Counteracting the Resistance Epidemic in Vietnam
## A Machine Learning Approach

Anaika Mehra, 2024, Computer Science
Oxford University Clinical Research Unit
Funded by CHW under the Princeton Global Health Fellows program

Introduction

- In Vietnam, there is an ongoing epidemic where a lot of patients are developing resistance to life-saving last-resort antibiotics.
- This resistance is normally developed during their stay at hospitals and can potentially be linked to the use of other antibiotic combinations for treatments of tropical diseases.

Objective of the Study

Discover the potential causes for resistance development and create a predictive model that would suggest best combinations to be used to avoid future resistance.
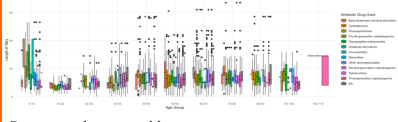

Summary plot created by me..

Methods

- In this study, we used data science methods to measure how important each patient feature (age/sex/ward of admission etc.) was in correlation to future antibiotic resistance development.
- We also developed a predictive ML model to study best antibiotic combinations to use.
- I fully ideated and executed the project myself under my supervisor Huong and with guidance from my colleague Trang.

Explanation and Results

- The feature importance was calculated in 3 ways: random forest feature importance (gini impurity), permutation importance (randomization testing), and SHAP values (cooperative game theory). The results of each method are displayed in the graphs created by me.
- All methods highlight different features as being the most/least important as shown.
- The most significant result is that across all three methods, combinations of antibiotics used was highlighted as a relatively important feature. This means that the medicines given at start of treatment could lead to resistance to medicines that may be needed later in the treatment of tropical diseases.
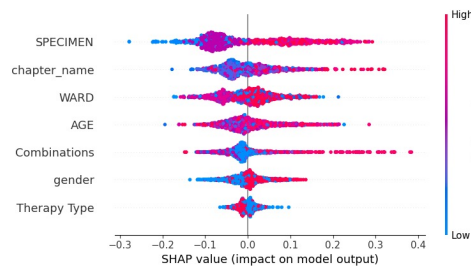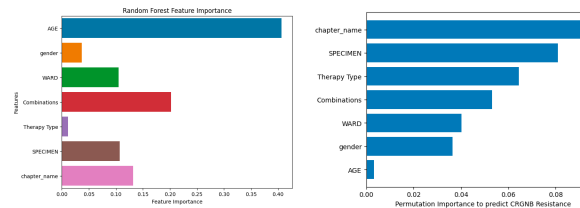
- Another important result was that combinations of medicines correlated relatively less with other features as shown in my graph, making this feature less biased. However, hospitals only collect specific feature data so we cannot eliminate bias fully.

- Our most important result was creating a Random Forest predictive model that can predict how much an antibiotic combination is likely to cause resistance to other antibiotics that may be needed in future treatments pertinent to tropical diseases. We achieved a predictive accuracy of 0.79. Though constrained by the small dataset size, this can be a promising foundation for future research.
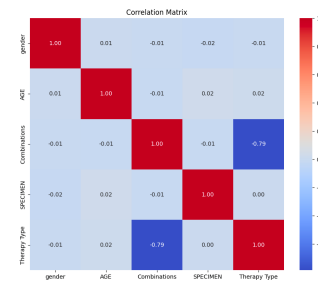
Discussion

The biggest implication of my findings is that the combinations of antibiotics that doctors currently prescribe to patients for tropical diseases in Vietnam might be leading to resistance development to other last-resort antibiotics that the patients may need during the duration of their treatment (esp if they colonize more diseases in the hospital)/later on in their life, often resulting in death.

Questions

- Are the research results biased by metrics not noted by the hospitals?
- Are the results affected by racial and socioeconomic factors that may vary by the hospital choice of patients?

Conclusion

My main conclusion is that there needs to be significant research in antibiotic combination induced resistance in Vietnam using ML methods and predictive modeling.

Acknowledgements

I want to thank OUCRU for giving me a summer full of learning and experiences. I especially want to thank my supervisor Huong and my colleague Trang for their constant support and guidance. Finally, I want to acknowledge CHW and GHP for giving me this once in a lifetime opportunity.