**Kevin Zhou**

Kevin.zhou.1@stonybrook.edu

2023

Project Name: Falls Patients Drug & Treatment Regimens and Readmissions

# OVERVIEW

Falls are a significant concern in healthcare settings, especially among elderly patients and those with specific health conditions. Falls can result in severe injuries, prolonged hospital stays, increased healthcare costs, and a decline in patients' overall well-being. They are considered a leading cause of morbidity and mortality, highlighting the importance of preventive measures. By conducting a capstone project focused on identifying predictors of falls, valuable insights can be gained to enhance fall risk assessment and prevention strategies. By understanding the factors that contribute to falls, healthcare professionals can implement targeted interventions, such as improved patient monitoring, environmental modifications, and tailored patient education, to mitigate the risk of falls and minimize their impact on patient outcomes.

The primary objective of this capstone project is to investigate and analyze the relationship between medical & treatment regimens and falls readmissions. By leveraging both falls & non-fall patient datasets, this project aims to uncover significant correlations and patterns related to falls. The project seeks to utilize advanced data analysis techniques, including exploratory data analysis, statistical modeling, and predictive analytics, to develop a robust understanding of the factors contributing to falls. Ultimately, the findings of this capstone project will inform the development of targeted fall prevention strategies, interventions, and protocols, enhancing patient safety and improving the quality of care provided within the hospital.

# SMART GOALS

1. **SMART Goal 1:** By the end of the first half of Summer Session II, we aim to complete the data collection and cleaning phase of the project.
   1. **How is it (S):** Use TriNetX to obtain a dataset comprising clinical records, treatment history and lab results of individuals presenting with a primary diagnosis of a fall and individuals not presenting with a primary diagnosis of a fall.
   2. **How is it (M):** Collect data from a randomized sample of 1000 patients who are diagnosed with a fall and a randomized sample of 1000 patients who have not been diagnosed with a fall.
   3. **How is it (A)**: Utilizing TriNetX to obtain deidentified patient data that meets the following standards:
      1. Diagnosed with a fall on or after July 14th, 2022.
      2. Has a code of R29.6 – repeated falls or z91.81 – history of falling or w18 – other slipping, tripping and stumbling and falls.
2. **SMART Goal:** By the end of the second half of Summer Session II, we aim to complete data wrangling, data visualization, descriptive and exploratory data analysis and interpretation phases of the project.
   1. **How is it (S):** Data will be collected and cleaned utilizing TriNetX.
   2. **How is it (M):** We aim to identify a correlation between drug and treatment regimens and fall readmission.
   3. **How is it (A)**: We have the necessary tools and knowledge to perform the analysis and machine learning.
   4. **How is it (R):** This goal is relevant to the project because it will result in the necessary information to answer the question.
   5. **How is it (T):** We aim to complete this project by the end of Summer Session II.

# GENERAL TASKS AND SUBTASKS

* **Task 0: Question Exploration**
  + Subtask 0-1:potential questions
    - Deliverable: A document containing feasible questions to be selected
* **Task 1: Collect data from TriNetX**
  + Subtask 1-1: Inclusion/Exclusion Criteria
    - **Deliverable:** a document that states the inclusion and exclusion criteria.
* **Task 2: Potentially enrich data**
  + Subtask 2-1: Search for a suitable dataset that can be used to enrich the data collected from TriNextX
    - **Deliverable:** If dataset is found a document that explains why it is beneficial to include
* **Task 3: Evaluate/Clean Data**
  + Subtask 3-1: Clean the dataset
    - **Deliverable:** A clean CSV file where missing data is excluded or dealt with accordingly
* **Task 4: Data Analysis**
  + Subtask 4-1: Visualize the data using histograms and scatter plots
    - **Deliverable:** Visuals such as scatter plots and histograms displaying any trends or outliers in the data
  + Subtask 4-2: Perform bivariate analysis and summary statistics
    - **Deliverable:** A document summarizing the correlations between variables
  + Subtask 4-3: Variable and Model Selection
    - **Deliverable:** A document demonstrating the target, predictor variables and models that will be used.
  + Subtask 4-4: Model Training
    - **Deliverable:** A python script that entails the models and parameters utilized
  + Subtask 4-5: Model Performance Evaluation
    - **Deliverable:** A document that summarizes the model’s performance metrics of the trained models
* **Task 5: Results and Recommendations**
  + Subtask 5-1: Identify key insights and trends
    - **Deliverable:** A document summarizing the key findings that were identified during analysis
  + Subtask 5-2: Recommendations for patient care outcomes
    - **Deliverable:** Document listing recommendations involving patient care outcomes based on the insights.
  + Subtask 5-3: Summary report
    - **Deliverable:** A report summarizing the entirety of the project such as methodology, data analysis and recommendations. It will include any images and graphs that were generated.