Cluster 3. This pattern extends to quality of life measures, where gradual improvement is seen in both SWLS Score and QoLIBRIOS Score scores. PROMIS Intensity intensity shows a similar intermediate pattern, declining from 8.5 to 6, while insomnia symptoms improve from moderate to mild severity. These outcomes suggest that severe TBI, while imposing a significant initial burden, does not preclude meaningful recovery when.

Cluster 5: Early Mortality (Purple, n=142, 6.1%)

Using the mortality investigation conducted in the preliminary analysis phase, we have confirmed that all patients in Cluster 5 experienced fatal outcomes, with the vast majority dying before the two-week assessment point. This cluster, therefore, represents the most severely injured non-survivors. Given that nearly all trajectory data for this cluster results from statistical imputation rather than actual patient assessments, meaningful trajectory analysis is not possible. We therefore focus exclusively on the baseline demographic and injury characteristics that defined this fatal outcome group.

This cluster comprised 142 patients with distinctive demographic features: they were the second-oldest group (55.35 $\pm$ 18.08 years), overwhelmingly male (77.5%), and predominantly White (78.9%). Notably, 21.8% had a documented psychiatric history, and their prior TBI status showed mixed patterns—50% with no prior history, though 14.8% had experienced previous significant brain injuries.

The injury mechanisms reveal important patterns: 49.3% resulted from road traffic incidents while 36.6% followed incidental falls—the second-highest fall rate across all clusters, consistent with the older age demographic. The neurological presentation was catastrophic: 78.9% experienced loss of consciousness, though post-traumatic amnesia documentation was problematic, with 44.4% having unknown status, likely reflecting the severity of injury

preventing assessment. Only 26.8% had confirmed PTA, possibly representing brief periods of responsiveness before deterioration.

The injury severity classification starkly illustrates the fatal nature of these injuries: only 19.7% were classified as mild TBI, 5.6% moderate, and 40.8% severe—the highest severe TBI rate among all clusters. The mean GCS score of 7.5 ± 4.86 (n=84) indicates severe neurological compromise. Hospital utilization patterns reflect the critical nature of these injuries: 96.5% required ICU admission (the highest rate), with only 2.8% admitted to regular hospital wards and a negligible 0.7% discharged from emergency. These utilization patterns indicate that, despite maximal medical intervention, these patients' injuries were incompatible with survival.

The apparent "trajectories" observed in this cluster—characterized by floor effects across cognitive measures, near-zero psychiatric symptoms, and minimal post-concussion symptoms—likely reflect artifacts from imputation procedures rather than genuine patient data. For instance, the GOSE Score trajectory demonstrates consistently catastrophic outcomes, beginning around 2 and declining to 1. These scores correspond to vegetative state or, at best, severe disability with complete dependence, aligning with expectations for a cohort in which all patients ultimately died, most within the first two weeks.

However, the DRS Score trajectory presents a contradictory pattern. It begins near 6.1, indicating moderate disability, and appears to improve to around 1.4 by 12 months, suggesting mild disability. Such an "improvement" is clinically implausible, as patients in vegetative states—or deceased—cannot simultaneously demonstrate increasing functional independence. This paradox is likely the result of imputation procedures and information loss rather than true clinical change. The stark disconnection between these two trajectories in a cohort of patients with fatal severe TBI highlights critical methodological challenges in longitudinal TBI research.

Cluster 6: Delayed Mortality (Red, n=30, 1.3%)

Cluster 6 represents a small but clinically important group of 30 patients who survived initial hospitalization but subsequently died during the follow-up period. Mortality verification confirmed that all patients in this cluster experienced fatal outcomes, though, unlike Cluster 5, some survived long enough to contribute limited early assessment data. The small sample size and progressive mortality throughout the follow-up period result in extremely limited trajectory data.

This cluster demonstrated distinctive demographic characteristics: they were the oldest group (56.69 ± 18.45 years), predominantly male (76.7%), and remarkably homogeneous in racial composition (96.7% White). All patients had complete education data (13.5 ± 2.78 years). Notably, 30% had documented psychiatric history—a relatively high prevalence.

The injury profile reveals a pattern consistent with elderly falls: 56.7% of injuries resulted from incidental falls—dramatically the highest rate across all clusters—while only 33.3% involved road traffic incidents. This mechanism distribution, combined with advanced age, suggests a population with baseline frailty and vulnerability. Despite the high fall rate typically associated with lower-energy trauma, 73.3% experienced loss of consciousness and 66.7% had documented post-traumatic amnesia, indicating significant brain injury even from seemingly lower-impact mechanisms.

The injury severity classification presents an apparent paradox: 76.7% were classified as mild TBI, with 10% severe and no moderate injuries recorded. The mean GCS of 13.54 ± 3.64 (n=24) further supports predominantly mild initial presentation. However, the clinical course—with 43.3% requiring ICU admission and 40% admitted to regular hospital wards—suggests injury severity classification may not have captured the full clinical picture. In elderly patients, even "mild" TBI can trigger cascading medical complications.

The limited trajectory data available showed initial moderate function followed by a precipitous decline. For the small number of patients with early assessments, GOSE Score scores began around 5.1 before declining. Cognitive measures showed similar patterns of initial moderate impairment followed by severe deterioration. However, interpreting these patterns is problematic: the trajectories reflect a survivorship bias where only patients well enough to complete assessments contributed data, while those deteriorating rapidly were lost to follow-up or death.

The combination of advanced age (the oldest cluster), high fall rate, and delayed mortality pattern suggests this cluster represents elderly patients experiencing fatal complications following what initially appeared to be survivable injuries. The progression from mild TBI classification to universal mortality highlights the inadequacy of traditional severity measures in predicting outcomes in elderly populations.

Chapter 4

An Interactive Analysis Tool

This chapter presents the design and implementation of three specialized Shiny applications developed to support: **(1)** the link function selection process, **(2)** the determination of the optimal number of classes of LCMM, and **(3)** the initial clustering explanatory analysis. Each application addresses specific analytical challenges by providing both visualizations and summary statistics in a unified interface, thereby facilitating decision-making.

4.1 The R Shiny Framework

The `R` Shiny framework, originally developed by RStudio, now Posit [46], provides a reactive programming paradigm for constructing interactive web applications directly from `R` code and is widely used for data science applications. The framework's architecture implements a reactive graph where inputs and outputs are connected through reactive expressions, enabling automatic dependency tracking and minimal re-computation when input values change. This programming model, inspired by functional reactive programming Functional Reactive Programming (FRP), ensures computational efficiency and automatic caching of intermediate results.

The R Shiny framework consists of two primary components: User Interface (UI) specification that defines the application layout using `R` functions that generate HyperText Markup Language (HTML), and a server function containing reactive expressions and observers that respond to user interactions. The communication between UI and the server occurs through WebSocket connections, enabling bidirectional real-time data transfer with minimal latency.

4.2 Application 1: Link Function Comparison

4.2.1 UI Architecture

The link function comparison interface implements a hierarchical layout structure using the shinydashboard grid system and consists of four primary panels arranged in the bootstrap grid:

- **Sidebar Panel:** Contains user selection dropdown, model navigation controls, and action button for confirming selection of a model.
- **Model Statistics Panel:** Interactive DataTable with sorting, filtering, and pagination.
- **Visualization Panel:** Dynamic plot rendering area for link function fit plot and residuals. Also has an action button to plot link function "fit plot" for all available link function for a selected outcome.
- **Selection Panel:** Model selection dropdowns for 2 best model selected for the further compare and area dedicated field for a comment and two action buttons for saving section and navigating to the next model.

A page UI of the application is presented in Figure 4.1.

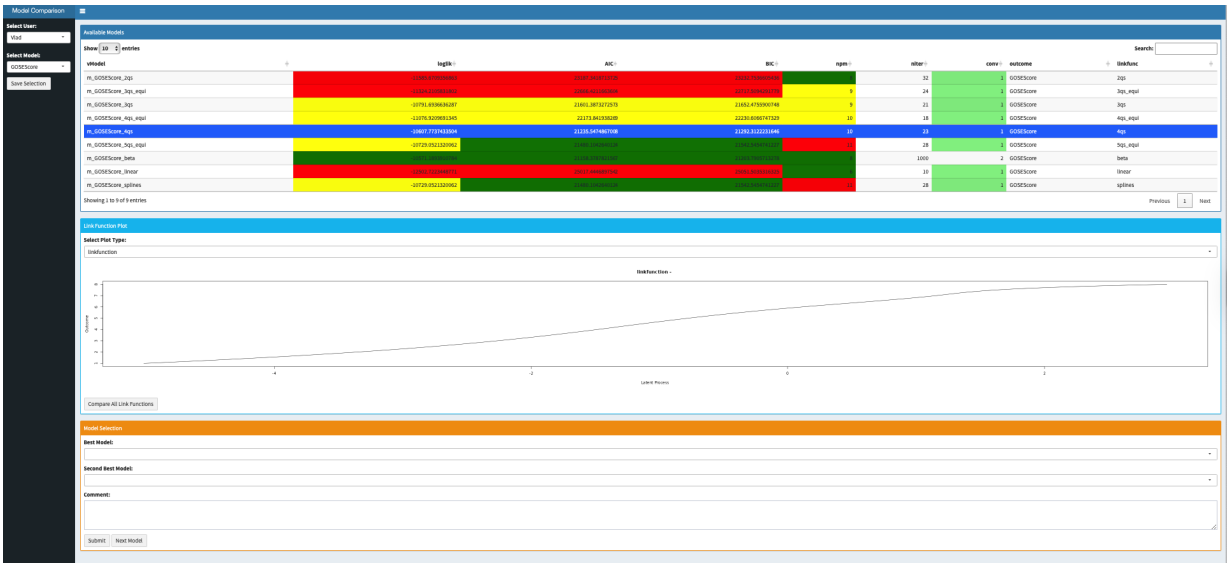


Figure 4.1: Sidebar Panel

4.2.2 Use Cases and Scenarios

A user analyzing an outcome can compare different link functions for a selected model. The workflow proceeds as follows:

1. Select "Outcome model" from the dropdown menu
2. Review AIC, BIC, loglik, npar, niter values in the colour-coded table where green indicates better values.
3. Click rows to select models with similar information criteria but different link specifications
4. Alter the plot type between "residuals", and "fit" to assess:
    - Link function shape and transformation characteristics
    - Residual patterns indicating model misspecification
5. Click "Compare All Link Functions" to visualize simultaneous overlay
6. Select 2 best variants and document selection rationales in comment field
7. Submit selection for recording in Comma-separated Values (CSV) file and navigate to the next outcome model.

Collaborative model review mode is available for a few users. It can be done by switching between users in the Sidebar Panel, and each user's selections are logged with a timestamp and user identifier, allowing comparison of selections from a few perspectives.

4.2.3 Main Components

Model Parsing Algorithm. The link function comparison application implements a deterministic finite automaton for parsing model identifiers into semantic components. The parsing function $\phi : S \to (O, L)$ maps string identifiers $s \in S$ to tuples of outcome $o \in O$ and link function $l \in L$. The parsing algorithm uses regular expression matching to extract model specifications from the model name. Although this method does not retrieve specifications directly from the model object, it offers a relatively simple implementation and provides users with flexibility. By relying solely on model names, users can easily organize models and apply their own grouping preferences.

DataTable Configuration. The DataTable implementation provides filtering and sorting capabilities for capturing the models and their appropriate grouping by outcomes and model identification itself. Also, the same approach used for the rest 2 application modules. Here is a structure realization example:

```
datatable(df,
    filter = 'top',            # Column-specific filters
    selection = list(
        mode = 'multiple',     # Multiple row selection
        target = 'row'         # Row-based selection
    ),
    options = list(
        dom = 'tprl',         # Table, pagination, records, length
        pageLength = 10,       # Default rows per page
        scrollY = "400px",     # Vertical scrolling
        scrollX = TRUE         # Horizontal scrolling
    )
)
```

Plot Rendering Feature. The visualization system is designed using a multi-stage rendering pipeline. First, the system extracts data from the given model (mounted by model list pattern). Depending on the specified plot type, different procedures are applied. If the plot type is set to link function, the system generates a grid specific, transforms them using the model's link function, and then plots the results. For the residuals plot type, the system computes residuals from the model and plots them against time. For these plotting tasks, internal `lcmm` package functions used (`lcmm` package call plot its `ggplot` based function for model object) therefore there is no specific algorithms established to perform the calculation.

4.3 Application 2: Single Model Comparison

4.3.1 UI and Layout

The class enumeration compare application implements a multi-panel dashboard architecture with synchronized components organized into functional regions. The interface presented in Fig 4.2 and primarily consist of:

- **Control Sidebar:** Group management system with Create, Read, Update, Delete (CRUD) operations for collections

- **Primary Display Grid:** Responsive layout containing model summary table with real-time filtering feature.
- **Visualization Matrix:** A grid of diagnostic plots updated through reactive bindings
- **Tabbed Analysis Panel:** Multi-tab interface for coefficients, diagnostics, and demographics

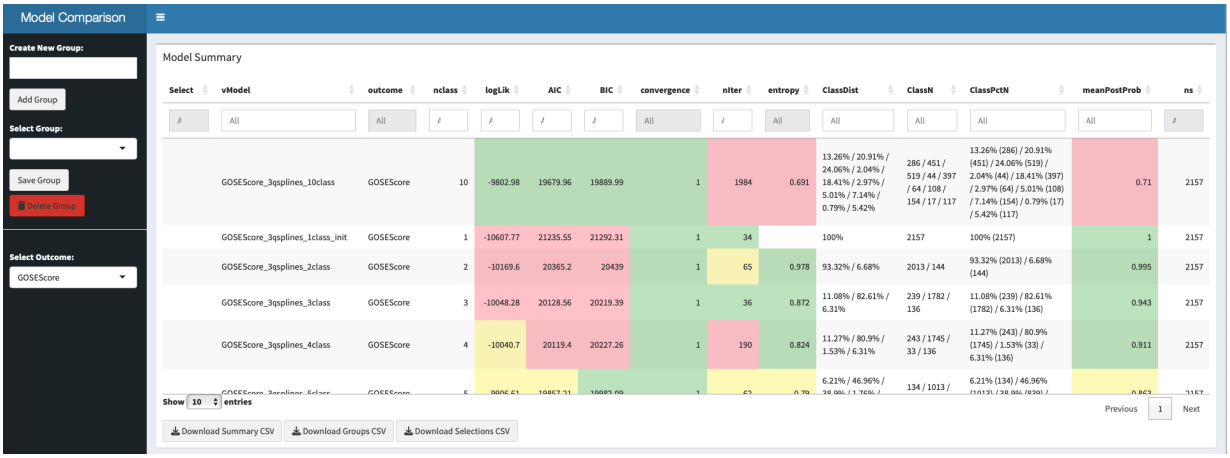


Figure 4.2: Sidebar Panel Model Summary

The navigation system implements both sequential and random access patterns:

```
# Navigation state management
nav_index <- reactiveVal(1)
nav_list <- reactive({
    if (input$nav_mode == "all") {
        all_outcomes
    } else {
        group_filtered_outcomes
    }
})
```

The group management system enables organization of models into logical collections for comparative analysis. The implementation maintains a persistent CSV-based storage with the following schema:

$$\text{GroupDef} = \langle \text{group\_name}, \text{vModel}, \text{timestamp} \rangle \tag{4.1}$$

Group operations implement ACID properties through file-based transactions:

- **Atomicity:** Write operations use temporary files with atomic rename
- **Consistency:** Schema validation before persistence
- **Isolation:** File locking prevents concurrent modifications
- **Durability:** CSV format ensures human-readable backup

4.3.2 Interactive Model Comparison Features

Multi-Model Selection

The DataTable supports multi-row selection through Shift+Click and Ctrl+Click interactions, maintaining selection state in a reactive value:

The visualization matrix displays four synchronized plots that update based on model selection:

1. **Metrics Evolution:** Line plot showing AIC/BIC/entropy across class numbers
2. **Class Distribution:** Stacked bar chart of class proportions
3. **Predicted Trajectories:** Class-specific predicted curves with confidence bands
4. **Individual Trajectories:** Actual patient data grouped by class assignment

4.3.3 Use Cases and Workflows

Use Case 1: Optimal Class Number Determination

A user investigating some outcome's trajectories needs to determine the optimal number of classes:

1. Create group "Outcome for ... Domain" through sidebar interface
2. Add models to group
3. Review model summary table with colour-coded metrics:
    - Green cells indicate favourable values (lower AIC/BIC, higher entropy)

- Convergence status highlighted (green=success, red=failure)

4. Select all models for simultaneous comparison
5. Examine metrics evolution plot for elbow points indicating optimal model complexity
6. Toggle to "Individual Trajectories" tab to assess within-class heterogeneity
7. Document decision in trajectory description panel with tentative interpretation

Use Case 2: Cross-Domain Model Integration

Analyzing relationships between outcome domains:

1. Create group "... Domain Group" containing relevant models
2. Navigate between outcomes using "Next" button or dropdown selection
3. Compare class distributions across domains using "Show Alluvial Plot" feature
4. Compare Sankey diagram showing class membership transitions between outcomes

Trajectory Description System

The trajectory description interface implements a hierarchical categorization of trajectory patterns:

```
categories <- list(
    "By Shape" = c("Linear increase", "U-shape", ...),
    "By Clinical" = c("Favourable recovery", ...),
    "By Statistical" = c("High posterior probability", ...),
    "By Temporal" = c("Rapid early change", ...)
)
```

Each model can be tagged with multiple descriptors, stored persistently in trajectory_log.CSV:

$$\text{TrajDesc} = \langle \text{vModel}, \text{descriptors}[], \text{comment}, \text{timestamp} \rangle \tag{4.2}$$

Demographics Profiling Module

The demographics module computes statistics (information criterias) for each model class and presented in Algorithm 2.

