

- University of North Carolina at Charlotte
- M.S. in Data Science and Business Analytics
- DSBA 6400
- Jameson Ellis





## Introduction

- Objective: To utilize predictive analytics to evaluate and compare the outcomes of patients treated operatively versus non-operatively following pelvic and acetabular fractures.
- Pelvic and acetabular fractures present significant challenges in treatment, and understanding the outcomes of different treatment approaches is crucial for optimizing patient care.

### Background and Significance

- Pelvic Ring Injuries: Represent approximately 9% of all fractures from blunt trauma.
- Challenges: Associated injuries can confound results; limited large, multicenter prospective studies.
- Study Goals: Capture data on functional, clinical, and performance outcomes to optimize patient care.

# Posterior Wall Posterior Column Anterior Wall Posterior Column Transverse Hemi Transverse Both Columns

# Study Protocol

Data Collection:
Performance assessments
(TUG, 10mWT) and patientreported outcomes
(PROMIS-29, Majeed Pelvis
Score).

Sample Size: 1,000 patients.

Primary Outcomes: Return to work/activities, PROMIS-29 scores.

Secondary Outcomes: Fracture healing, pain, mobilization, complications, cost.

# Project Plan Outline

- Data Acquisition and Management
- Exploratory Data Analysis (EDA)
- Predictive Analytics
- Model Evaluation
- Reporting and Presentation



# **Data Preparation**

- Exporting/Importing Data
- Initial Cleanup
- Investigating Missing Data
- Creating Summary Tables
- Identifying Unique Patients and Features
- Checking Data Consistency

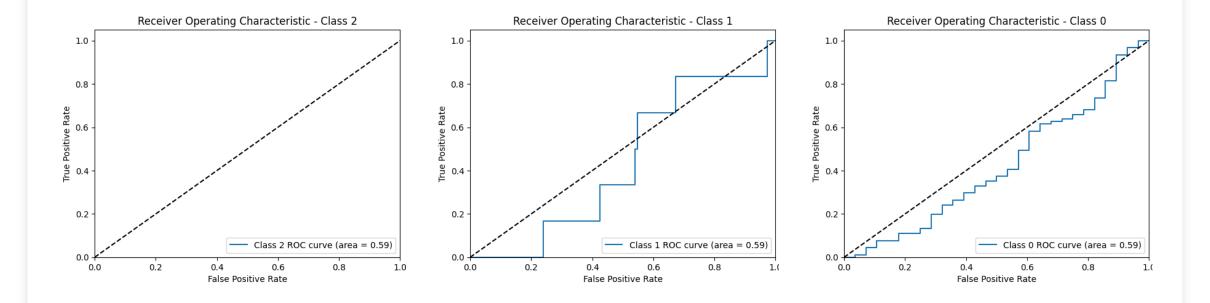
# **Predictive Analytics**

- Variables Chosen for the Model:
  - Gender
  - Age
  - Initial Definitive Treatment
  - Total Injury Severity Score (ISS)
  - Charleston Comorbidity Index (CCI)
  - Race
  - Insurer
  - Independent Living Status
  - Fracture Type
  - Total PROMIS-29 Score
  - Majeed Score
  - Complications Since Last Visit

- Model Selection: Logistic regression for binary outcomes.
- Data Splitting: 80-20 split for training and testing sets.
- Model Training: Logistic regression model built and trained using training data.

