Superheroes for Superbugs: Amplifying Domain Expertise to Combat Antimicrobial Resistance

Digitization of health records has provided an avenue for applying data-driven algorithms to healthcare challenges. One such challenge is antibiotic resistance, which the World Health Organization (WHO) classifies as a major health threat alongside cancer and dementia. When bacteria are exposed to an antibiotic, they develop resistance to it and are harder to treat. Antibiotic resistant organisms or "superbugs" kill around 700,000 people annually (35,000 in the United States). This number is projected to rise to 10 million by 2050 [1]. While developing new drug therapies is one avenue, a more urgent concern is antimicrobial stewardship. Many institutions have implemented an antibiotic stewardship program (ASP) consisting of infectious disease (ID) domain experts (e.g., microbiologists, physicians, pharmacists) to improve antimicrobial prescriptions [2]. However, these domain experts lack tools that allow for sophisticated interactions with the data. Domain experts have different skills and requirements from other end-users (such as crowd workers and data scientists), which must be taken into account when designing systems for them. There is thus a need to develop tools and methods to combat antibiotic resistance.

To further motivate the need for an ASP tool, consider the case of an infectious disease domain expert who wants to study and improve *empiric* prescription. Empiric prescription refers to the treatment prescribed by healthcare providers before the microbiology lab results are available, which typically take two days for *Urinary Tract Infections (UTI)*. The lab culture report contains the infection-causing organism and the results of antibiotic susceptibility testing, i.e., whether an antibiotic will cover the infection. There is an opportunity for reducing resistance during empiric prescription since healthcare providers are working with incomplete information and can assign an unnecessarily broad antibiotic. To evaluate empiric treatment, we are interested in two metrics: coverage and breadth. Coverage refers to if an infection was successfully treated with the empiric antibiotic. On the other hand, breadth refers to the range of organisms covered by an antibiotic. Broad-spectrum antibiotics cover a wide range of organisms but prescribing them increases community resistance and promotes the growth of superbugs, while narrow-spectrum antibiotics cover a smaller range of organisms and risks not covering the infection. To quantify antibiotic breadth, the University's ASP members rated the breadth of different antibiotics for UTI on a scale of 1-100, and each antibiotic's average score was used as its breadth score [3]. A higher breadth score indicates a broader spectrum antibiotic.
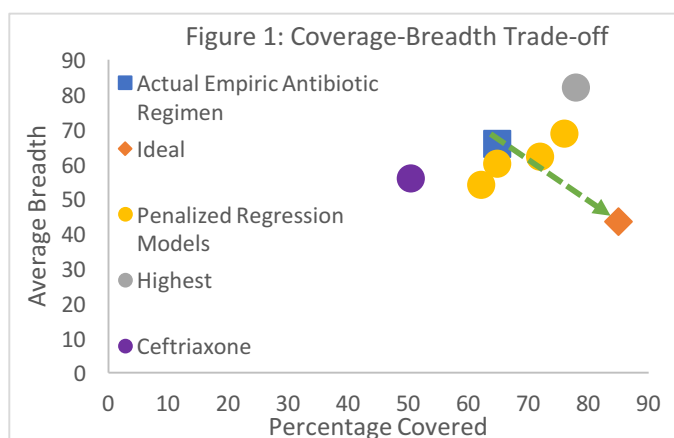

Figure 1: Coverage-Breadth Trade-off

Figure 1 shows the breadth-coverage trade-off for patients admitted to the Ohio State University's Wexner Medical Center between 2011-2016 with a positive UTI diagnosis. Medication assigned within the first 24-hours was designated as empiric treatment (Fig. 1: blue square). We also pulled microbiology lab results for these patients so we could calculate if an antibiotic covered the infection. The x-axis then

denotes the percentage of cases where the infection was covered with the treatment, while the y-axis denotes the average breadth score of the treatment antibiotic for all patient cultures. For comparison, we also assigned an "ideal" antibiotic (Fig. 1: orange diamond) for each patient culture. This is defined as the antibiotic with the lowest breadth score which covers the infection (based on culture results), i.e., the antibiotic a physician would prescribe if they had perfect knowledge. As we can see when comparing the blue square and orange diamond in Figure 1, there is scope for reducing breadth while improving coverage.

Naturally, the next question that the domain expert asks is "can we guide practitioners to do better?" Her goal is to build predictive models for empiric prescription based on patient factors such as demographics, comorbidities, prior infections, and antibiotic exposure. But before even thinking about modeling, she needs to do multi-step data processing. Microbiology lab data is mostly structured but some information needs to be extracted from semi-structured text fields. This requires tokenizing the text, extracting tokens of interest via regular expressions, and then annotating these tokens with SNOMED concepts to identify microbiology entities [4]. This requires interaction between the domain expert and the data scientist since the latter does not know what information is relevant.

