SIR-SIS Fusion: MRSA and Influenza Co-infection Forecasting

Lauren Fogel [1], Asma Khimani [1], Harshini Vummadi [1]
[1] School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, 30332, USA

Introduction

In healthcare settings, preventing the spread of disease is a critical concern, especially in populations of highly vulnerable patients. Healthcare Associated Infections (HAIs) refers to infections that people can develop while receiving care for another health condition [U.S. Department of Health Human Services 2024]. In hospitals, 1 in 31 patients have at least one HAI at any given time [Centers for Disease Control and Prevention 2020]. Among HAIs, Methicillin-Resistant Staphylococcus aureus (MRSA) is a common infection that is associated with significant morbidity, mortality, length of hospital stay, and cost burden [Siddiqui and Koirala 2018], [Cosgrove et al. 2003], [Cosgrove et al . 2005]. Alongside this, seasonal outbreaks, such as influenza (flu), can further complicate infection control efforts, particularly when co-infection occurs between these viral and bacterial pathogens leading to worse patient outcomes [Liu et al . 2021]. However, hospitals do not always have effective ways to predict when and how these infections will spread concurrently. Therefore, it is important to understand the conditions that lead to co-infections, so hospitals can prevent these co-infection cases and improve patient outcomes.

Problem Definition

The problem this project attempts to solve is the lack of predictive models for understanding and assessing the risk of co-infections between MRSA and flu in healthcare settings. Co-infections have been shown to complicate patient conditions, with research indicating that 20.5-24.5% of patients with flu also experienced bacterial co-infections, particularly involving MRSA [Blyth et al. 2013]. These co-infections likely lead to more severe cases requiring intensive care, yet not much is known of the transmission dynamics between these viral and bacterial pathogens in a healthcare setting. Ultimately, this project addresses the need to better predict co-infection risks by developing model to simulate the spread of both MRSA and flu within healthcare environments. By exploring how these infections interact, we aim to inform hospitals with insights they need to reduce the risk of co-infections.

Dataset

The dataset we used is Healthcare Personnel Movement Data [Jang et al. 2021], which is based on a network of sensors to track the movement of healthcare workers in a hospital setting, specifically a dialysis unit. The dataset contains 10 days of the movements of healthcare workers and has fields such as badge ID, time, location of the nursing station, hand- washing station and dialysis chairs, and time intervals of patients in and out of the dialysis chairs. We created contact networks based on proximity thresholds, spatial locations, and occupation of specific zones within the dialysis unit. Our output was 30 days of simulated data for the most intensive day (Day 10) in the dialysis unit and includes two groups of patients on different schedules along with healthcare workers. The authors' original model for COVID-19 assumes close contact within a 6-foot distance for transmission, which follows closely with established flu transmission guidelines, thus making it an ideal foundation. Further modifications were required to adapt the data for modeling MRSA transmission which has a contact distance of 1 foot and therefore has a much smaller contact network.Using this dataset, we created a 30-day synthetic dataset from the most intensive day and used that to build the network of our models.