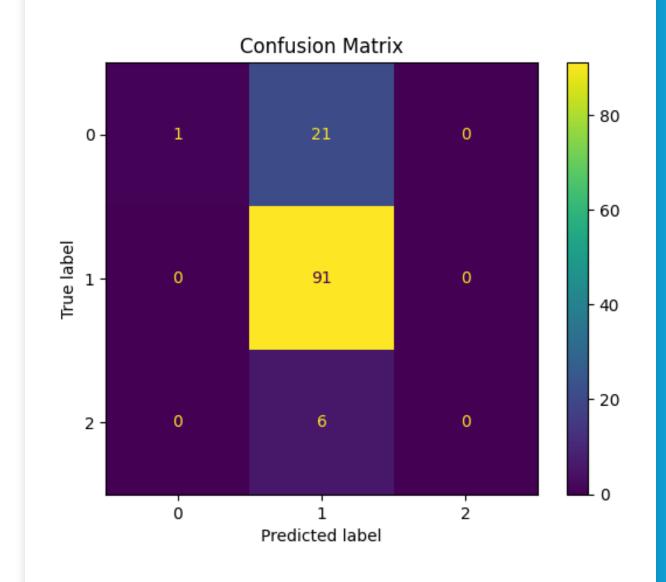
### **Model Evaluation**

- Evaluation Metrics:
  - Accuracy: 0.7731
  - Precision: 0.5904
  - Recall: 0.3485
  - F1 Score: 0.3193
  - ROC AUC Score: 0.5886



# Visualizations:

# Confusion Matrix



### **Coefficients Interpretation**

- Females and older patients are less likely to experience complications.
- Higher injury severity and comorbidity scores increase complication risks.
- Multi-race and white races, and private insurance types reduce complication likelihood.
- Independent living decreases the likelihood of complications.
- Higher PROMIS-29 and Majeed scores increase the likelihood of complications.

#### PACE Data Elements - Shell Tables

Table 1. Demographics

Demographics	Coefficient	Odds Ratio
Age: median	-0.0059	0.9941
(95% CI)		
Sex (#)	-0.0461	0.9549
Race	-0.0409	0.9599
Independent	-0.2474	0.7808
Living		
Insurer	-0.2680	0.7649

Table 2. Comorbidities

Comorbidities	Coefficient	Odds Ratio
CCI: median	0.0041	1.0041
(95% CI)		

Table 3 Injury Characteristics

. a. a		
Injury	Coefficient	Odds Ratio
Characteristics		
Fracture Type	0.2237	1.2507

Table 4. Functional Outcomes

Functional	Coefficient	Odds Ratio
Outcomes		
PROMIS 29	0.0047	1.0047
Score		
Majeed Score	0.0075	1.0076

Table 5. ISS Scores

ISS Scores	Coefficient	Odds Ratio
Total ISS	0.0343	1.0349

Table 6. Fracture Management

iable of Fracture Management		
Fracture	Coefficient	Odds Ratio
Management		
Initial Definitive	0.2161	1.2412
Treatment		

# Key Findings and Implications

- Operative vs. Non-Operative Treatments: Operative treatments have higher complication risks.
- Key Factors: Age, sex, ISS score, CCI score, race, insurer, independent living status, fracture type, PROMIS-29 score, and Majeed score influence complications.
- Clinical Implications: Tailoring treatment plans to individual profiles can reduce complications.
- Impact on Patient Care: Understanding predictors helps improve recovery and reduce adverse outcomes.

# Model Improvements

### Potential Improvements and Future Work:

- Baseline Data: Include repeat analyses at multiple intervals to capture longitudinal trends.
- Random Forest Model: Identify key predictors, handle non-linearity, and reduce overfitting.
- Additional Models: Explore other machine learning models for better accuracy and robustness.
- Evaluation Metrics: Current metrics are not great due to potential issues with the dataset.
- Future Work: Determine complication rate and account for missing cases by checking the "final status form."

# **Questions and Comments**



# Acknowledgements

- Data Team: Special thanks to my mentor, Dr. Susan Odum, and her team for their invaluable guidance and support throughout this project.
- MSKI Department: Appreciation to the MSKI department for allowing me to use their resources and facilities.
- UNC Charlotte: Heartfelt thanks to all my professors and the faculty and staff in the M.S. in Data Science and Business Analytics program at UNC Charlotte for their continuous support and encouragement.