Algorithm 2 Demographics Extracting

ComputeDemographicssubjects, class $filtered \leftarrow$ subjects[subjects.class = class] $stats \leftarrow \{\}$ $var \in$ continuous_vars $stats[var] \leftarrow$ mean $\pm$ sd $var \in$ categorical_vars $stats[var] \leftarrow$ count(percentage) $stats$

4.3.4 Trajectory Plots

The trajectory visualization component constructs a temporal grid with adaptive resolution to represent predicted outcome trajectories over time. Let $T = [t_{\min}, t_{\max}]$ denote the observed time interval. The temporal grid points are defined as:

$$\mathcal{G} = \{t_i : t_i = t_{\min} + i \cdot \delta, \quad i = 0, ..., g-1\}, \tag{4.3}$$

where $\delta = (t_{\max} - t_{\min})/g$, and $g \in \{20, ..., 200\}$ is a user-specified resolution parameter controlling the density of the grid. When confidence intervals are requested, the number of grid points is limited to $g = 50$ in order to reduce the computational load.

The application retrieves data and model information dynamically using a multi-environment evaluation strategy. For a given model m with call environment $\mathcal{E}_m$, the evaluation hierarchy is defined as:

$$\text{data} = \text{eval}(\text{expr}, \mathcal{E}) \quad \text{where } \mathcal{E} \in \{\mathcal{E}_{\text{formula}}, \mathcal{E}_{\text{global}}, \mathcal{E}_{\text{parent}}\}. \tag{4.4}$$

The system iterates through the available environments until successful evaluation or exhaustion of candidates, ensuring robust data access across different execution contexts.

All trajectories displayed in the application are generated directly from the fitted LCMM, using the `predictY()` function of the `lcmm` package [14]. This function provides class-specific predicted values of the outcome variable across the defined temporal grid, computed from the fixed and random effects estimated in the fitted model. The predicted means are performed using Gauss–Hermite or Monte Carlo methods [14].

When CIs are required, `predictY()` performs a Monte Carlo approximation of the posterior distribution of the model parameters. Specifically, it generates a large number

of random draws from the asymptotic posterior of the estimated parameters and recomputes predicted outcome values for each draw. Then the 2.5%, 50% and 97.25% persentiles provide the mean prediction and its 0.95% CIs. The predicted class-specific mean trajectories and their CIs are subsequently visualized using `ggplot2`, with time standardized to evenly spaced anchors corresponding to follow-up points (2 weeks, 3 months, 6 months, and 12 months).

Accordingly, the `Shiny` module serves solely as a dynamic visualization and control layer. All computations of predicted means, uncertainty propagation, and posterior approximations are fully handled internally by the `lcmm` package, ensuring that the displayed trajectories accurately reflect the model-based predictions derived under the estimated parameter distribution.

4.4 Application 3: Multi-outcome Clustering

4.4.1 UI Architecture

The clustering exploration system implements a complex multi-tab interface with four primary analytical modules accessible through the sidebar menu navigation (4.3). The interface state management employs a tabbed architecture where each tab maintains independent reactive contexts while sharing common data structures through global reactive values.

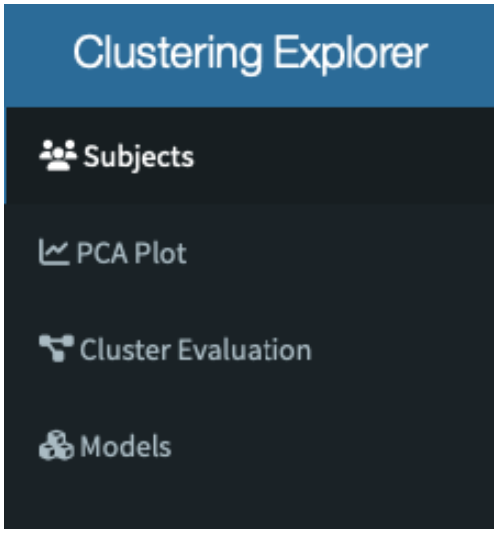
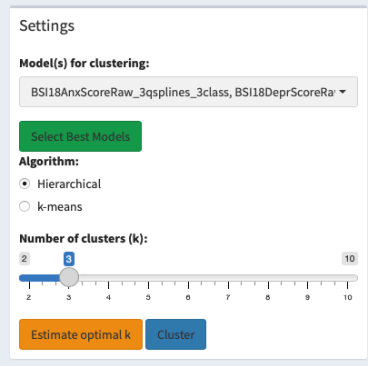


Figure 4.3: Multi-tab Interface

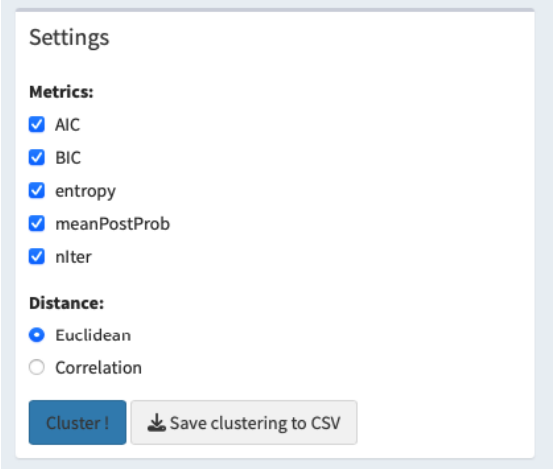
Tab Structure and Navigation

The application implements four interconnected analytical views (Figure 4.4):

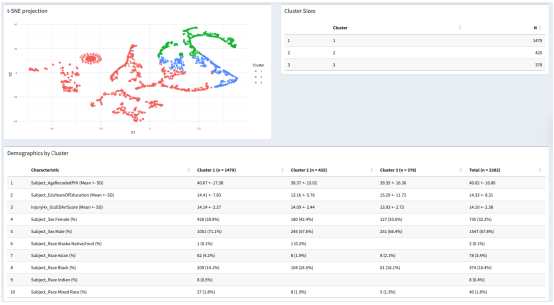
1. **Subjects Tab:** Primary clustering interface with algorithm selection and parameter controls
2. **PCA Plot Tab:** Linear dimensionality reduction visualization
3. **Cluster Evaluation Tab:** Elbow and silhouette methods for optimal k determination
4. **Models Tab:** Model-level clustering based on statistical metrics



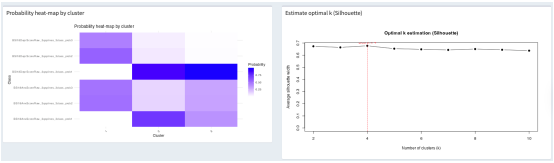
Clustering Set-up



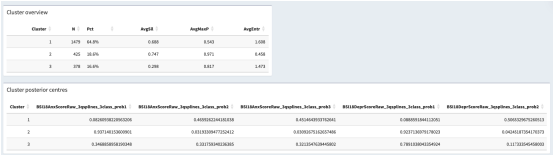
Clustering Settins



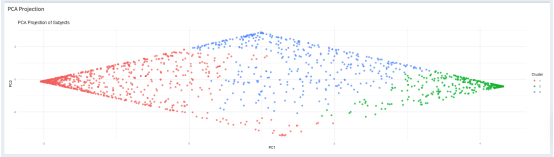
Cluster Demograhic Table



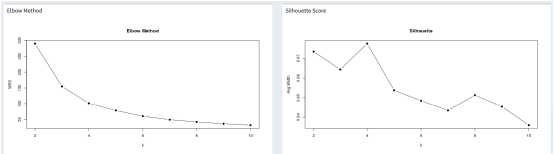
Profiling



Profiling continued



PCA View



Cluster Metrics

Figure 4.4: Multi-tab Interface

4.4.2 Clustering Control Panel

The control panel implements a hierarchical configuration system with the following components:

Model Selection Interface

The pickerInput widget provides multi-select functionality with search and batch operations:

```
pickerInput(
    inputId = "subj_models",
    label = "Model(s) for clustering:",
    choices = split(model_meta$vModel, model_meta$outcome),
    multiple = TRUE,
    options = list(
        `actions-box` = TRUE,       # Select/deselect all
        `live-search` = TRUE,        # Real-time filtering
        `selected-text-format` = "count > 3"  # Compact display
    )
)
```

The "Select Best Models" button automatically populates optimal models based on predefined criteria, implementing a one-click configuration for standard analyses.

Algorithm Selection and Parameters

The clustering algorithm selection employs radio buttons for algorithm choice and dynamic UI updates for algorithm-specific parameters:

$$\text{AlgorithmParams} = \begin{cases} \{k, \text{linkage\_method}\} & \text{if hierarchical} \\ \{k, \text{nstart}, \text{iter.max}\} & \text{if k-means} \end{cases} \tag{4.5}$$

4.4.3 Visualization Components

Dendrogram with Cluster Rectangles

For hierarchical clustering, the dendrogram visualization implements colored rectangles delineating cluster boundaries:

```
plot(hc, labels = FALSE, hang = -0.01)
rect.hclust(hc, k = input$k,
            border = rainbow(input$k, alpha = 0.8))
```

t-SNE Projection with Adaptive Parameters

The t-SNE visualization automatically adjusts parameters based on data characteristics:

Algorithm 3 Adaptive t-SNE Configuration

ConfigureTSNEn_subjects $n\_subjects \leq 3$ "Insufficient data" $perplexity \leftarrow \min(30, \lfloor (n\_subjects - 1)/3 \rfloor)$ $perplexity < 1$ $perplexity \leftarrow 1$ Rtsne(data, perplexity = $perplexity$)

Probability Heatmap

The heatmap visualization displays posterior probability distributions across clusters:

```
ggplot(pprob_long, aes(x = Cluster, y = Class,
                       fill = Probability)) +
    geom_tile() +
    scale_fill_gradient(low = "white", high = "blue") +
    facet_wrap(~ Outcome, scales = "free_y")
```

4.4.4 Use Cases and Analytical Workflows

Use Case 1: Multi-outcome Trajectories Discovery

A user can explore preliminary idea of clustering:

1. Select model set using "Select Best Models" button:
2. Choose hierarchical clustering with Ward.D2 linkage
3. Use "Estimate optimal k" to determine cluster number via silhouette analysis
4. Review t-SNE projection for visual cluster separation
5. Examine demographic profiles to characterize identified trajectories
6. Export cluster assignments for external validation

Use Case 2: Threshold-Based Classification Refinement

Refining cluster assignments based on posterior probability thresholds:

1. Perform initial clustering with specific k
2. Navigate to "Threshold-based classification" section
3. Set probability threshold to 0.7 for high-certainty classification
4. Review subjects with MaxProb < threshold indicating ambiguous assignment
5. Calculate certainty scores as MaxProb - SecondProb
6. Identify boundary subjects requiring clinical review
7. Export filtered assignments for additional analysis

4.4.5 Additional Functional Features

Dynamic Trajectory Aggregation

The trajectory aggregation system handles missing data and variable time points:

```
aggregate_trajectories <- function(data, clusters) {
    data %>%
        group_by(Cluster, Outcome, Time) %>%
        summarise(
            mean_y = mean(Value_norm, na.rm = TRUE),
            se_y = sd(Value_norm, na.rm = TRUE) / sqrt(n()),
            n_obs = n(),
            .groups = "drop"
        ) %>%
        filter(n_obs >= min_observations)
}
```

Cluster Quality Metrics Dashboard

The cluster overview table provides comprehensive quality metrics:

- **Silhouette Width:** Within-cluster cohesion vs separation

- **Maximum Posterior Probability:** Assignment certainty
- **Entropy:** Classification uncertainty measure
- **Cluster Centers:** Mean posterior probability profiles
- **Within-Cluster Variance:** Homogeneity assessment

4.5 System Integration and Deployment

The three applications are deployed using Shiny Server Open Source on a Linux-based infrastructure with the following configuration:

```
# /etc/shiny-server/shiny-server.conf
run_as shiny;
server {
    listen 3838;
    location /link_function {
        app_dir /srv/shiny-server/link_function;
        log_dir /var/log/shiny-server;
    }
    location /class_enum {
        app_dir /srv/shiny-server/class_enum;
        log_dir /var/log/shiny-server;
    }
    location /clustering {
        app_dir /srv/shiny-server/clustering;
        log_dir /var/log/shiny-server;
    }
}
```

4.6 Conclusion

The three Shiny modules introduced in this chapter provide an interactive multi-module framework for enhanced analysis of LCMM models and for the visualization of clustering results. The system architecture is built on Shiny's reactive graph, ensuring consistency across multiple analytical components while reducing computational overhead through strategic caching and lazy evaluation. The technical contributions of this framework include creative methods for multi-outcome clustering via posterior probability integration,

adaptive parameter selection in dimensionality reduction algorithms, and efficient visualization of high-dimensional trajectory data. Importantly, the clustering module is not limited to posterior matrices; it operates primarily with two key data sources—cluster assignments and subject characteristics—which makes it adaptable for analyzing different datasets and applicable to similar research tasks. For instance, the module can be configured in a time-efficient way to analyze different probabilistic models.

