Multi-Drug-Resistant Organism Infection Rate in Patients Undergoing Immunomodulatory Therapies for Solid Organ Malignancies.

Team Members:
- Stein Ingebretsen, u0460309, u0460309@utah.edu
- David Stone, u1025397, david.stone@hci.utah.edu
- David Villarreal, u1531841, david.villarreal@utah.edu
- Lindsey Kwon, u0895615, lindsey.k@utah.edu
- Lindsey McGuire, u1204515, lindsey.mcguire@utah.edu

GitHub Link:
https://github.com/dstone42/Huzzah

Background and Motivation

The immune response impaired by cancer and antineoplastic treatments (e.g., myelosuppressive chemotherapy) renders patients with cancer particularly susceptible to severe infections (https://pubmed.ncbi.nlm.nih.gov/34224061/). Sepsis, bacteremia, neutropenia, rank among the most critical complications in this population and contribute substantially to morbidity and mortality (https://pubmed.ncbi.nlm.nih.gov/16575919/). In addition to the immune-related toxicities associated with immunomodulatory antineoplastic treatments, patients exhibit specific immune dysfunctions and a high risk of severe infections that frequently require admission to the intensive care unit (ICU) (https://pubmed.ncbi.nlm.nih.gov/35950720/). Therefore, infections in patients with solid tumors undergoing immunotherapy are clinically relevant, and even modest advances in their management could result in substantial improvements in survival.

Infections caused by multidrug-resistant organisms (MDRO) are associated with a higher risk of mortality compared to infections caused by susceptible pathogens, possibly because initial antibiotic therapy is more frequently inappropriate in the presence of resistance (https://pubmed.ncbi.nlm.nih.gov/36355907/). In critical care settings such as the ICU, the prevalence of MDRO is higher, increasing the risk of hospital-acquired infections and prolonging the length of hospital stay among critically ill patients (https://pubmed.ncbi.nlm.nih.gov/29023567/; https://pubmed.ncbi.nlm.nih.gov/16575919/). Moreover, the increase in antimicrobial resistance correlates directly with higher rates of mortality, morbidity, and costs associated with hospitalizations, particularly in the ICU (https://pubmed.ncbi.nlm.nih.gov/28701159/). In this context, knowledge of local antimicrobial resistance epidemiology enables the early identification of patients at risk for MDRO infections and facilitates the timely initiation of effective therapy (https://pubmed.ncbi.nlm.nih.gov/31438593/).

The present study focuses on a specific group of critically ill patients, namely those with solid tumors who are receiving active immunotherapy and developing sepsis with bacteremia requiring hospitalization. Patients receiving immunomodulatory treatments for malignancy may be at higher risk due to additional immunosuppressive treatments (i.e. corticosteroids or TNF-α inhibitors) to manage adverse effects of their antineoplastic treatments (https://pubmed.ncbi.nlm.nih.gov/27501841/;

https://pubmed.ncbi.nlm.nih.gov/31912796/). Given the increasing use of immunotherapies in oncology, there is an epidemiological need to quantify the risk of MDRO infections in this emerging cohort of patients.

Overall, current knowledge regarding infections caused by multidrug-resistant pathogens in patients with solid tumors is limited (https://pubmed.ncbi.nlm.nih.gov/31438593/), and no published study to date has specifically quantified whether oncologic immunotherapy confers an additional risk of developing MDRO infections. This lack of data leaves the true magnitude of the problem uncertain and hinders the development of evidence-based clinical recommendations for the management of infections in this subgroup. Therefore, the central question motivating the present study is whether patients with solid tumors receiving active immunomodulatory therapy have an increased risk of MDRO infections during episodes of hospital-acquired sepsis and bacteremia. Addressing this question will fill an existing knowledge gap by providing epidemiological evidence on the burden of MDRO in a population that is becoming increasingly prevalent in clinical practice.

Project Objectives

Are patients admitted to the hospital with sepsis and bacteremia undergoing active treatment with immunomodulators for solid organ malignancies at higher risk of developing multi-drug-resistant infections (MDRO).

- Improving empiric coverage practice
- Decreasing mortality
- Reducing ICU time
- Reducing length of stay in hospital
- Early recognition and intervention with improved screening.

The primary research question: Among patients hospitalized with sepsis and bacteremia who are undergoing active treatment with immunomodulatory therapies for solid organ malignancies, is there an increased risk of infections caused by multidrug-resistant organisms (MDROs) compared to similar patients not receiving such therapies?

Using data from the MIMIC-IV database, our study will examine whether active treatment with immunomodulatory agents (like checkpoint inhibitors or other immunotherapies) adds extra risk for multidrug-resistant organism (MDRO) infections in patients with solid tumors who develop hospital-acquired sepsis or bacteremia.

- Identify the subset of patients with solid organ malignancies on immunomodulator therapy who had episodes of sepsis or bacteremia, identify which patients had MDROs by reviewing culture results for organism identification and resistance patterns.

- Compare the rates of MDRO infections between those who were on immunotherapy at the time versus those who weren't, while adjusting for potential confounders.

- Look at downstream clinical outcomes depending on whether the infection was an MDRO or not: in-hospital death, time spent in the ICU, total hospital length of stay, days on mechanical ventilation, and other indicators of illness severity, 30-day survival.