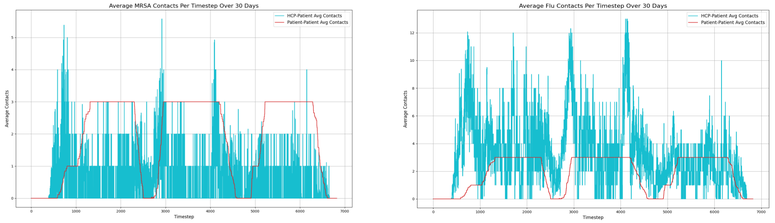


Figure 1: Average Contact Over Time Steps

Experiments and Results

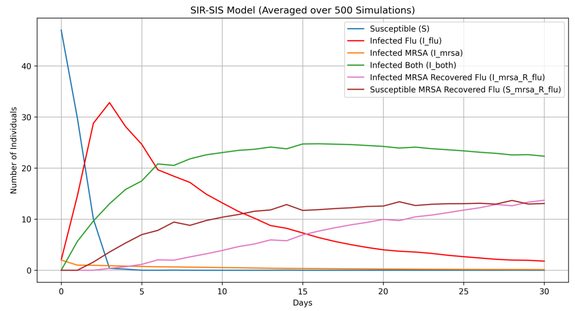


Figure 2: SIR-SIS Model

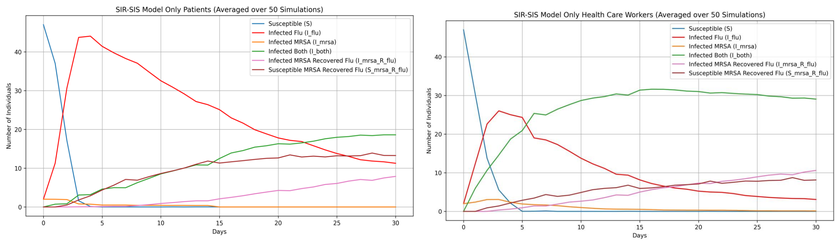


Figure 3: Patient only and Healthcare worker only SIR-SIS model

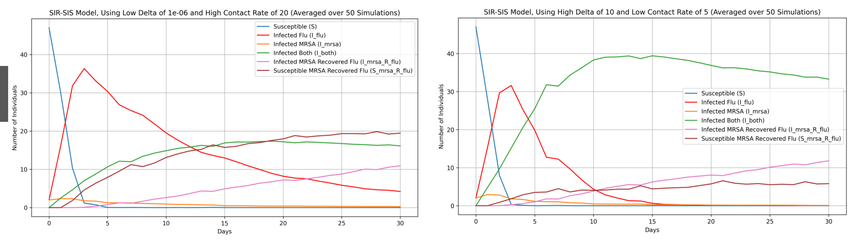


Figure 4: Low and High Co-infection SIR-SIS Model

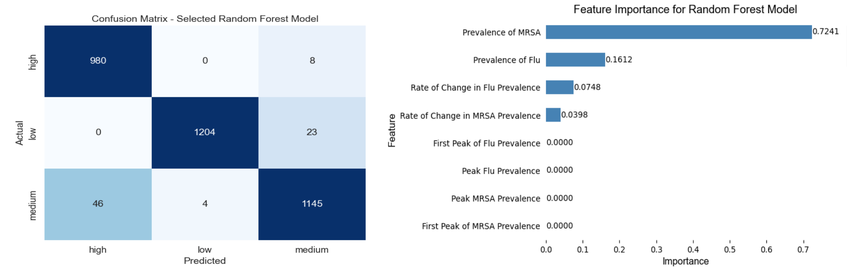


Figure 5: Confusion Matrix and Feature Importance Plot

This confusion matrix displays the minimal false negatives, 8, of the selected random forest model. This feature importance plot shows that the prevalence of MRSA and prevalence of flu contributed the most to the decisions of the model, followed by the rate of change metrics.

Scenario-Based Risk Evaluation To validate the model's predictions and provide actionable insights, we generated scenarios by varying important features, prevalence of MRSA and prevalence of flu across realistic ranges observed in the simulation data. The RF model identified 16 high-risk scenarios, primarily associated with high flu prevalence combined with moderate MRSA prevalence. These scenarios consistently showed high-risk probabilities exceeding 72%, which suggests that simultaneous surges in flu and MRSA prevalence significantly increase the likelihood of co-infection.

Methods

SIR-SIS Co-Infection Model Development

Parameters

- β_{flu}: Transmission rate of influenza
- β_{MRSA}: Transmission rate of MRSA
- γ_{flu}: Recovery rate of influenza
- γ_{MRSA}: Recovery rate of MRSA
- γ_{both}: Recovery rate from both diseases, taking in co-infection dynamics
- α: Co-infection penalty factor
- δ: Increased susceptibility to MRSA due to current infection with influenza

Differential Equations of SIR-SIS Model

$$\frac{dS}{dt} = -\beta_{\text{flu}} S I_{\text{flu}} - \beta_{\text{MRSA}} S I_{\text{MRSA}}$$

$$\frac{dI_{\text{flu}}}{dt} = \beta_{\text{flu}} S I_{\text{flu}} - \gamma_{\text{flu}} I_{\text{flu}} - \delta \beta_{\text{MRSA}} I_{\text{flu}} I_{\text{MRSA}}$$

$$\frac{dI_{\text{MRSA}}}{dt} = \beta_{\text{MRSA}} S I_{\text{MRSA}} - \gamma_{\text{MRSA}} I_{\text{MRSA}} + \gamma_{\text{flu}} I_{\text{both}} - \delta \beta_{\text{flu}} I_{\text{MRSA}} I_{\text{flu}}$$

$$\frac{dI_{\text{both}}}{dt} = \delta \beta_{\text{MRSA}} I_{\text{flu}} I_{\text{MRSA}} + \delta \beta_{\text{flu}} I_{\text{MRSA}} I_{\text{flu}} - \gamma_{\text{both}} I_{\text{both}}$$