Chapter 5

Discussion and Contribution

This research identifies six multi-outcomes trajectories among **2,337** subjects with TBI, contributing to change the conventional understanding of TBI recovery and extending previous works where often studies focused on specific domains– such as fatigue [22], function independence [27] or psychiatric symptoms [1] —whereas our study integrates seventeen outcome measures across multiple domains to provide a wider spectrum of recovery. The heterogeneity revealed through latent class growth analysis demonstrates that trajectory is determined not merely by injury severity but by complex interactions between pre-injury vulnerabilities, demographic factors, and injury characteristics. This observation resonates with concerns expressed in the reviewed literature regarding overreliance on severity as a prognostic indicator. For instance, Maas *et al.* (2022) [4] emphasize that TBI is a "complex, multidimensional syndrome" in which secondary factors modulate the outcome progression. Similar conclusions were drawn by MacDonald *et al.* (2022) [26] and Molteni *et al.* (2021) [27], emphasized that recovery trajectories reflect the interplay between injury severity, demographic characteristics, and psychosocial context rather than linear relationships with acute clinical metrics.

The identification of two complete mortality clusters (Cluster 5 and Cluster 6, comprising **172** subjects) necessitates a new conception of TBI research methodology. None of the reviewed studies explicitly modeled death as a distinct trajectory outcome. For example, both MacDonald *et al.* (2022) [26] and Kowalski *et al.* (2021) [21] excluded deceased participants or censored data at last observation, introducing survival bias. Our study addresses that methodological limitation by incorporating mortality as a trajectory endpoint, aligning with concerns raised in the scoping review regarding imputation-related distortion of longitudinal outcomes [19, 29, 31].

This dichotomy mirrors previous findings from mild TBI cohorts, where psychological

and behavioral factors predicted persistent symptoms despite mild injuries. De Koning *et al.* (2018) [23] identified a "secondary complaints" trajectory marked by chronic fatigue and affective distress, while Rakers *et al.* (2021) [22] described fatigue–distress clusters resistant to conventional rehabilitation. Our results extend these observations by showing that pre-injury psychiatric history and lower educational attainment jointly define the non-improving trajectory. For instance, Cluster 1 optimal recovery trajectory—despite **86.9%** mild TBI—contrasts sharply with Cluster 3 persistent symptom burden across all domains, though both have predominantly mild injuries. The differentiating factors seem to be psychiatric history (**18.2% vs 39.7%**), educational attainment (**14.09 vs 12.78 years**), and prior TBI exposure (**24.1% vs 33.8%**). Cluster 3 complete resistance to improvement—with flat trajectories across anxiety, depression, post-concussion symptoms, and quality of life measures—suggests that conventional rehabilitation paradigms fail when neurological injury intersects with pre-existing psychiatric vulnerability. This population requires fundamentally different approaches integrating psychiatric and neurological care from initial presentation, not as an adjunct but as the primary treatment framework.

The temporal architecture of recovery reveals critical intervention windows previously obscured by aggregate analyses. The three-month inflection point, consistently observed across multiple clusters and outcome domains, represents a biological and functional turning point where recovery trajectories crystallize into either continued improvement or chronic disability. This finding has immediate clinical implications: intensive interventions must be front-loaded within this critical window, as trajectories rarely improve after this consolidation phase. Equally important, the complete mortality of Cluster 6—elderly subjects with predominantly "mild" TBI who survived initial hospitalization only to deteriorate fatally—exposes the inadequacy of traditional severity classifications in predicting outcomes for vulnerable populations, particularly those over **55** years with fall-related injuries.

The methodological implications of the mortality clusters extend beyond this study to challenge fundamental assumptions in longitudinal TBI research. The impossible paradox observed in Cluster 5—where GOSE Score scores indicate vegetative state while DRS Score scores suggest functional improvement—exemplifies how standard imputation procedures create artificial trajectories that obscure the reality of fatal outcomes. This finding mandates that future TBI outcome studies employ competing risk models that treat death as a distinct endpoint rather than missing data, and that trajectory analysis explicitly stratifies by survival status. The apparent "good performance" on subjective measures in mortality clusters represents not recovery but the absence of data, highlighting how survivor-ship bias and differential imputation can create entirely fictitious recovery patterns.

The clinical translation of these findings might require fundamental restructuring of

TBI care pathways. Current practice typically stratifies subjects by initial injury severity—mild, moderate, severe—but our analysis demonstrates that demographic and pre-injury factors may be equally or more predictive of outcomes. Early identification of cluster membership through readily available baseline characteristics (age, psychiatric history, education, injury mechanism) could enable precision medicine approaches: Cluster 1 subjects could enter accelerated return-to-function protocols; Cluster 3 subjects would receive immediate integrated psychiatric-neurological care; Cluster 4 subjects with severe injury but absent psychiatric comorbidity would receive sustained intensive rehabilitation knowing that meaningful recovery remains possible; and elderly subjects with fall-related injuries matching Cluster 5 and Cluster 6 profiles would receive aggressive monitoring for delayed complications and realistic prognostication regarding mortality risk.

The multidimensional nature of recovery patterns revealed across **17** outcome measures demonstrates that no single metric adequately captures TBI complex progression. The dissociation between domains—where cognitive recovery does not predict functional outcomes, and psychiatric improvement does not ensure quality of life gains—mandates comprehensive assessment batteries rather than reliance on global measures alone. This multidimensional heterogeneity also explains why uniform treatment protocols show variable efficacy: interventions effective for one cluster may be irrelevant or harmful for another. The persistent elevation of all symptom domains in Cluster 3, despite standard care, suggests that some populations require entirely novel therapeutic approaches that address the synergistic interaction between brain injury and psychiatric vulnerability rather than treating these as separate conditions.

Overall, this discussion aligns with the broader trends identified in the scoping review, where methodological innovation has revealed reproducible patterns of recovery heterogeneity across diverse cohorts [18, 19]. This research's primary contribution lies not in confirming TBI heterogeneity—which clinicians have long recognized—but in empirically demonstrating that this heterogeneity follows predictable patterns identifiable at presentation. The six clusters represent not random variation but distinct clinical phenotypes with characteristic trajectories, vulnerability factors, and intervention needs. By moving beyond average treatment effects to cluster-specific patterns, this work enables precision medicine approaches that match intervention intensity and focus to individual recovery phenotypes, potentially improving outcomes while optimizing resource allocation in a healthcare system strained by the complexity and chronicity of TBI sequelae.

Chapter 6

Limitations and Future Considerations

This study provides further evidence of the importance of analyzing recovery trajectories in order to advance an understanding of subjects recovery patterns. However, as with all data-driven research at the intersection of public health and data science, there are important limitations that should be acknowledged. Below, we describe the key limitations in this work, focusing on methodological choices, statistical considerations, data-related challenges, and interpretative constraints.

Methodological limitations. One of the central challenges we faced relates to the methodological choices for trajectory modeling. Currently, as our literature review showed, very limited research has been conducted on multi-outcome trajectory modeling specifically for TBI populations, meaning that our approach had to be largely exploratory. Our initial strategy involved using the LCMMs framework, which is particularly well-suited for heterogeneous subjects cohorts. We first implemented single-outcome models, but a key limitation quickly emerged: each of the selected outcomes under investigation produced a different optimal number of trajectory classes. This diversity of trajectory structures prevented us from directly applying a straightforward multivariate or joint LCMMs, which assumes a common class structure across all outcomes [14].

As a result, we turned to clustering methods as an alternative way of integrating posterior probabilities from single-outcome LCMMs. Specifically, we selected hierarchical clustering because it is widely used in biomedical research, relatively interpretable, and less dependent on heavy data transformation [47]. While this choice was practical, it also brought methodological compromises. The posterior probabilities from LCMMs models are compositional data (vectors whose components sum to one), a structure that classical

clustering methods like hierarchical clustering or k-means do not explicitly account for. Although these methods remain popular, their statistical properties are difficult to interpret in settings with compositional constraints [48].

For the purpose of this study, hierarchical clustering served as a sufficient and pragmatic approach. Nevertheless, more advanced probabilistic clustering strategies, such as the Dirichlet mixture model framework, may provide a better methodological fit in future work [47]. Such models can explicitly accommodate probability-based data, allow more nuanced handling of multidimensional structures, and reduce the risk of artificially hard subjects assignments.

Data-related limitations. The second domain of limitation is directly linked to the dataset. We relied on the TRACK-TBI, which is one of the most comprehensive and widely used datasets in TBI research. Despite its strengths, the dataset presents several well-documented challenges that affected our analysis. In particular, longitudinal follow-up assessments exhibited substantial missing data across different outcomes and time points, and some participants had complete outcome loss for various reasons, ranging from dropouts to attrition due to clinical severity. As described in 3.1.3, missingness was especially problematic for certain outcomes at the six-month follow-up window.

We applied multiple imputation strategies to address data gaps; however, the extent of missingness at some time points limited our ability to conduct thorough robustness checks. This constrains confidence in certain derived trajectories. In future research, more advanced imputation techniques—particularly domain-specific approaches designed for clinical longitudinal data—could improve data quality and better preserve statistical power.

Another related limitation is generalizability. The TRACK-TBI cohort, while diverse, does not fully represent the global TBI population. It is constrained to the US. Level I trauma center patients and may not capture recovery patterns in patients treated in other healthcare environments or in different cultural and socioeconomic contexts. This limitation means that our findings should be interpreted within the boundaries of the TRACK-TBI cohort rather than being generalized universally. A promising direction for future research is to combine multiple large-scale datasets across different clinical settings and countries, thereby broadening representation and ensuring greater external validity.

Interpretation. A further limitation concerns the interpretation of the modeled trajectories. Describing recovery pathways in TBI is inherently complex, requiring not only statistical insight but also deep clinical expertise. In this study, we restricted our analysis to the key demographic and clinical variables available within our timeline. This limitation means that additional aspects of patient profiles that may influence recovery—including

comorbidities, rehabilitation environments, or psychosocial factors—were not included, potentially narrowing interpretability. In addition, post-discharge treatment information, such as rehabilitation intensity, therapy duration, and outpatient support, was not available in the dataset. The absence of such data limits the ability to account for variability in recovery that may stem from differences in post-acute care. Future studies incorporating standardized indicators, such as the Therapy Intensity Level (TIL) index, could help clarify how treatment exposure after discharge contributes to long-term recovery heterogeneity.

Another important consideration relates to the temporal frequency of data collection. The follow-up schedule in the TRACK-TBI dataset includes relatively wide intervals between assessment points (e.g., at 2 weeks, 3 months, 6 months, and 12 months), which constrains the temporal precision with which inflection points in the recovery trajectories can be identified. For instance, while the 3-month time point consistently appears as a major turning point in our analyses, it should be interpreted as an approximate period of transition rather than a precise inflection moment. Denser longitudinal sampling in future studies would allow more accurate characterization of the timing and shape of recovery changes over time.

Moreover, because we used clustering-based methods, subjects grouping required hard class assignments. While we attempted to provide some visualization of uncertainty by including confidence intervals, the method does not allow soft assignment or partial membership across classes in the multi-outcome analysis. This creates the possibility that some patients could be misaligned across groups, especially when confidence intervals of different trajectories substantially overlapped. Nonetheless, given that such subjects still exhibit broadly similar recovery patterns, the clinical implications remain informative. Furthermore, our initial reliance on LCMMs models, which are considered state-of-the-art for trajectory analysis, helped mitigate this concern by grounding our approach in established statistical methodology.

Another promising direction for future research involves modeling treatment as a time-varying covariate within trajectory analyses. Because rehabilitation intensity, pharmacological interventions, and psychological support can change throughout the recovery period, incorporating these dynamic effects could reveal how treatment modifications influence recovery trajectories over time.

Beyond these extensions, future work could also explore *multi-state modeling* as a complementary analytical framework. While the competing-risk structure adopted in this study accounts for mortality as a terminal event, multi-state models enable the representation of transitions between multiple clinically meaningful states—such as partial recovery, sustained disability, or death—rather than focusing solely on final endpoints. This approach allows for the estimation of transition-specific hazards, transition probabilities, and covari-

ate effects, thereby offering a more advanced understanding of the temporal dynamics of recovery.

In summary, while this study represents an important step toward understanding the multidimensional recovery patterns of TBI, it is inevitably constrained by methodological, data-related, and interpretive limitations. Acknowledging these boundaries provides a foundation for meaningful refinement in future research—through methodological innovation, improved data collection practices, and broader collaborations that bridge statistical modeling with clinical insight.

Chapter 7

Conclusion

This research challenges the idea of TBI as a uniform condition. Using data from TRACK-TBI and applying LCMMs together with hierarchical clustering, six distinct recovery trajectories were identified. These results show that TBI recovery is highly heterogeneous and cannot be explained by injury severity alone.

Several important findings emerged. **First**, patients with similar initial injury can follow very different recovery paths, highlighting the role of pre-injury factors and psychiatric history. In particular, Cluster 3 showed little to no recovery despite mainly mild injuries, suggesting the need for early integrated psychiatric and neurological care. **Second**, the analysis revealed a consistent "turning point" around three months after injury, when most recovery trajectories stabilize. This emphasizes the importance of intensive intervention during the early recovery window. The study also identified two clusters defined by mortality, showing that traditional outcome measures can misrepresent fatal trajectories. These results can support a shift from severity-based treatment pathways toward more personalized care based on demographic and pre-injury characteristics. Recognizing cluster membership at baseline could allow clinicians to tailor rehabilitation intensity and focus, moving closer to precision medicine in TBI care.

Methodologically, the study demonstrates the value of combining LCMMs with clustering on the posterior probability for analyzing complex, multi-outcome longitudinal data. The interactive tools developed facilitate the use of such approaches in TBI.

Nevertheless, limitations include the geographic scope of the TRACK-TBI cohort, challenges with missing data, and the constraints of clustering compositional outcomes. Future research should expand to multi-site transition modeling, develop improved clustering methods, and explore causal mechanisms of exposure to different injury characteristics and treatments.

In summary, this thesis shows that TBI recovery follows **predictable, cluster-specific patterns**. By identifying these trajectories early, healthcare systems can provide more effective, tailored interventions, ultimately improving patient outcomes and resource use.

References

[1] Emerson M Wickwire, Jennifer S Albrecht, Vincent F Capaldi, 2nd, Sonia O Jain, Raquel C Gardner, J Kent Werner, Pratik Mukherjee, Ashlee B McKeon, Michael T Smith, Joseph T Giacino, Lindsay D Nelson, Scott G Williams, Jacob Collen, Xiaoying Sun, David M Schnyer, Amy J Markowitz, Geoffrey T Manley, Andrew D Krystal, and Transforming Research and Clinical Knowledge in Traumatic Brain Injury (TRACK-TBI) Investigators. Trajectories of insomnia in adults after traumatic brain injury. *JAMA Netw. Open*, 5(1):e2145310, January 2022. doi: 10.1001/jamanetworkopen.2021.45310.

[2] Michael C Dewan, Abbas Rattani, Saksham Gupta, Ronnie E Baticulon, Ya-Ching Hung, Maria Punchak, Amit Agrawal, Amos O Adeleye, Mark G Shrime, Andrés M Rubiano, Jeffrey V Rosenfeld, and Kee B Park. Estimating the global incidence of traumatic brain injury. *J. Neurosurg.*, 130(4):1080–1097, April 2019. doi: 10.3171/2017.10.JNS17352.

[3] GBD 2016 Traumatic Brain Injury and Spinal Cord Injury Collaborators. Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990-2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol.*, 18(1):56–87, January 2019. doi: 10.1016/S1474-4422(18)30415-0.