Practical benefits:

- Smarter antibiotic choices upfront: If we know immunotherapy might make multi-drug-resistant infections more likely, clinicians could start with more directed antibiotics in those patients.
- Fewer deaths and less time in the hospital/ICU: Spotting this risk earlier means we can act faster—better antibiotics sooner, quicker source control, less time on a ventilator or pressors, decreased ICU stay. All of that adds up to lower mortality and shorter, less complicated stays.
- Better use of screening and antibiotics overall: This kind of data could help build smarter risk-stratified approaches for patient care to appropriately treat patients while being conscientious of antibiotic stewardship.
- Bigger picture payoff: It could feed into updated guidelines, help hospitals save money on unnecessarily long stays, and set the stage for better risk-prediction tools or bigger prospective studies down the road.

Data

- The data for this project is the MIMIC-IV Database. It contains retrospectively collected, deidentified data of patients admitted to the emergency department or intensive care unit at the Beth Israel Deaconess Medical Center in Boston, MA.
- Johnson, Alistair, et al. "MIMIC-IV" (version 3.1). *PhysioNet* (2024). RRID:SCR_007345. https://doi.org/10.13026/kpb9-mt58

Data Processing

We expect moderate data cleanup. The data is structured and organized well. The data cleanup challenges will come from missing values, inconsistencies, and temporal information.

We plan to derive:

- Patients with malignancies who are taking immunotherapy
    - Within that subset, patients who have infections and receive antibiotics
    - Measures of outcome for the patient

The data processing will be implemented in a scripting programming language such as python or R.

Design

The general design will follow a stage-based approach, as to demonstrate how the data evolves throughout the wrangling process. This will initially be discovering and highlighting the data issues (missing, implausible, and irrelevant data) and ending with proof of a cleaned and ready dataset. This will be through a combination of tables, simple plots (bar plots, histograms, etc.) and schematic diagrams generated from the cleaned and intermediate datasets. Clarity and interpretability will be prioritized over visual complexity, using consistent colors and styles for final readability.

Some example ideas of visualizations include:

- A flow diagram showing cohort selection and filtering
- Bar charts and/or heatmaps highlighting missing data
- Histograms and/or box plots for timestamped data
- Tables summarizing variable definitions and dataset structure

These are justified as they can easily represent counts and proportions, as well as highlight outliers, patterns, and other relevant information as we clean the data.

An alternative approach would prioritize overall comprehension, such as a single consolidated summary in one view. This could be a more concise way to digest our process, however, would likely be less detailed.

Must-Have Features

- ICD-10/ICD-9 Diagnostic codes
- Microbiology data
    - Blood cultures
    - Urine cultures
    - Sputum cultures
    - Tissue/fluid cultures
    - Antimicrobial speciation
    - Antimicrobial Resistances
- Medications
    - Antibiotics
    - Antifungals
    - Pain control
    - Antithrombotic
    - Blood product transfusions
    - Antihyperglycemics
    - Antihypertensives
    - Antineoplastics
        - Immunomodulators
        - Others
- Physiologic parameters
    - Heart rate
    - Temperature
    - Blood pressure
    - Respiratory rate
    - Oxygen needs
- Laboratory Data
    - CBC
    - BMP
    - CMP
    - CRP
    - ESR
    - INR
- In hospital mortality

Optional Features

- 60- or 90-day mortality
- Prognosis
- Grading/staging of cancer

Project Schedule

- **Week 0:** Project Proposal draft, MIMIC-IV Access application, and credential setup.
- **Week 1: Submit Proposal (Feb 2nd).** Data reconstruction/download and initial schema exploration.
- Week 2: Write SQL scripts to get the correct subset and tables for analysis.
- **Weeks 3-4: Data Quality Assessment (DQA) - Baseline.** Use **Weiskopf's Rubric** (Completeness, Plausibility, etc.) to audit the raw data. Identify features which need addressing (e.g., missing heart rates, impossible lab values). Team meeting for GitHub branching and best practices for reproducible scripts. Assign tables/features for cleaning.
- **Week 5-6:** Begin Data Cleaning.
- **Intermediate Presentation.** Demo the "cleaned" data vs. the "raw" baseline. Incorporate peer/instructor feedback on your cleaning logic.
- **Week 6: Verification & Validation.** Reevaluate the DQA rubric on the *cleaned* data to verify improvements.
    - **Project Update (March 16th):** Submit progress report and "Before vs. After" quality metrics.
- **Week 7: Spring Break.** (Rest or optional light documentation).
- **Week 8: Advanced Verification.** Focus on **Concordance** (cross-referencing tables) and **Correctness**. Draft 1 of the technical report (Introduction and Methodology).
- **Week 9: Refining the Pipeline.** Finalize any edge-case cleaning. Team review of the technical report. Start integrating final data tables/figures.
- **Week 10: Documentation & Reporting.** Finalize all cleaning logs. Complete Draft 2 of the report, focusing on the "Data Quality Profile" (how much the data improved).
- **Week 11: Repository Cleanup.** Finalize README.md and ensure the code is "one-click" runnable. Polish the final presentation slides and visuals.
- **Week 12: Presentation (April 20th). Final Submission (April 21st):** Submit the final cleaning code, data quality report, and slides.