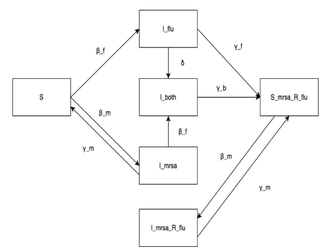
$$\frac{dI_{\text{MRSA\_}R_{\text{flu}}}}{dt} = \beta_{\text{MRSA}} (S_{\text{MRSA\_}R_{\text{flu}}})(I_{\text{MRSA}}) - \gamma_{\text{MRSA}} (I_{\text{MRSA\_}R_{\text{flu}}})$$

$$\frac{dS_{\text{MRSA\_}R_{\text{flu}}}}{dt} = -\beta_{\text{MRSA}} (S_{\text{MRSA\_}R_{\text{flu}}})(I_{\text{MRSA}}) + \gamma_{\text{flu}} I_{\text{flu}}$$

Model Compartments

- S: Susceptible to both influenza and MRSA
- I_{flu}: Infected with influenza only
- I_{MRSA}: Infected with MRSA only
- I_{both}: Co-infected with both influenza and MRSA
- $I_{\text{MRSA\_}R_{\text{flu}}}$: Recovered from influenza but infected with MRSA
- $S_{\text{MRSA\_}R_{\text{flu}}}$: Recovered from influenza but susceptible to MRSA

Flow Diagram of the Model



Recovery Rate of Co-Infected Individuals

$$\gamma_{\text{both}} = \frac{1}{\frac{1}{\gamma_{\text{flu}}} + \frac{1}{\gamma_{\text{MRSA}}} + \alpha}$$

Machine Learning Model Development

To assess co-infection risk and identify under which conditions the risk is highest, we utilized ML algorithms, such as Logistic Regression (LR), Random Forest (RF), and XGB. This will be multi-class classification to predict the risk level of co-infection (high,medium, and low).

Feature Engineering and Target Label Generation

We used the simulation outputs from our custom SIR-SIS model as features in our ML models, such as prevalence of MRSA and flu, rate of change in MRSA and flu, and peak MRSA and flu prevalence. The target variable, risk level, was derived from the number of individuals co-infected with both flu and MRSA (I_both). Thresholds for categorizing risk levels were determined based on quantiles.

Train-Test Split The dataset was split into training (80%) and testing (20%) sets grouped by simulation.

Hyperparameter Optimization To optimize the performance of the models, hyperparameter tuning was performed using random search.

Model Evaluation

Model performance was evaluated using recall for the high-risk class and minimizing false negatives. We selected the best model to conduct scenario-based risk evaluation to extract insights to inform hospitals.

Conclusion

Our model shows the benefits and drawbacks of using both ODE co-infection model and machine learning to model and evaluate co-infection dynamics of flu and MRSA. Many factors and parameters affect co-infection rates in the SIR-SIS model, as is evidenced by the outcome. In future explorations, we may want to add an exposed state to the ODE model and explore other methods such as an agent based model. Modifying certain parameters leads to findings that show the significant variance in co-infection rates. Our findings highlight the importance of monitoring increases in both flu and MRSA prevalence to identify high-risk scenarios of co-infection. The model's most important features, prevalence of MRSA and flu, suggests that real-time infection monitoring could help mitigate co-infection risk in healthcare settings. Also, the scenario-based risk analysis provides valuable insights and a potential resource for hospitals to evaluate conditions under which co-infection risk is highest allowing for timely intervention.

References

Hankyu Jang, et al. 2021. COVID-19 modeling and non-pharmaceutical interventions in an outpatient dialysisunit. PLoS Computational Biology (2021).

Yingzhi Liu, et al. 2021. Outcomes of respiratory viral-bacterial co-infection in adult hospitalized patients. EClinicalMedicine 37 (2021).

Sara E Cosgrove, et al. 2003. Comparison of mortality associated with methicillin-resistant and methicillin-susceptible Staphylococcus aureus bacteremia: a meta-analysis. Clinical infectious diseases 36, 1 (2003), 53–59.