[4] Andrew I R Maas, David K Menon, P David Adelson, Nada Andelic, Michael J Bell, Antonio Belli, Peter Bragge, Alexandra Brazinova, András Büki, Randall M Chesnut, Giuseppe Citerio, Mark Coburn, D Jamie Cooper, A Tamara Crowder, Endre Czeiter, Marek Czosnyka, Ramon Diaz-Arrastia, Jens P Dreier, Ann-Christine Duhaime, Ari Ercole, Thomas A van Essen, Valery L Feigin, Guoyi Gao, Joseph Giacino, Laura E Gonzalez-Lara, Russell L Gruen, Deepak Gupta, Jed A Hartings, Sean Hill, Ji-Yao Jiang, Naomi Ketharanathan, Erwin J O Kompanje, Linda Lanyon, Steven Laureys, Fiona Lecky, Harvey Levin, Hester F Lingsma, Marc Maegele, Marek Majdan, Geoffrey Manley, Jill Marsteller, Luciana Mascia, Charles McFadyen, Stefania Mondello, Virginia Newcombe, Aarno Palotie, Paul M Parizel, Wilco Peul, James

Piercy, Suzanne Polinder, Louis Puybasset, Todd E Rasmussen, Rolf Rossaint, Peter Smielewski, Jeannette Söderberg, Simon J Stanworth, Murray B Stein, Nicole von Steinbüchel, William Stewart, Ewout W Steyerberg, Nino Stocchetti, Anneliese Synnot, Braden Te Ao, Olli Tenovuo, Alice Theadom, Dick Tibboel, Walter Videtta, Kevin K W Wang, W Huw Williams, Lindsay Wilson, Kristine Yaffe, and InTBIR Participants and Investigators. Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research. *Lancet Neurol.*, 16(12):987–1048, December 2017. doi: 10.1016/S1474-4422(17)30371-X.

[5] Raquel C Gardner, Kristen Dams-O'Connor, Molly Rose Morrissey, and Geoffrey T Manley. Geriatric traumatic brain injury: Epidemiology, outcomes, knowledge gaps, and future directions. *J. Neurotrauma*, 35(7):889–906, April 2018. doi: 10.1089/neu.2018.5864.

[6] Kaj Blennow, David L Brody, Patrick M Kochanek, Harvey Levin, Ann McKee, Gerard M Ribbers, Kristine Yaffe, and Henrik Zetterberg. Traumatic brain injuries. *Nat. Rev. Dis. Primers*, 2(1):16084, November 2016. doi: 10.1038/nrdp.2016.84.

[7] Jun-Feng Zou, Hai-Lan Fang, Jing Zheng, Yu-Qiang Ma, Chu-Wei Wu, Gao-Jian Su, Xian-Sheng Liu, Jun Liu, Jie Gao, Jie-Hua Zhang, Dong-Liang Zhu, Xin Shi, and Xian-Jian Huang. The epidemiology of traumatic brain injuries in the fastest-paced city in china: a retrospective study. *Front. Neurol.*, 14:1255117, November 2023. doi: 10.3389/fneur.2023.1255117.

[8] Anne Williams, Ellie Fossey, John Farhall, Fiona Foley, and Neil Thomas. Impact of jointly using an e-mental health resource (Self-Management and recovery technology) on interactions between service users experiencing severe mental illness and community mental health workers: Grounded theory study. *JMIR Ment. Health*, 8(6):e25998, June 2021. doi: 10.2196/25998.

[9] Andrew I R Maas, David K Menon, Geoffrey T Manley, Mathew Abrams, Cecilia Åkerlund, Nada Andelic, Marcel Aries, Tom Bashford, Michael J Bell, Yelena G Bodien, Benjamin L Brett, András Büki, Randall M Chesnut, Giuseppe Citerio, David Clark, Betony Clasby, D Jamie Cooper, Endre Czeiter, Marek Czosnyka, Kristen Dams-O'Connor, Véronique De Keyser, Ramon Diaz-Arrastia, Ari Ercole, Thomas A van Essen, Éanna Falvey, Adam R Ferguson, Anthony Figaji, Melinda Fitzgerald, Brandon Foreman, Dashiell Gantner, Guoyi Gao, Joseph Giacino, Benjamin Gravesteijn, Fabian Guiza, Deepak Gupta, Mark Gurnell, Juanita A Haagsma, Flora M Hammond, Gregory Hawryluk, Peter Hutchinson, Mathieu van der Jagt, Sonia Jain, Swati Jain, Ji-Yao Jiang, Hope Kent, Angelos Kolias, Erwin J O Kompanje,

Fiona Lecky, Hester F Lingsma, Marc Maegele, Marek Majdan, Amy Markowitz, Michael McCrea, Geert Meyfroidt, Ana Mikolić, Stefania Mondello, Pratik Mukherjee, David Nelson, Lindsay D Nelson, Virginia Newcombe, David Okonkwo, Matej Orešič, Wilco Peul, Dana Pisică, Suzanne Polinder, Jennie Ponsford, Louis Puybasset, Rahul Raj, Chiara Robba, Cecilie Røe, Jonathan Rosand, Peter Schueler, David J Sharp, Peter Smielewski, Murray B Stein, Nicole von Steinbüchel, William Stewart, Ewout W Steyerberg, Nino Stocchetti, Nancy Temkin, Olli Tenovuo, Alice Theadom, Ilias Thomas, Abel Torres Espin, Alexis F Turgeon, Andreas Unterberg, Dominique Van Praag, Ernest van Veen, Jan Verheyden, Thijs Vande Vyvere, Kevin K W Wang, Eveline J A Wiegers, W Huw Williams, Lindsay Wilson, Stephen R Wisniewski, Alexander Younsi, John K Yue, Esther L Yuh, Frederick A Zeiler, Marina Zeldovich, Roger Zemek, and InTBIR Participants and Investigators. Traumatic brain injury: progress and challenges in prevention, clinical care, and research. *Lancet Neurol.*, 21(11):1004–1060, November 2022. doi: 10.1016/S1474-4422(22)00309-X.

[10] Gabrielle F Miller, Lara DePadilla, and Likang Xu. Costs of nonfatal traumatic brain injury in the united states, 2016. *Med. Care*, 59(5):451–455, May 2021. doi: 10.1097/MLR.0000000000001511.

[11] Rabea Iris Pantelatos, Jonas Stenberg, Turid Follestad, Oddrun Sandrød, Cathrine Elisabeth Einarsen, Anne Vik, and Toril Skandsen. Improvement in functional outcome from 6 to 12 months after moderate and severe traumatic brain injury is frequent, but may not be detected with the glasgow outcome scale extended. *Neurotrauma Rep.*, 5(1):139–149, February 2024. doi: 10.1089/neur.2023.0109.

[12] Dhaval Shukla, B Indira Devi, and Amit Agrawal. Outcome measures for traumatic brain injury. *Clin. Neurol. Neurosurg.*, 113(6):435–441, July 2011. doi: 10.1016/j.clineuro.2011.02.013.

[13] Raquel C Gardner, Jing Cheng, Adam R Ferguson, Ross Boylan, John Boscardin, Ross D Zafonte, Geoffrey T Manley, and Citicoline Brain Injury Treatment Trial Investigators. Divergent six month functional recovery trajectories and predictors after traumatic brain injury: Novel insights from the citicoline brain injury treatment trial study. *J. Neurotrauma*, 36(17):2521–2532, September 2019. doi: 10.1089/neu.2018.6167.

[14] Cécile Proust-Lima, Viviane Philipps, and Benoit Liquet. Estimation of extended mixed models using latent classes and latent processes: The R package lcmm. *J. Stat. Softw.*, 78(2), 2017. doi: 10.18637/jss.v078.i02.

[15] John K Yue, Mary J Vassar, Hester F Lingsma, Shelly R Cooper, David O Okonkwo, Alex B Valadka, Wayne A Gordon, Andrew I R Maas, Pratik Mukherjee, Esther L Yuh, Ava M Puccio, David M Schnyer, Geoffrey T Manley, and TRACK-TBI Investigators. Transforming research and clinical knowledge in traumatic brain injury pilot: multicenter implementation of the common data elements for traumatic brain injury. *J. Neurotrauma*, 30(22):1831–1844, November 2013. doi: 10.1089/neu.2013.2970.

[16] Amit Saxena, Mukesh Prasad, Akshansh Gupta, Neha Bharill, Om Prakash Patel, Aruna Tiwari, Meng Joo Er, Weiping Ding, and Chin-Teng Lin. A review of clustering techniques and developments. *Neurocomputing*, 267:664–681, December 2017. doi: 10.1016/j.neucom.2017.06.053.

[17] Hilary Arksey and Lisa O'Malley. Scoping studies: towards a methodological framework. *Int. J. Soc. Res. Methodol.*, 8(1):19–32, February 2005. doi: https://doi.org/10.1080/1364557032000119616.

[18] Karen A Martinez, Ehri Ryu, Christopher J Patrick, Nancy R Temkin, Murray B Stein, Brooke E Magnus, Michael A McCrea, Geoffrey T Manley, Lindsay D Nelson, and TRACK-TBI Investigators. Distinct trajectories of neuropsychiatric symptoms in the 12 months following traumatic brain injury (TBI): a TRACK-TBI study. *Psychol. Med.*, 54(11):3089–3098, August 2024. doi: 10.1017/S0033291724001211.

[19] Hamid Ghaderi, Brandon Foreman, Chandan K Reddy, and Vignesh Subbian. Discovery of generalizable TBI phenotypes using multivariate time-series clustering. *Comput. Biol. Med.*, 180(108997):108997, September 2024. doi: doi.org/10.1016/j.compbiomed.2024.108997.

[20] Rhonda Orr, Tina Bogg, Andrew Fyffe, Lawrence T Lam, and Gary J Browne. Graded exercise testing predicts recovery trajectory of concussion in children and adolescents. *Clin. J. Sport Med.*, 31(1):23–30, January 2021. doi: 10.1097/JSM.0000000000000683.

[21] Robert G Kowalski, Flora M Hammond, Alan H Weintraub, Risa Nakase-Richardson, Ross D Zafonte, John Whyte, and Joseph T Giacino. Recovery of consciousness and functional outcome in moderate and severe traumatic brain injury. *JAMA Neurol.*, 78(5):548–557, May 2021. doi: 10.1001/jamaneurol.2021.0084.

[22] Sandra E Rakers, Marieke E Timmerman, Myrthe E Scheenen, Myrthe E de Koning, Harm J van der Horn, Joukje van der Naalt, and Jacoba M Spikman. Trajectories of fatigue, psychological distress, and coping styles after mild traumatic brain injury: A 6-month prospective cohort study. *Arch. Phys. Med. Rehabil.*, 102(10):1965–1971.e2, October 2021. doi: 10.1016/j.apmr.2021.06.004.

[23] M E De Koning, M E Scheenen, H J Van Der Horn, J M Spikman, and J Van Der Naalt. From 'miserable minority' to the 'fortunate few': the other end of the mild traumatic brain injury spectrum. *Brain Inj.*, 32(5):540–543, April 2018. doi: doi.org/10.1080/02699052.2018.1431844.

[24] Chiara S Haller. Twelve-month prospective cohort study of patients with severe traumatic brain injury and their relatives: Coping, satisfaction with life and neurological functioning. *Brain Inj.*, 31(13-14):1903–1909, December 2017. doi: 10.1080/02699052.2017.1346295.

[25] Juan Lu, Cecilie Roe, Solrun Sigurdardottir, Nada Andelic, and Marit Forslund. Trajectory of functional independent measurements during first five years after moderate and severe traumatic brain injury. *J. Neurotrauma*, 35(14):1596–1603, July 2018. doi: 10.1089/neu.2017.5299.

[26] Christine L Mac Donald, Jason Barber, Ann Johnson, Jana Patterson, and Nancy Temkin. Global disability trajectories over the first decade following combat concussion. *J. Head Trauma Rehabil.*, 37(2):63–70, 2022. doi: 10.1097/HTR.0000000000000738.

[27] Erika Molteni, Marta Bianca Maria Ranzini, Elena Beretta, Marc Modat, and Sandra Strazzer. Individualized prognostic prediction of the long-term functional trajectory in pediatric acquired brain injury. *J. Pers. Med.*, 11(7):675, July 2021. doi: 10.3390/jpm11070675.

[28] Joshua P McGeown, Stephen Kara, Mark Fulcher, Hannah Crosswell, Robert Borotkanics, Patria A Hume, Kenneth L Quarrie, and Alice Theadom. Predicting sport-related mTBI symptom resolution trajectory using initial clinical assessment findings: A retrospective cohort study. *Sports Med.*, 50(6):1191–1202, June 2020. doi: 10.1007/s40279-019-01240-4.

[29] Ross C Puffer, John K Yue, Matthew Mesley, Julia B Billigen, Jane Sharpless, Anita L Fetzick, Ava Puccio, Ramon Diaz-Arrastia, and David O Okonkwo. Long-term outcome in traumatic brain injury patients with midline shift: a secondary analysis of the phase 3 COBRIT clinical trial. *J. Neurosurg.*, 131(2):596–603, August 2019. doi: 10.3171/2018.2.JNS173138.

[30] Sara M Lippa, Rael T Lange, Jason M Bailie, Jan E Kennedy, Tracey A Brickell, and Louis M French. Utility of the validity-10 scale across the recovery trajectory following traumatic brain injury. *J. Rehabil. Res. Dev.*, 53(3):379–390, 2016. doi: 10.1682/JRRD.2015.01.0009.

[31] Lien A Bui, Dacosta Yeboah, Louis Steinmeister, Sima Azizi, Daniel B Hier, Donald C Wunsch, Gayla R Olbricht, and Tayo Obafemi-Ajayi. Heterogeneity in blood biomarker trajectories after mild TBI revealed by unsupervised learning. *IEEE/ACM Trans. Comput. Biol. Bioinform.*, 19(3):1365–1378, May 2022. doi: 10.1109/TCBB.2021.3091972.

[32] Geoffrey T Manley, Kristen Dams-O'Connor, Michael L Alosco, Hibah O Awwad, Jeffery J Bazarian, Peter Bragge, John D Corrigan, Adele Doperalski, Adam R Ferguson, Christine L Mac Donald, David K Menon, Molly M McNett, Joukje van der Naalt, Lindsay D Nelson, Dana Pisică, Noah D Silverberg, Nsini Umoh, Lindsay Wilson, Esther L Yuh, Henrik Zetterberg, Andrew I R Maas, Michael A McCrea, and NIH-NINDS TBI Classification and Nomenclature Initiative. A new characterisation of acute traumatic brain injury: the NIH-NINDS TBI classification and nomenclature initiative. *Lancet Neurol.*, 24(6):512–523, June 2025.

[33] Misti Timpson, Erinn M Hade, Cynthia Beaulieu, Susan D Horn, Flora M Hammond, Juan Peng, Erin Montgomery, Clare Giuffrida, Kamie Gilchrist, Aubrey Lash, Marcel Dijkers, John D Corrigan, and Jennifer Bogner. Advanced therapy in traumatic brain injury inpatient rehabilitation: Effects on outcomes during the first year after discharge. *Arch. Phys. Med. Rehabil.*, 100(10):1818–1826, October 2019. doi: 10.1016/j.apmr.2018.11.015.

[34] Hermine Lore Nguena Nguefack, M Gabrielle Pagé, Joel Katz, Manon Choinière, Alain Vanasse, Marc Dorais, Oumar Mallé Samb, and Anaïs Lacasse. Trajectory modelling techniques useful to epidemiological research: A comparative narrative review of approaches. *Clin. Epidemiol.*, 12:1205–1222, October 2020.

[35] Michael A McCrea, Joseph T Giacino, Jason Barber, Nancy R Temkin, Lindsay D Nelson, Harvey S Levin, Sureyya Dikmen, Murray Stein, Yelena G Bodien, Kim Boase, Sabrina R Taylor, Mary Vassar, Pratik Mukherjee, Claudia Robertson, Ramon Diaz-Arrastia, David O Okonkwo, Amy J Markowitz, Geoffrey T Manley, TRACK-TBI Investigators, Opeolu Adeoye, Neeraj Badjatia, M Ross Bullock, Randall Chesnut, John D Corrigan, Karen Crawford, Ann-Christine Duhaime, Richard Ellenbogen, V Ramana Feeser, Adam R Ferguson, Brandon Foreman, Raquel Gardner, Etienne Gaudette, Dana Goldman, Luis Gonzalez, Shankar Gopinath, Rao Gullapalli, J Claude Hemphill, Gillian Hotz, Sonia Jain, C Dirk Keene, Frederick K Korley, Joel ramer, Natalie Kreitzer, Chris Lindsell, Joan Machamer, Christopher Madden, Alastair Martin, Thomas McAllister, Randall Merchant, Laura B Ngwenya, Florence Noel, Amber Nolan, Eva Palacios, Daniel Perl, Ava Puccio, Miri Rabinowitz, Jonathan Rosand, Angelle Sander, Gabriella Satris, David Schnyer, Seth Seabury, Mark Sherer, Arthur

Toga, Alex Valadka, Kevin Wang, John K Yue, Esther Yuh, and Ross Zafonte. Functional outcomes over the first year after moderate to severe traumatic brain injury in the prospective, longitudinal TRACK-TBI study. *JAMA Neurol.*, 78(8):982–992, August 2021. doi: 10.1001/jamaneurol.2021.2043.

[36] P David Adelson, Jose Pineda, Michael J Bell, Nicholas S Abend, Rachel P Berger, Christopher C Giza, Gillian Hotz, Mark S Wainwright, and Pediatric TBI Demographics and Clinical Assessment Working Group. Common data elements for pediatric traumatic brain injury: recommendations from the working group on demographics and clinical assessment. *J. Neurotrauma*, 29(4):639–653, March 2012. doi: 10.1089/neu.2011.1952.

[37] Sophie Richter, Susan Stevenson, Tom Newman, Lindsay Wilson, Andrew I R Maas, Daan Nieboer, Hester Lingsma, Ewout W Steyerberg, and Virginia F J Newcombe. Study design features associated with patient attrition in studies of traumatic brain injury: A systematic review. *J. Neurotrauma*, 37(17):1845–1853, September 2020. doi: 10.1089/neu.2020.7000.

[38] Jessica L Nielson, Shelly R Cooper, Seth A Seabury, Davide Luciani, Anthony Fabio, Nancy R Temkin, Adam R Ferguson, and TRACK-TBI Investigators. Statistical guidelines for handling missing data in traumatic brain injury clinical research. *J. Neurotrauma*, 38(18):2530–2537, September 2021. doi: 10.1089/neu.2019.6702.

[39] Xuan Vinh To, Hien D Nguyen, Paul Cumming, Peter Curpen, and Fatima A Nasrallah. Contemporaneous outcomes multiple imputation (KOMI) can communicate missing glasgow outcome scale - extended (GOSE) scores: A multi-domain imputation for GOSE in the transforming research and clinical knowledge in traumatic brain injury (TRACK-TBI) data. *J. Neurol. Sci.*, 475(123564):123564, August 2025. doi: https://doi.org/10.1016/j.jns.2025.123564.

[40] Nan M Laird and James H Ware. Random-effects models for longitudinal data. *Biometrics*, pages 963–974, 1982.

[41] Hannah Lennon, Scott Kelly, Matthew Sperrin, Iain Buchan, Amanda J Cross, Michael Leitzmann, Michael B Cook, and Andrew G Renehan. Framework to construct and interpret latent class trajectory modelling. *BMJ open*, 8(7):e020683, 2018.

[42] Gavin van der Nest, Valéria Lima Passos, Math JJM Candel, and Gerard JP van Breukelen. An overview of mixture modelling for latent evolutions in longitudinal data: Modelling approaches, fit statistics and software. *Advances in Life Course Research*, 43:100323, 2020.

[43] Fionn Murtagh and Pierre Legendre. Ward's hierarchical agglomerative clustering method: Which algorithms implement ward's criterion? *J. Classif.*, 31(3):274–295, October 2014. doi: 10.1007/s00357-014-9161-z.

[44] Joe H Ward, join(' '. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.*, 58(301):236, March 1963. doi: 10.2307/2282967.

[45] Peter J Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20:53–65, November 1987. doi: 10.1016/0377-0427(87)90125-7.

[46] Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. *shiny: Web Application Framework for R*. URL https://github.com/rstudio/shiny.

[47] Samyajoy Pal and Christian Heumann. Clustering compositional data using dirichlet mixture model. *PLoS One*, 17(5):e0268438, May 2022. doi: 10.1371/journal.pone.0268438.

[48] Arno Fritsch and Katja Ickstadt. Improved criteria for clustering based on the posterior similarity matrix. *Bayesian Anal.*, 4(2):367–391, June 2009. doi: 10.1214/09-BA414.

[49] Peter C Austin. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav. Res.*, 46(3):399–424, May 2011. doi: 10.1080/00273171.2011.568786.

[50] Daniel G Lynch, Raj K Narayan, and Chunyan Li. Multi-mechanistic approaches to the treatment of traumatic brain injury: A review. *J. Clin. Med.*, 12(6), March 2023. doi: 10.3390/jcm12062179.

[51] Sahithi Reddi, Smita Thakker-Varia, Janet Alder, and Anna O Giarratana. Status of precision medicine approaches to traumatic brain injury. *Neural Regen. Res.*, 17(10):2166–2171, October 2022. doi: 10.4103/1673-5374.335824.

[52] Christopher J Miller, Miya L Barnett, Ana A Baumann, Cassidy A Gutner, and Shannon Wiltsey-Stirman. The FRAME-IS: a framework for documenting modifications to implementation strategies in healthcare. *Implement. Sci.*, 16(1):36, April 2021. doi: 10.1186/s13012-021-01105-3.

[53] Ji Hong Min and Yong-Il Shin. Treatment and rehabilitation for traumatic brain injury: Current update. *Brain Neurorehabilit.*, 15(2):e14, July 2022. doi: 10.12786/bn.2022.15.e14.

[54] Fabrice I Mowbray, Donna Manlongat, and Meghna Shukla. Sensitivity analysis: A method to promote certainty and transparency in nursing and health research. *Can. J. Nurs. Res.*, 54(4):371–376, December 2022. doi: 10.1177/08445621221107108.

[55] Neil H Young and Peter J D Andrews. Developing a prognostic model for traumatic brain injury–a missed opportunity? *PLoS Med.*, 5(8):e168, August 2008. doi: 10.1371/journal.pmed.0050168.

[56] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.

[57] Miceline Mésidor, Marie-Claude Rousseau, Jennifer O'Loughlin, and Marie-Pierre Sylvestre. Does group-based trajectory modeling estimate spurious trajectories? *BMC Med. Res. Methodol.*, 22(1):194, July 2022. doi: https://doi.org/10.1186/s12874-022-01622-9.

APPENDICES

Appendix A

Table 1: List of Articles Included in the Scoping Review

Title	Authors
Heterogeneity in Blood Biomarker Trajectories After Mild TBI Revealed by Unsupervised Learning	Bui *et al.*, 2021 [31]
Trajectories of Fatigue, Psychological Distress, and Coping Styles After Mild TBI: A 6-Month Prospective Cohort Study	Rakers *et al.*, 2021 [22]
Distinct trajectories of neuropsychiatric symptoms in the 12 months following TBI: a TRACK-TBI study	Martinez *et al.*, 2024 [18]
Discovery of Generalizable TBI Phenotypes Using Multivariate Time-Series Clustering	Ghaderi *et al.*, 2024 [19]
Trajectories of Insomnia in Adults After TBI	Wickwire *et al.*, 2022 [1]
From 'miserable minority' to the 'fortunate few': the other end of the mild TBI spectrum	De Koning *et al.*, 2018 [23]
Graded Exercise Testing Predicts Recovery Trajectory of Concussion in Children and Adolescents	Orr *et al.*, 2021 [20]
Recovery of Consciousness and Functional Outcome in Moderate and Severe TBI	Kowalski *et al.*, 2021 [21]
Predicting Sport-related mTBI Symptom Resolution Trajectory Using Initial Clinical Assessment Findings: A Retrospective Cohort Study	McGeown *et al.*, 2019 [28]

Continued on next page

Title	Authors
Twelve-month prospective cohort study of patients with severe TBI and their relatives: Coping, satisfaction with life and neurological functioning	Haller *et al.*, 2017 [24]
Divergent Six Month Functional Recovery Trajectories and Predictors after TBI: Novel Insights from the Citicoline Brain Injury Treatment Trial Study	Gardner *et al.*, 2019 [5]
Trajectory of Functional Independent Measurements during First Five Years after Moderate and Severe TBI	Lu *et al.*, 2018 [25]
Long-term outcome in TBI patients with midline shift: a secondary analysis of the Phase 3 COBRIT clinical trial	Puffer *et al.*, 2019 [29]
Individualized Prognostic Prediction of the Long-Term Functional Trajectory in Pediatric Acquired Brain Injury	Molteni *et al.*, 2021 [27]
Utility of the Validity-10 Scale Across the Recovery Trajectory Following TBI	Lippa *et al.*, 2016 [30]
Global Disability Trajectories Over the First Decade Following Combat Concussion	MacDonald *et al.*, 2022 [26]

Appendix B

Table 2: Comparison of Link Functions for Selected Outcomes

Outcome Variable	Link Function	Conv.	Log-likelihood	Param.	AIC	BIC
BSI-18 Anxiety Score	Linear	Y	$-24,876.81$	6	49,765.63	49,800.17
BSI-18 Anxiety Score	2-quant splines	Y	$-22,061.65$	8	44,139.30	44,185.35
BSI-18 Anxiety Score	**3-quant splines**	**Y**	$-16,064.07$	**9**	32,146.14	32,197.95
BSI-18 Anxiety Score	3-equi splines	Y	$-20,605.67$	9	41,229.35	41,281.15
BSI-18 Anxiety Score	4-equi splines	Y	$-19,864.17$	10	39,748.34	39,805.90

Continued on next page

Table 2 – continued from previous page

Outcome Variable	Link Function	Conv.	Log-likelihood	Param.	AIC	BIC
BSI-18 Anxiety Score	5-equi splines	Y	$-19,354.76$	11	38,731.53	38,794.85
BSI-18 Depression Score	Linear	Y	$-25,149.11$	6	50,310.22	50,344.76
BSI-18 Depression Score	2-quant splines	Y	$-22,471.81$	8	44,959.63	45,005.68
BSI-18 Depression Score	**3-quant splines**	**Y**	$-15,973.42$	**9**	31,964.84	32,016.65
BSI-18 Depression Score	3-equi splines	Y	$-20,944.41$	9	41,906.81	41,958.62
BSI-18 Depression Score	4-equi splines	Y	$-20,075.13$	10	40,170.26	40,227.83
BSI-18 Depression Score	5-equi splines	Y	$-19,525.39$	11	39,072.78	39,136.10
DRS Physical Indep.	Linear	Y	$-22,941.36$	6	45,894.71	45,929.25
DRS Physical Indep.	2-quant splines	Y	$-18,110.92$	8	36,237.83	36,283.89
DRS Physical Indep.	3-equi splines	Y	$-15,873.70$	9	31,765.40	31,817.21
DRS Physical Indep.	4-equi splines	Y	$-14,921.51$	10	29,863.02	29,920.59
DRS Physical Indep.	**5-equi splines**	**Y**	$-14,244.09$	**11**	28,510.18	28,573.50
GOSE Score	Linear	Y	$-16,182.72$	6	32,377.45	32,411.99
GOSE Score	2-quant splines	Y	$-14,919.95$	8	29,855.89	29,901.94
GOSE Score	3-quant splines	Y	$-13,793.90$	9	27,605.79	27,657.60
GOSE Score	**4-quant splines**	**Y**	$-13,560.83$	**10**	27,141.66	27,199.22
GOSE Score	3-equi splines	Y	$-14,583.78$	9	29,185.56	29,237.37
GOSE Score	4-equi splines	Y	$-14,217.20$	10	28,454.41	28,511.97
GOSE Score	5-equi splines	Y	$-13,763.25$	11	27,548.50	27,611.82
ISI Total Score	Linear	Y	$-29,675.55$	6	59,363.11	59,397.65

Continued on next page

Table 2 – continued from previous page

Outcome Variable	Link Function	Conv.	Log-likelihood	Param.	AIC	BIC
ISI Total Score	2-quant splines	Y	$-28,311.84$	8	56,639.69	56,685.74
ISI Total Score	3-quant splines	Y	$-26,625.30$	9	53,268.59	53,320.40
ISI Total Score	4-quant splines	Y	$-25,745.39$	10	51,510.77	51,568.34
ISI Total Score	**5-quant splines**	**Y**	$-25,393.78$	**11**	50,809.55	50,872.88
ISI Total Score	3-equi splines	Y	$-27,689.80$	9	55,397.59	55,449.40
ISI Total Score	4-equi splines	Y	$-27,247.01$	10	54,514.02	54,571.58
ISI Total Score	5-equi splines	Y	$-26,901.34$	11	53,824.68	53,888.00
PCL-5 Total Score	Linear	Y	$-36,940.68$	6	73,893.36	73,927.90
PCL-5 Total Score	Beta	Y	$-32,054.86$	8	64,125.73	64,171.78
PCL-5 Total Score	2-quant splines	Y	$-35,082.11$	8	70,180.23	70,226.28
PCL-5 Total Score	3-quant splines	Y	$-33,262.49$	9	66,542.98	66,594.79
PCL-5 Total Score	4-quant splines	Y	$-32,775.91$	10	65,571.82	65,629.38
PCL-5 Total Score	**5-quant splines**	**Y**	$-32,322.95$	**11**	64,667.90	64,731.23
PCL-5 Total Score	3-equi splines	Y	$-34,231.37$	9	68,480.75	68,532.56
PCL-5 Total Score	4-equi splines	Y	$-33,912.10$	10	67,844.20	67,901.77
PCL-5 Total Score	5-equi splines	Y	$-33,717.78$	11	67,457.56	67,520.88
PROMIS Pain Intensity	Linear	Y	$-21,746.04$	6	43,504.08	43,538.62
PROMIS Pain Intensity	2-quant splines	Y	$-20,773.88$	8	41,563.75	41,609.80
PROMIS Pain Intensity	**3-quant splines**	**Y**	$-19,442.47$	**9**	38,902.93	38,954.74

Continued on next page

Table 2 – continued from previous page

Outcome Variable	Link Function	Conv.	Log-likelihood	Param.	AIC	BIC
PROMIS Pain Intensity	3-equi splines	Y	$-20,162.13$	9	40,342.27	40,394.08
PROMIS Pain Intensity	4-equi splines	Y	$-19,746.32$	10	39,512.64	39,570.21
PROMIS Pain Intensity	5-equi splines	Y	$-19,360.77$	11	38,743.53	38,806.86
QOLIBRI Overall Scale	2-quant splines	Y	$-10,972.88$	8	21,961.76	22,007.81
QOLIBRI Overall Scale	3-quant splines	Y	$-10,866.22$	9	21,750.44	21,802.25
QOLIBRI Overall Scale	4-quant splines	Y	$-10,544.12$	10	21,108.24	21,165.80
QOLIBRI Overall Scale	**5-quant splines**	**Y**	$-10,397.08$	**11**	20,816.15	20,879.48
QOLIBRI Overall Scale	3-equi splines	Y	$-10,877.31$	9	21,772.61	21,824.42
QOLIBRI Overall Scale	4-equi splines	Y	$-10,501.70$	10	21,023.40	21,080.97
QOLIBRI Overall Scale	5-equi splines	Y	$-10,344.77$	11	20,711.54	20,774.86
RAVLT Sum	Linear	Y	$-26,311.89$	6	52,635.78	52,670.31
RAVLT Sum	Beta	Y	$-25,633.73$	8	51,283.46	51,329.51
RAVLT Sum	2-quant splines	Y	$-25,908.32$	8	51,832.63	51,878.69
RAVLT Sum	3-quant splines	Y	$-25,892.85$	9	51,803.69	51,855.50
RAVLT Sum	**4-quant splines**	**Y**	$-25,764.64$	**10**	51,549.29	51,606.85
RAVLT Sum	5-quant splines	Y	$-25,720.02$	11	51,462.05	51,525.37
RAVLT Sum	3-equi splines	Y	$-25,830.97$	9	51,679.94	51,731.75
RAVLT Sum	4-equi splines	Y	$-25,615.51$	10	51,251.03	51,308.60
RAVLT Sum	5-equi splines	Y	$-25,520.06$	11	51,062.12	51,125.44
RPQ-13 Score	Linear	Y	$-33,686.26$	6	67,384.52	67,419.06
RPQ-13 Score	2-quant splines	Y	$-31,950.12$	8	63,916.24	63,962.30
RPQ-13 Score	3-quant splines	Y	$-29,106.89$	9	58,231.79	58,283.60

Continued on next page

Table 2 – continued from previous page

Outcome Variable	**Link Function**	**Conv.**	**Log-likelihood**	**Param.**	**AIC**	**BIC**
RPQ-13 Score	**4-quant splines**	**Y**	$-27,373.13$	**10**	$54,766.25$	$54,823.82$
RPQ-13 Score	3-equi splines	Y	$-30,956.17$	9	$61,930.35$	$61,982.16$
RPQ-13 Score	4-equi splines	Y	$-30,435.69$	10	$60,891.38$	$60,948.95$
RPQ-13 Score	5-equi splines	Y	$-30,074.51$	11	$60,171.01$	$60,234.33$

Appendix C

Table 3: Subject's Characteristics by Cluster

Characteristic	Total (n=2337)	Cluster 1 (n=1082)	Cluster 2 (n=505)	Cluster 3 (n=456)	Cluster 4 (n=122)	Cluster 5 (n=142)	Cluster 6 (n=30)
Demographics							
Age (Mean ± SD)	41.36 ± 17.41	38.83 ± 17.18	43.03 ± 16.82	39.61 ± 15.02	43.66 ± 18.9	55.35 ± 18.08	56.69 ± 18.45
Sex - Male	1585 (67.8%)	775 (71.6%)	320 (63.4%)	271 (59.4%)	86 (70.5%)	110 (77.5%)	23 (76.7%)
Sex - Female	752 (32.2%)	307 (28.4%)	185 (36.6%)	185 (40.6%)	36 (29.5%)	32 (22.5%)	7 (23.3%)
Race/Ethnicity							
Race - White	1798 (76.9%)	856 (79.1%)	373 (73.9%)	333 (73%)	95 (77.9%)	112 (78.9%)	29 (96.7%)
Race - Black	377 (16.1%)	135 (12.5%)	101 (20%)	109 (23.9%)	20 (16.4%)	11 (7.7%)	1 (3.3%)
Race - Mixed Race	35 (1.5%)	19 (1.8%)	10 (2%)	6 (1.3%)	0 (0%)	0 (0%)	0 (0%)
Race - Asian	85 (3.6%)	56 (5.2%)	11 (2.2%)	5 (1.1%)	5 (4.1%)	8 (5.6%)	0 (0%)
Race - Indian	8 (0.3%)	5 (0.5%)	3 (0.6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Race - Unknown	4 (0.2%)	1 (0.1%)	2 (0.4%)	0 (0%)	0 (0%)	1 (0.7%)	0 (0%)
Ethnicity - Hispanic or Latino	445 (19%)	205 (18.9%)	104 (20.6%)	93 (20.4%)	25 (20.5%)	14 (9.9%)	4 (13.3%)
Ethnicity - Non Hispanic or Latino	1870 (80%)	873 (80.7%)	397 (78.6%)	362 (79.4%)	96 (78.7%)	116 (81.7%)	26 (86.7%)
Ethnicity - Unknown	10 (0.4%)	3 (0.3%)	3 (0.6%)	1 (0.2%)	0 (0%)	3 (2.1%)	0 (0%)
Education (Mean ± SD)	13.52 ± 2.81	14.09 ± 2.85	13.3 ± 2.91	12.78 ± 2.37	12.72 ± 2.55	12.68 ± 2.54	13.5 ± 2.78
Psychiatric History	533 (22.8%)	197 (18.2%)	89 (17.6%)	181 (39.7%)	26 (21.3%)	31 (21.8%)	9 (30%)
Injury History							
TBI History - No	1701 (72.8%)	821 (75.9%)	393 (77.8%)	302 (66.2%)	92 (75.4%)	71 (50%)	22 (73.3%)
TBI History - Yes ED Visit	266 (11.4%)	133 (12.3%)	39 (7.7%)	76 (16.7%)	8 (6.6%)	8 (5.6%)	2 (6.7%)
TBI History - Yes Hospital Admit	170 (7.3%)	58 (5.4%)	36 (7.1%)	56 (12.3%)	5 (4.1%)	13 (9.2%)	2 (6.7%)
Cause of Injury							
Act of mass violence	1 (0%)	0 (0%)	0 (0%)	1 (0.2%)	0 (0%)	0 (0%)	0 (0%)
Incidental fall	645 (27.6%)	300 (27.7%)	153 (30.3%)	99 (21.7%)	24 (19.7%)	52 (36.6%)	17 (56.7%)
Other	104 (4.5%)	57 (5.3%)	15 (3%)	21 (4.6%)	5 (4.1%)	6 (4.2%)	0 (0%)
Other non-intentional injury	146 (6.2%)	83 (7.7%)	31 (6.1%)	24 (5.3%)	4 (3.3%)	2 (1.4%)	2 (6.7%)
Road traffic incident	1285 (55%)	586 (54.2%)	271 (53.7%)	265 (58.1%)	83 (68%)	70 (49.3%)	10 (33.3%)
Suicide attempt	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Violence/assault	136 (5.8%)	46 (4.3%)	31 (6.1%)	45 (9.9%)	6 (4.9%)	7 (4.9%)	1 (3.3%)
Clinical Care							
Highest Level of Care - ED Discharge	519 (22.2%)	321 (29.7%)	99 (19.6%)	93 (20.4%)	0 (0%)	1 (0.7%)	5 (16.7%)
Highest Level of Care - Hospital admit no ICU	841 (36%)	433 (40%)	193 (38.2%)	188 (41.2%)	11 (9%)	4 (2.8%)	12 (40%)
Highest Level of Care - Hospital admit with ICU	977 (41.8%)	328 (30.3%)	213 (42.2%)	175 (38.4%)	111 (91%)	137 (96.5%)	13 (43.3%)
Clinical Features							
Loss of Consciousness - No	455 (19.5%)	282 (26.1%)	93 (18.4%)	60 (13.2%)	7 (5.7%)	7 (4.9%)	6 (20%)
Loss of Consciousness - Suspected	107 (4.6%)	38 (3.5%)	27 (5.3%)	29 (6.4%)	6 (4.9%)	5 (3.5%)	2 (6.7%)
Loss of Consciousness - Unknown	99 (4.2%)	32 (3%)	25 (5%)	28 (6.1%)	6 (4.9%)	8 (5.6%)	0 (0%)
Loss of Consciousness - Yes	1659 (71%)	726 (67.1%)	357 (70.7%)	339 (74.3%)	103 (84.4%)	112 (78.9%)	22 (73.3%)
Post-traumatic Amnesia - No	502 (21.5%)	306 (28.3%)	108 (21.4%)	63 (13.8%)	9 (7.4%)	13 (9.2%)	3 (10%)
Post-traumatic Amnesia - Suspected	155 (6.6%)	47 (4.3%)	31 (6.1%)	31 (6.8%)	25 (20.5%)	18 (12.7%)	3 (10%)
Post-traumatic Amnesia - Unknown	310 (13.3%)	91 (8.4%)	58 (11.5%)	57 (12.5%)	37 (30.3%)	63 (44.4%)	4 (13.3%)
Post-traumatic Amnesia - Yes	1353 (57.9%)	634 (58.6%)	305 (60.4%)	305 (66.9%)	51 (41.8%)	38 (26.8%)	20 (66.7%)
TBI Severity							
TBI Severity - Mild	1803 (77.2%)	940 (86.9%)	408 (80.8%)	379 (83.1%)	25 (20.5%)	28 (19.7%)	23 (76.7%)
TBI Severity - Moderate	85 (3.6%)	27 (2.5%)	25 (5%)	12 (2.6%)	13 (10.7%)	8 (5.6%)	0 (0%)
TBI Severity - Severe	200 (8.6%)	38 (3.5%)	30 (5.9%)	28 (6.1%)	43 (35.2%)	58 (40.8%)	3 (10%)
GCS at ED Admission (Mean ± SD)	13.6 ± 3.15	14.3 ± 1.98	13.84 ± 2.69	13.95 ± 2.67	8.44 ± 4.76	7.5 ± 4.86	13.54 ± 3.64
GCS Post-Stabilization (Mean ± SD)	13.06 ± 3.76	13.67 ± 2.54	14.1 ± 2.85	14.45 ± 1.51	6.5 ± 4.95	8 ± 6.4	15 ± NA

Appendix D

Table 4: Demographic Comparison Between Dropout and Retained Groups

Characteristic	**Dropout**	**Retained**
Age (Mean ± SD)	48.02 ± 19.12 (n=469)	38.35 ± 17.18
Sex		
Male (%)	349 (73.8%)	1900 (67.3%)
Female (%)	124 (26.2%)	923 (32.7%)
Race		
White (%)	376 (79.5%)	2138 (75.7%)
Black (%)	50 (10.6%)	470 (16.6%)
Mixed Race (%)	1 (0.2%)	64 (2.3%)
Asian (%)	20 (4.2%)	101 (3.6%)
Indian (%)	2 (0.4%)	10 (0.4%)
Race Unknown (%)	1 (0.2%)	5 (0.2%)
Ethnicity		
Hispanic or Latino (%)	76 (16.1%)	605 (21.4%)
Non Hispanic (%)	372 (78.6%)	2196 (77.8%)
Unknown (%)	5 (1.1%)	8 (0.3%)
Education		
Years (Mean ± SD)	12.8 ± 3.02 (n=392)	13.16 ± 3.44 (n=2677)
Psychiatric History		
History Present (%)	101 (21.4%)	612 (21.7%)
TBI History		
No (%)	292 (61.7%)	1867 (66.1%)
Yes - ED Visit (%)	36 (7.6%)	279 (9.9%)
Yes - Hospital Admit (%)	39 (8.2%)	169 (6%)
Cause of Injury		
Act of mass violence (%)	0 (0%)	1 (0%)
Incidental fall (%)	156 (33%)	709 (25.1%)
Other (%)	18 (3.8%)	131 (4.6%)
non-intentional injury (%)	23 (4.9%)	161 (5.7%)
Road traffic incident (%)	243 (51.4%)	1352 (47.9%)
Suicide attempt (%)	0 (0%)	2 (0.1%)
Violence/assault (%)	23 (4.9%)	150 (5.3%)
Highest Level of Care		
ED Discharge (%)	54 (11.4%)	588 (20.8%)

Characteristic	Dropout	Retained
Hospital admit no ICU (%)	140 (29.6%)	965 (34.2%)
Hospital admit with ICU (%)	279 (59%)	970 (34.4%)
Loss of Consciousness		
No (%)	97 (20.5%)	538 (19.1%)
Suspected (%)	16 (3.4%)	127 (4.5%)
Unknown (%)	22 (4.7%)	106 (3.8%)
Yes (%)	323 (68.3%)	1743 (61.7%)
PTA		
No (%)	101 (21.4%)	561 (19.9%)
Suspected (%)	45 (9.5%)	146 (5.2%)
Unknown (%)	107 (22.6%)	305 (10.8%)
Yes (%)	205 (43.3%)	1502 (53.2%)
GCS Scores		
ED Admission (Mean ± SD)	12.45 ± 4.33 (n=364)	13.92 ± 2.66 (n=2134)
Post-Stabilization (Mean ± SD)	11.08 ± 5.47 (n=12)	13.59 ± 3.12 (n=49)

Appendix E

Table 5: Fit Metrics for Fitted Models

Model	Classes	BIC	AIC	logLik	Converged	ClassProportions
BSI18AnxScoreRaw	1	32197.9506959737	32146.1410859629	-16064.0705429814	T	1: 100%
BSI18AnxScoreRaw	2	31822.4006490664	31753.3211690519	-15864.660584526	T	1: 29.9%; 2: 70.1%
BSI18AnxScoreRaw	3	31845.6710505347	31759.3217005166	-15864.6608502583	T	1: 33.8%; 2: NA; 3: 66.2%
BSI18AnxScoreRaw	4	31785.380377067	31681.7611570453	-15822.8805785226	T	1: 32.1%; 2: 2.9%; 3: 65%; 4: NA
BSI18AnxScoreRaw	5	31808.6502428626	31687.7611528373	-15822.8805764186	T	1: 33.8%; 2: 3.6%; 3: NA; 4: 62.6%; 5: NA
BSI18AnxScoreRaw	6	31831.9201128962	31693.7611528673	-15822.8805764337	T	1: 35.3%; 2: 4%; 3: NA; 4: 60.7%; 5: NA; 6: NA
BSI18AnxScoreRaw	7	31855.1900038051	31699.7611737726	-15822.8805868863	T	1: 36.9%; 2: 4.8%; 3: NA; 4: NA; 5: 58.3%; 6: NA; 7: NA
BSI18AnxScoreRaw	8	31878.3101831264	31705.6114830902	-15822.8057415451	T	1: 36.8%; 2: 4.7%; 3: NA; 4: NA; 5: NA; 6: NA; 7: 58.5%; 8: NA
BSI18AnxScoreRaw	9	31658.2574922897	31468.28892225	-15701.144461125	T	1: 33.8%; 2: 9.8%; 3: 8.2%; 4: NA; 5: 48.2%; 6: NA; 7: NA; 8: NA; 9: NA
BSI18AnxScoreRaw	10	31925.37211179	31718.1336717466	-15823.0668358733	T	1: 33.7%; 2: 3.6%; 3: NA; 4: 62.7%; 5: NA; 6: NA; 7: NA; 8: NA; 9: NA; 10: NA

Model	Classes	BIC	AIC	logLik	Converged	ClassProportions
BSI18DeprScoreRaw	1	32016.646198478	31964.8365884671	-15973.4182942336	T	1: 100%
BSI18DeprScoreRaw	2	31551.2354333653	31482.1559533509	-15729.0779766754	T	1: 33%; 2: 67%
BSI18DeprScoreRaw	3	31574.5044830948	31488.1551330767	-15729.0775665384	T	1: 36.1%; 2: 3.9%; 3: 60%
BSI18DeprScoreRaw	4	31484.2237016767	31380.604481655	-15672.3022408275	T	1: 33.5%; 2: 8.8%; 3: 57.7%; 4: NA
BSI18DeprScoreRaw	5	31507.4934791474	31386.6043891221	-15672.3021945611	T	1: 8.8%; 2: 33.5%; 3: 57.7%; 4: NA; 5: NA
BSI18DeprScoreRaw	6	31530.7633438675	31392.6043838385	-15672.3021919193	T	1: 10.8%; 2: 35%; 3: NA; 4: 54.3%; 5: NA; 6: NA
BSI18DeprScoreRaw	7	31554.0332137375	31398.6043837049	-15672.3021918525	T	1: 35.7%; 2: 11.4%; 3: NA; 4: NA; 5: 52.9%; 6: NA; 7: NA
BSI18DeprScoreRaw	8	31577.3030837395	31404.6043837033	-15672.3021918517	T	1: 36.2%; 2: 11.7%; 3: NA; 4: NA; 5: 52%; 6: NA; 7: NA; 8: NA
BSI18DeprScoreRaw	9	31600.7158159484	31410.7472459086	-15672.3736229543	T	1: 12.5%; 2: 36.3%; 3: NA; 4: NA; 5: 51.2%; 6: NA; 7: NA; 8: NA; 9: NA
BSI18DeprScoreRaw	10	31428.1426562875	31220.9042162441	-15574.452108122	T	1: 10.4%; 2: 12.6%; 3: NA; 4: NA; 5: NA; 6: 46.7%; 7: NA; 8: 30.3%; 9: NA; 10: NA
DRS	1	28573.5045901972	28510.1817335172	-14244.0908667586	T	1: 100%
DRS	2	27554.967524496	27474.3747978124	-13723.1873989062	T	1: 13.9%; 2: 86.1%
DRS	3	27465.4436100346	27367.5810133475	-13666.7905066737	T	1: 13.3%; 2: 18.1%; 3: 68.6%
DRS	4	27488.7130669786	27373.5806002879	-13666.7903001439	T	1: 13.4%; 2: 22.6%; 3: NA; 4: 63.9%
DRS	5	27511.9833534869	27379.5810167925	-13666.7905083963	T	1: 13.5%; 2: 22.7%; 3: 63.8%; 4: NA; 5: NA
DRS	6	27276.6388344883	27126.9666277903	-13537.4833138951	T	1: 6.8%; 2: 12.9%; 3: 23.6%; 4: NA; 5: 56.7%; 6: NA
DRS	7	27558.5228624708	27391.5807857692	-13666.7903928846	T	1: 13.7%; 2: 27.1%; 3: NA; 4: NA; 5: NA; 6: 59.2%; 7: NA
DRS	8	27581.7923409983	27397.5803942931	-13666.7901971465	T	1: 13.7%; 2: 26.7%; 3: NA; 4: NA; 5: 59.6%; 6: NA; 7: NA; 8: NA
DRS	9	27604.1947032106	27402.7128865017	-13666.3564432509	T	1: 13.9%; 2: 23.7%; 3: NA; 4: NA; 5: NA; 6: 62.4%; 7: NA; 8: NA; 9: NA
DRS	10	27591.1656608398	27372.4139741274	-13648.2069870637	T	1: 12.5%; 2: 1%; 3: 32.7%; 4: NA; 5: 53.8%; 6: NA; 7: NA; 8: NA; 9: NA; 10: NA
GOSEScore	1	27199.2235876513	27141.6573543059	-13560.828677153	T	1: 100%
GOSEScore	2	27222.4685911338	27147.6324877848	-13560.8162438924	T	1: 48.9%; 2: 51.1%
GOSEScore	3	26096.8648747294	26004.7589013768	-12986.3794506884	T	1: 21.4%; 2: 70.6%; 3: 8.1%
GOSEScore	4	26120.1353209544	26010.7594775981	-12986.3797387991	T	1: 22.3%; 2: 69.6%; 3: NA; 4: 8.1%
GOSEScore	5	25812.7082736441	25686.0625602842	-12821.0312801421	T	1: 38.6%; 2: 48.2%; 3: NA; 4: 6.3%; 5: 6.9%
GOSEScore	6	25474.0943721219	25330.1787887585	-12640.0893943792	T	1: 33.3%; 2: 51.1%; 3: 2.1%; 4: 5.6%; 5: 1.4%; 6: 6.6%
GOSEScore	7	25497.3618400625	25336.1763866954	-12640.0881933477	T	1: 35.6%; 2: 48.7%; 3: NA; 4: 2.2%; 5: 5.6%; 6: 1.4%; 7: 6.6%
GOSEScore	8	25520.6313141378	25342.1759907671	-12640.0879953836	T	1: NA; 2: NA; 3: 48.4%; 4: 2.4%; 5: 35.7%; 6: 5.6%; 7: 1.4%; 8: 6.6%
GOSEScore	9	25534.2966816976	25338.5714883233	-12635.2857441617	T	1: NA; 2: 46.2%; 3: NA; 4: 3%; 5: 30.9%; 6: 6.5%; 7: 5.4%; 8: 1.4%; 9: 6.6%
GOSEScore	10	25547.4733551869	25334.478291809	-12630.2391459045	T	1: NA; 2: 45.2%; 3: NA; 4: 3.3%; 5: 34.6%; 6: 0.9%; 7: 2.7%; 8: 5.3%; 9: 1.4%; 10: 6.6%
ISITotalScore	1	50872.875565964	50809.5527092841	-25393.776354642	T	1: 100%
ISITotalScore	2	50837.9400485086	50757.3473218251	-25364.6736609126	T	1: 25.1%; 2: 74.9%
ISITotalScore	3	50765.2891859321	50667.4265892449	-25316.7132946225	T	1: 28.7%; 2: 3.5%; 3: 67.8%
ISITotalScore	4	50737.1705459568	50622.038079266	-25291.019039633	T	1: 12.2%; 2: 3.5%; 3: 53.1%; 4: 31.2%
ISITotalScore	5	50760.4404175028	50628.0380808084	-25291.0190404042	T	1: 12.2%; 2: 3.9%; 3: 56.4%; 4: 27.6%; 5: NA
ISITotalScore	6	50783.7102848237	50634.0380781257	-25291.0190390629	T	1: 12.2%; 2: 4.5%; 3: 59.5%; 4: 23.8%; 5: NA; 6: NA
ISITotalScore	7	50730.3201916305	50563.3781149289	-25252.6890574645	T	1: 24.2%; 2: 6.2%; 3: 2.2%; 4: 42.1%; 5: 8.9%; 6: 16.4%; 7: NA
ISITotalScore	8	50753.5922933401	50569.3803466348	-25252.6901733174	T	1: 6.2%; 2: 2.1%; 3: 24.3%; 4: 42.5%; 5: 9.5%; 6: NA; 7: NA; 8: 15.4%
ISITotalScore	9	50755.7323806418	50554.250563933	-25242.1252819665	T	1: 19.4%; 2: 5.3%; 3: 2.3%; 4: 34.5%; 5: 14.9%; 6: 11.1%; 7: NA; 8: NA; 9: 12.5%

Model	Classes	BIC	AIC	logLik	Converged	ClassProportions
ISITotalScore	10	50792.5908291137	50573.8391424013	-25248.9195712006	T	1: 6.2%; 2: 0.6%; 3: 24.5%; 4: 41.5%; 5: 14%; 6: NA; 7: 2.1%; 8: NA; 9: 11%; 10: NA
PCL5TotalScore	1	64731.226433793	64667.903577113	-32322.9517885565	T	1: 100%
PCL5TotalScore	2	64707.8774204536	64627.28469377	-32299.642346885	T	1: 36.3%; 2: 63.7%
PCL5TotalScore	3	64663.3123086938	64565.4497120066	-32265.7248560033	T	1: 34.4%; 2: 1.8%; 3: 63.8%
PCL5TotalScore	4	64686.5817367897	64571.4492700989	-32265.7246350495	T	1: 39.6%; 2: 2.1%; 3: 58.3%; 4: NA
PCL5TotalScore	5	64661.0190483801	64528.6167116857	-32241.3083558429	T	1: 1.9%; 2: 16.1%; 3: 57.5%; 4: 24.5%; 5: NA
PCL5TotalScore	6	64684.2888495148	64534.6166428168	-32241.3083214084	T	1: 59.8%; 2: 16.1%; 3: 2.1%; 4: NA; 5: NA; 6: 22%
PCL5TotalScore	7	64684.4681870773	64517.5261103757	-32229.7630551879	T	1: 58%; 2: 16%; 3: 2.4%; 4: 0.9%; 5: NA; 6: 22.7%; 7: NA
PCL5TotalScore	8	64685.2224226947	64501.0104759894	-32218.5052379947	T	1: 28%; 2: 5.3%; 3: 1.7%; 4: 45.8%; 5: NA; 6: NA; 7: NA; 8: 19.3%
PCL5TotalScore	9	64695.4450645108	64493.9632478019	-32211.981623901	T	1: 11.8%; 2: 2.1%; 3: 1.5%; 4: 25.7%; 5: 40.2%; 6: 18.7%; 7: NA; 8: NA; 9: NA
PCL5TotalScore	10	64718.7149345142	64499.9632478017	-32211.9816239009	T	1: 11.8%; 2: 2.1%; 3: 1.5%; 4: 40.6%; 5: 25.7%; 6: 18.3%; 7: NA; 8: NA; 9: NA; 10: NA
PROMISIntensityRawScore	1	38954.7406192122	38902.9310092014	-19442.4655046007	T	1: 100%
PROMISIntensityRawScore	2	38972.4337104721	38903.3542304577	-19439.6771152288	T	1: 0.3%; 2: 99.7%
PROMISIntensityRawScore	3	38747.9074722882	38661.5581222701	-19315.779061135	T	1: 29.9%; 2: 19.7%; 3: 50.4%
PROMISIntensityRawScore	4	38771.2390530492	38667.6198330275	-19315.8099165137	T	1: 32.6%; 2: 27.5%; 3: 0.9%; 4: 39%
PROMISIntensityRawScore	5	38794.4480713201	38673.5589812948	-19315.7794906474	T	1: 35.3%; 2: 15.1%; 3: NA; 4: NA; 5: 49.5%
PROMISIntensityRawScore	6	38817.7170586684	38679.5580986395	-19315.7790493197	T	1: 36.6%; 2: NA; 3: 21%; 4: NA; 5: 42.4%; 6: NA
PROMISIntensityRawScore	7	38840.9882849588	38685.5594549262	-19315.7797274631	T	1: 38.6%; 2: 17.2%; 3: 4.1%; 4: NA; 5: 40.1%; 6: NA; 7: NA
PROMISIntensityRawScore	8	38864.2568492471	38691.558149211	-19315.7790746055	T	1: 39%; 2: 27%; 3: NA; 4: NA; 5: NA; 6: NA; 7: NA; 8: 34.1%
PROMISIntensityRawScore	9	38887.526668647	38697.5580986072	-19315.7790493036	T	1: 36.7%; 2: NA; 3: 31.2%; 4: 32.2%; 5: NA; 6: NA; 7: NA; 8: NA; 9: NA
PROMISIntensityRawScore	10	38910.7965386506	38703.5580986072	-19315.7790493036	T	1: 39.1%; 2: 29.1%; 3: NA; 4: NA; 5: 31.8%; 6: NA; 7: NA; 8: NA; 9: NA; 10: NA
QoLIBRIOSTotalScore	1	20879.4778100228	20816.1549533429	-10397.0774766714	T	1: 100%
QoLIBRIOSTotalScore	2	20580.8862642336	20500.29353755	-10236.146768775	T	1: 92.1%; 2: 7.9%
QoLIBRIOSTotalScore	3	20468.7545346071	20370.89193792	-10168.44596896	T	1: 90.5%; 2: 1.6%; 3: 8%
QoLIBRIOSTotalScore	4	20391.0519235885	20275.9194568978	-10117.9597284489	T	1: 1.7%; 2: 45.9%; 3: 44.9%; 4: 7.6%
QoLIBRIOSTotalScore	5	20351.7812831736	20219.3789464792	-10086.6894732396	T	1: 1.7%; 2: 45.7%; 3: 43.7%; 4: 1.6%; 5: 7.4%
QoLIBRIOSTotalScore	6	20325.3977320139	20175.7255253159	-10061.862762658	T	1: 1.6%; 2: 26%; 3: 52.1%; 4: 12%; 5: 1.4%; 6: 6.9%
QoLIBRIOSTotalScore	7	20297.1547955515	20130.2127188499	-10036.1063594249	T	1: 39.1%; 2: 19%; 3: 27.9%; 4: 5%; 5: 0.9%; 6: 1.5%; 7: 6.6%
QoLIBRIOSTotalScore	8	20320.424665893	20136.2127191877	-10036.1063595939	T	1: 1.5%; 2: 19%; 3: 39.1%; 4: 27.9%; 5: 0.9%; 6: 5%; 7: NA; 8: 6.5%
QoLIBRIOSTotalScore	9	20306.5316330108	20105.0498163019	-10017.5249081509	T	1: 39.2%; 2: 19.2%; 3: 26.8%; 4: 1%; 5: 1.5%; 6: 1.2%; 7: 4.6%; 8: NA; 9: 6.5%
QoLIBRIOSTotalScore	10	20329.8015016938	20111.0498149813	-10017.5249074907	T	1: 39.2%; 2: 19.2%; 3: 26.8%; 4: 1%; 5: 1.5%; 6: 1.2%; 7: 4.6%; 8: NA; 9: NA; 10: 6.5%
RAVLTSum	1	51525.3690949119	51462.046238232	-25720.023119116	T	1: 100%
RAVLTSum	2	51093.3316348223	51012.7389081388	-25492.3694540694	T	1: 93.7%; 2: 6.3%
RAVLTSum	3	50546.0156719724	50448.1530752853	-25207.0765376426	T	1: 1.3%; 2: 92.6%; 3: 6.1%
RAVLTSum	4	50991.5028472793	50876.3703805886	-25418.1851902943	T	1: 28.2%; 2: 16.5%; 3: 49%; 4: 6.3%
RAVLTSum	5	50422.1452168174	50289.742880123	-25121.8714400615	T	1: 1.3%; 2: 15%; 3: 43.7%; 4: 33.8%; 5: 6.1%
RAVLTSum	6	50411.862364002	50262.190157304	-25105.095078652	T	1: 1.3%; 2: 14.2%; 3: 43%; 4: 34.8%; 5: 0.5%; 6: 6.1%
RAVLTSum	7	50414.1708135469	50247.2287368453	-25094.6143684226	T	1: 1.3%; 2: 9%; 3: 31.7%; 4: 32.1%; 5: 19.3%; 6: 0.5%; 7: 6.1%
RAVLTSum	8	50427.5705086635	50243.3585619582	-25089.6792809791	T	1: 1.3%; 2: 6.2%; 3: 20.3%; 4: 31.5%; 5: 26.7%; 6: 7.5%; 7: 0.5%; 8: 6.1%

Model	Classes	BIC	AIC	logLik	Converged	ClassProportions
RAVLTSum	9	50409.1962001981	50207.7143834893	-25068.8571917446	T	1: 1.3%; 2: 6.2%; 3: 20.4%; 4: 31%; 5: 7.2%; 6: 26.6%; 7: 0.5%; 8: 0.8%; 9: 6.1%
RAVLTSum	10	50383.7807504186	50165.0290637061	-25044.5145318531	T	1: 1.3%; 2: 6%; 3: 19.7%; 4: 30.9%; 5: 24.7%; 6: 10%; 7: 0.5%; 8: 0.8%; 9: 0.1%; 10: 6.1%
RPQ13Score	1	54823.8208282371	54766.2545948917	-27373.1272974459	T	1: 100%
RPQ13Score	2	54543.5258013167	54468.6896979677	-27221.3448489839	T	1: 35.5%; 2: 64.5%
RPQ13Score	3	54528.5259732058	54436.4199998532	-27202.2099999266	T	1: 36.2%; 2: 8.3%; 3: 55.5%
RPQ13Score	4	54551.7362374045	54442.3603940482	-27202.1801970241	T	1: 39%; 2: 12.6%; 3: 29%; 4: 19.4%
RPQ13Score	5	54367.0968456865	54240.4511323267	-27098.2255661633	T	1: 15.1%; 2: 6.2%; 3: 40.4%; 4: NA; 5: 38.4%
RPQ13Score	6	54390.3677794621	54246.4521960986	-27098.2260980493	T	1: 39.7%; 2: 15.1%; 3: 8%; 4: 37.2%; 5: NA; 6: NA
RPQ13Score	7	54314.1230745684	54152.9376212014	-27048.4688106007	T	1: 5.6%; 2: 10%; 3: 21.1%; 4: NA; 5: 34.3%; 6: NA; 7: 29%
RPQ13Score	8	54435.4985922458	54257.0432688751	-27097.5216344375	T	1: 15.2%; 2: 0.1%; 3: 10%; 4: NA; 5: NA; 6: NA; 7: 35.1%; 8: 39.5%
RPQ13Score	9	54460.1758556014	54264.4506622271	-27098.2253311135	T	1: 41%; 2: 15.1%; 3: 10.1%; 4: NA; 5: 33.8%; 6: NA; 7: NA; 8: NA; 9: NA
RPQ13Score	10	54367.3149447913	54154.3198814134	-27040.1599407067	T	1: 21.3%; 2: 5.8%; 3: 1%; 4: 16.6%; 5: 29%; 6: NA; 7: 26.3%; 8: NA; 9: NA; 10: NA
RPQ3Score	1	33402.4107299227	33339.0878732428	-16658.5439366214	T	1: 100%
RPQ3Score	2	32677.3403911022	32596.7476644187	-16284.3738322093	T	1: 30.2%; 2: 69.8%
RPQ3Score	3	32700.6101476611	32602.747550974	-16284.373775487	T	1: 30.9%; 2: 69.1%; 3: NA
RPQ3Score	4	32692.0871939544	32576.9547272636	-16268.4773636318	T	1: 30.8%; 2: 11.4%; 3: 57.8%; 4: NA
RPQ3Score	5	32715.3570369887	32582.9547002943	-16268.4773501472	T	1: 31.7%; 2: 14.7%; 3: 53.6%; 4: NA; 5: NA
RPQ3Score	6	32738.6269071681	32588.9547004701	-16268.477350235	T	1: 32.2%; 2: 15.2%; 3: NA; 4: 52.6%; 5: NA; 6: NA
RPQ3Score	7	32545.6916895026	32378.749612801	-16160.3748064005	T	1: 9.8%; 2: 25.1%; 3: 43.7%; 4: NA; 5: NA; 6: NA; 7: 21.4%
RPQ3Score	8	32785.1666480455	32600.9547013402	-16268.4773506701	T	1: 32.6%; 2: 17.6%; 3: NA; 4: NA; 5: 49.9%; 6: NA; 7: NA; 8: NA
RPQ3Score	9	32808.4365168933	32606.9547001844	-16268.4773500922	T	1: 32.6%; 2: 20.1%; 3: NA; 4: NA; 5: 47.2%; 6: NA; 7: NA; 8: NA; 9: NA
RPQ3Score	10	32589.8176871587	32371.0660004463	-16147.5330002231	T	1: 20%; 2: 9.8%; 3: 27%; 4: 30.2%; 5: NA; 6: NA; 7: 12.9%; 8: NA; 9: NA; 10: NA
RPQTotalScore	1	58022.0871905737	57964.5209572283	-28972.2604786141	T	1: 100%
RPQTotalScore	2	57811.0854211258	57736.2493177768	-28855.1246588884	T	1: 34%; 2: 66%
RPQTotalScore	3	57793.8009285466	57701.694955194	-28834.847477597	T	1: 35.4%; 2: 3.8%; 3: 60.8%
RPQTotalScore	4	57817.0707982404	57707.6949548841	-28834.8474774421	T	1: 38.4%; 2: 5%; 3: 56.7%; 4: NA
RPQTotalScore	5	57840.3406682416	57713.6949548817	-28834.8474774409	T	1: 39.6%; 2: 5.6%; 3: 54.7%; 4: NA; 5: NA
RPQTotalScore	6	57863.6105384518	57719.6949550883	-28834.8474775442	T	1: 42.7%; 2: 7.2%; 3: 50.1%; 4: NA; 5: NA; 6: NA
RPQTotalScore	7	57686.1587006173	57524.9732472502	-28734.4866236251	T	1: 44.5%; 2: 12.2%; 3: 3.2%; 4: NA; 5: 40.1%; 6: NA; 7: NA
RPQTotalScore	8	57709.4285706219	57530.9732472512	-28734.4866236256	T	1: 45%; 2: 12.2%; 3: 3.7%; 4: NA; 5: 39.1%; 6: NA; 7: NA; 8: NA
RPQTotalScore	9	57732.698440625	57536.9732472507	-28734.4866236254	T	1: 45.1%; 2: 12.2%; 3: 3.8%; 4: NA; 5: NA; 6: NA; 7: 38.9%; 8: NA; 9: NA
RPQTotalScore	10	57755.9683106513	57542.9732472734	-28734.4866236367	T	1: 45.2%; 2: 12.2%; 3: 3.9%; 4: NA; 5: NA; 6: 38.7%; 7: NA; 8: NA; 9: NA; 10: NA
SWLSTotalScore	1	56666.4520480826	56620.3990614063	-28302.1995307031	T	1: 100%
SWLSTotalScore	2	56222.6186688106	56159.2958121307	-28068.6479060653	T	1: 92.6%; 2: 7.4%
SWLSTotalScore	3	56332.4474086592	56251.8546819757	-28111.9273409878	T	1: 81.9%; 2: 8%; 3: 10.1%
SWLSTotalScore	4	55814.1661742655	55716.3035775783	-27841.1517887891	T	1: 1.9%; 2: 88.7%; 3: 2.5%; 4: 6.8%
SWLSTotalScore	5	55808.422976306	55693.2905096152	-27826.6452548076	T	1: 2.1%; 2: 66.9%; 3: 22%; 4: 2.4%; 5: 6.6%
SWLSTotalScore	6	55798.3431834192	55665.9408467248	-27809.9704233624	T	1: 61.4%; 2: 12.8%; 3: 15.4%; 4: 1.5%; 5: 2.2%; 6: 6.6%
SWLSTotalScore	7	55782.6460839177	55632.9738772197	-27790.4869386098	T	1: 0.5%; 2: 13.4%; 3: 60.2%; 4: 15.4%; 5: 1.7%; 6: 2.1%; 7: 6.6%

Model	Classes	BIC	AIC	logLik	Converged	ClassProportions
SWLSTotalScore	8	55844.6360630507	55677.6939863491	-27809.8469931746	T	1: 18.7%; 2: NA; 3: 49.4%; 4: NA; 5: 2.2%; 6: 1.8%; 7: 21.2%; 8: 6.6%
SWLSTotalScore	9	55740.6019166469	55556.3899699417	-27746.1949849708	T	1: 0.4%; 2: 13.4%; 3: 21.9%; 4: 52.4%; 5: 2%; 6: 1.2%; 7: 1.5%; 8: 1.2%; 9: 6%
SWLSTotalScore	10	55730.1937443255	55528.7119276167	-27729.3559638083	T	1: 8.8%; 2: 1.5%; 3: 23.6%; 4: 33.4%; 5: 17.8%; 6: 4.9%; 7: 1.7%; 8: 1.1%; 9: 1.2%; 10: 5.9%
TMTAScaledScore	1	32989.0936658921	32925.7708092122	-16451.8854046061	T	1: 100%
TMTAScaledScore	2	32844.3366528857	32763.7439262022	-16367.8719631011	T	1: 90.5%; 2: 9.5%
TMTAScaledScore	3	32809.5461713414	32711.6835746543	-16338.8417873271	T	1: 37.4%; 2: 53.8%; 3: 8.8%
TMTAScaledScore	4	32629.2915456736	32514.1590789828	-16237.0795394914	T	1: 37%; 2: 53.9%; 3: 1.5%; 4: 7.7%
TMTAScaledScore	5	32685.1960359269	32552.7936992325	-16253.3968496163	T	1: 33.7%; 2: 34.4%; 3: 3.9%; 4: 19.4%; 5: 8.6%
TMTAScaledScore	6	32511.8164797936	32362.1442730956	-16155.0721365478	T	1: 33.3%; 2: 3.8%; 3: 34.6%; 4: 1.4%; 5: 19.3%; 6: 7.7%
TMTAScaledScore	7	32484.0698770756	32317.127800374	-16129.563900187	T	1: 1.4%; 2: 4%; 3: 35.8%; 4: 31.8%; 5: 18.1%; 6: 1.2%; 7: 7.6%
TMTAScaledScore	8	32481.0406051176	32296.8286584123	-16116.4143292062	T	1: 37.4%; 2: 2.1%; 3: 11.2%; 4: 25.5%; 5: 13.6%; 6: 1.2%; 7: 1.4%; 8: 7.7%
TMTAScaledScore	9	32531.3980814455	32329.9162647367	-16129.9581323683	T	1: 1.4%; 2: 37.7%; 3: 18%; 4: 30%; 5: 4.1%; 6: 1.2%; 7: NA; 8: 7.6%; 9: NA
TMTAScaledScore	10	32527.5825151752	32308.8308284627	-16116.4154142314	T	1: 1.4%; 2: 2.1%; 3: 11.2%; 4: 37.4%; 5: 25.5%; 6: 13.6%; 7: 1.2%; 8: NA; 9: NA; 10: 7.7%
TMTBScaledScore	1	32810.3031126106	32746.9802559307	-16362.4901279653	T	1: 100%
TMTBScaledScore	2	32644.7311817972	32564.1384551137	-16268.0692275568	T	1: 90.5%; 2: 9.5%
TMTBScaledScore	3	32443.5272421254	32345.6646454383	-16155.8323227191	T	1: 1.8%; 2: 89.9%; 3: 8.3%
TMTBScaledScore	4	32465.575466997	32350.4430003062	-16155.2215001531	T	1: 45.7%; 2: 45.5%; 3: 0.3%; 4: 8.4%
TMTBScaledScore	5	32325.154566901	32192.7522302066	-16073.3761151033	T	1: 24.6%; 2: 45.7%; 3: 20.1%; 4: 1.7%; 5: 7.9%
TMTBScaledScore	6	32149.6095898839	31999.9373831859	-15973.9686915929	T	1: 48.1%; 2: 13.2%; 3: 28.2%; 4: 0.4%; 5: 1.8%; 6: 8.3%
TMTBScaledScore	7	2000000224.94208	2000000058	-1e+09	F	1: NA; 2: NA; 3: NA; 4: NA; 5: NA; 6: NA; 7: NA
TMTBScaledScore	8	32023.3064597104	31839.0945130051	-15887.5472565026	T	1: 35.9%; 2: 8.9%; 3: 30.7%; 4: 14%; 5: 0.4%; 6: 1.8%; 7: 1%; 8: 7.2%
TMTBScaledScore	9	32168.0437025303	31966.5618858214	-15948.2809429107	T	1: 30.3%; 2: 5%; 3: 16.4%; 4: 26.1%; 5: 12.3%; 6: 0.4%; 7: 1.8%; 8: NA; 9: 7.7%
TMTBScaledScore	10	32044.4557924751	31825.7041057626	-15874.8520528813	T	1: 2.4%; 2: 10.1%; 3: 37.6%; 4: 25.5%; 5: 14%; 6: 0.4%; 7: 1%; 8: 1.8%; 9: 7.2%; 10: NA
WAISCodingTotalRawScore	1	55413.9823070246	55350.6594503446	-27664.3297251723	T	1: 100%
WAISCodingTotalRawScore	2	55437.2518131113	55356.6590864277	-27664.3295432139	T	1: 53.1%; 2: 46.9%
WAISCodingTotalRawScore	3	54837.2271784992	54739.364581812	-27352.682290906	T	1: 1.7%; 2: 92.1%; 3: 6.2%
WAISCodingTotalRawScore	4	54720.4367607499	54605.3042940591	-27282.6521470296	T	1: 1.7%; 2: 1%; 3: 91.1%; 4: 6.2%
WAISCodingTotalRawScore	5	54702.74297017	54570.3406334756	-27262.1703167378	T	1: 1.1%; 2: 0.9%; 3: 90.4%; 4: 6.2%; 5: 1.5%
WAISCodingTotalRawScore	6	54671.3293924031	54521.6571857051	-27234.8285928526	T	1: 1.7%; 2: 35.3%; 3: 0.1%; 4: 55.4%; 5: 1.3%; 6: 6.2%
WAISCodingTotalRawScore	7	54654.3948561343	54487.4527794327	-27214.7263897164	T	1: 1.6%; 2: 3.6%; 3: 46.3%; 4: 0.1%; 5: 40.9%; 6: 1.3%; 7: 6.1%
WAISCodingTotalRawScore	8	54580.5544826433	54396.342535938	-27166.171267969	T	1: 1.3%; 2: 1.9%; 3: 22.3%; 4: 1.2%; 5: 28.8%; 6: 0.5%; 7: 37.9%; 8: 6.1%
WAISCodingTotalRawScore	9	54572.5402267128	54371.0584100039	-27150.529205002	T	1: 1.3%; 2: 2.2%; 3: 20.9%; 4: 37.7%; 5: 0.1%; 6: 0.6%; 7: 1.4%; 8: 29.7%; 9: 6.1%
WAISCodingTotalRawScore	10	54675.7648067718	54457.0131200593	-27190.5065600297	T	1: 1.4%; 2: 2.9%; 3: 20.5%; 4: 31.5%; 5: 0.1%; 6: 30%; 7: 1%; 8: 2.1%; 9: 4.4%; 10: 6.1%
WAISSymbolTotalRawScore	1	47891.0347925114	47827.7119358314	-23902.8559679157	T	1: 100%
WAISSymbolTotalRawScore	2	47763.6095636702	47683.0168369867	-23827.5084184933	T	1: 1.4%; 2: 98.6%
WAISSymbolTotalRawScore	3	47616.7596646617	47518.8970679745	-23742.4485339873	T	1: 48.7%; 2: 43.6%; 3: 7.7%
WAISSymbolTotalRawScore	4	47598.4706464333	47483.3381797425	-23721.6690898712	T	1: 47.9%; 2: 43.2%; 3: 1.2%; 4: 7.7%

Model	Classes	BIC	AIC	logLik	Converged	ClassProportions
WAISSymbolTotalRawScore	5	47557.9676224232	47425.5652857288	-23689.7826428644	T	1: 42.4%; 2: 10.4%; 3: 1.2%; 4: 38.4%; 5: 7.7%
WAISSymbolTotalRawScore	6	47050.3050207801	46900.632814082	-23424.316407041	T	1: 36.7%; 2: 12.3%; 3: 1.6%; 4: 41.7%; 5: 1.4%; 6: 6.2%
WAISSymbolTotalRawScore	7	47015.2715391252	46848.3294624235	-23395.1647312118	T	1: 36.7%; 2: 12.5%; 3: 41.3%; 4: 1.6%; 5: 0.3%; 6: 1.4%; 7: 6.1%
WAISSymbolTotalRawScore	8	47019.3062867708	46835.0943400655	-23385.5471700328	T	1: 31.7%; 2: 8.3%; 3: 31.1%; 4: 1.4%; 5: 19.6%; 6: 0.3%; 7: 1.4%; 8: 6.1%
WAISSymbolTotalRawScore	9	47017.2593221978	46815.777505489	-23372.8887527445	T	1: 30.5%; 2: 8%; 3: 29.1%; 4: 22.8%; 5: 1.3%; 6: 0.5%; 7: 0.3%; 8: 1.4%; 9: 6.1%
WAISSymbolTotalRawScore	10	47030.6295742638	46811.8778875514	-23367.9389437757	T	1: 25.3%; 2: 4.8%; 3: 14.6%; 4: 18.7%; 5: 1.2%; 6: 27%; 7: 0.5%; 8: 0.3%; 9: 1.4%; 10: 6.1%