Once the data is re-structured to the required format, the domain expert is left with a missing data problem. This is because the microbiology lab only reports susceptibility tests for a subset of antibiotics. At the point of care, physicians can rely on their domain knowledge to infer the remaining susceptibilities, so the missing data do not pose a problem. But when using these data for predictive models, all susceptibilities need to be filled in. Again, expert interaction is needed, in this case, to encode their domain knowledge as "rules" to fill in the unreported data. Automatic augmentation methods would require external knowledge bases or biomedical literature extraction which lack transparency and do not take local patterns or regulations into account. The goal is to replicate physician decision-making to augment the dataset. To this end, I built Icarus [5], an interactive data completion system that enabled domain experts to fill in 50,000 unreported data points in a couple of hours. Icarus amplifies domain expertise by addressing two main challenges: first, it uses concepts from information theory to optimize experts' edits by guiding them to useful update sets, and second, it accelerates their task by suggesting general rules via a depth-first search algorithm on the database schema.

Due to the subjectivity of physician decision-making, multiple experts are needed to augment the dataset, who individually make decisions and then come together to resolve conflicts. Amplification for the former process is addressed by Icarus and to amplify expertise in the latter stage, I designed Delphi, an interactive rule editor [6] which visualizes relationships between rules of different experts and the impact of the rule on the data. Using Delphi, experts can interactively edit rule sets to consolidate conflicts and redundancies between their decisions. Delphi addresses the challenge of effectively visualizing and interactively editing multi-dimensional relationship data. The network representation maps to the expert's mental model of the antibiotic hierarchy and allows them to quickly navigate the ruleset.

Once multiple experts have filled in the dataset and their conflicts have been resolved, our domain expert finally has an analysis-ready dataset. She works with colleagues in biostatistics to model the resistance of five antibiotics, providing input on variables of interest and their clinical

significance. The output of the models gives her a probability of coverage for each of the five antibiotics. A strategy is now needed to recommend one antibiotic based on the five probabilities. Different strategies are shown in Figure 1. The grey dot corresponds to selecting the antibiotic with the highest probability of coverage, which is usually a broad-spectrum antibiotic. This strategy covers 78% of the infections but at almost double the breadth-score of the ideal case. As a baseline strategy, the purple dot shows the effect of always prescribing ceftriaxone, which is the practice of emergency department physicians if the patient is relatively healthy. While this strategy prevents unnecessary resistance exposure since ceftriaxone is a narrow-spectrum drug, it only covers 55% of the cases in our dataset, i.e., 45% of patient cultures grow an organism that is resistant to ceftriaxone. The yellow dots show different strategies for selecting an antibiotic based on the probabilities of coverage. An example strategy would be to select the narrowest antibiotic with at least a 60% probability of coverage. Other strategies involve selecting different probability thresholds for different antibiotics. As seen in the figure, increasing coverage requires increasing breadth, i.e., while they can improve on physician prescription (higher coverage at the same breadth or lower breadth at the same coverage), it is difficult to model the ideal.

Since the data can provide insight but we do not have a perfect oracle, the question then becomes how should this information be presented to practitioners or decision-makers. Domain experts would gain more value by seeing the impact of different variables on the probabilities. This can then help members of the ASP make policy decisions on what antibiotics should be prescribed to different populations to maximize coverage while minimizing the risk of resistance. To this end, my current research focuses on building a microbiology datamart which will provide an interactive interface for ASP domain experts to view, manipulate and run what-if analysis on microbiology data to improve antibiotic stewardship. The research challenges here are deciding what information should be presented, how it should be presented, and types of guidance and interactions which should be presented.

Research in this area empowers microbiology domain experts to interact and leverage patient data to improve patient care. They can study local patient trends and resistance patterns. This provides insight on antibiotics that are being overused and which organisms are commonly mistreated. Antimicrobial policymakers currently lack tools that allow for direct and sophisticated interactions with the data which is addressed by my research. Further, my experiences in working with microbiology domain experts have exposed me to the technical challenges of building systems for them. Domain experts are very different end-users from crowd workers and data scientists. They have advanced domain knowledge which needs to be amplified by the interface and the tasks. This can be done by summarizing information to provide an overview, guiding them to insights and accelerating their input by making inferences from their interactions. Additionally, domain experts have limited data literacy, i.e., they are sometimes unfamiliar with data science terms such as clustering or decision trees. Hence, it is important to design tasks and interactions that they are familiar with in their daily practice and match their mental model. As more fields move on to data-driven decision making, it is imperative to improve data literacy among domain experts. Tools such as Icarus can be used as exercises to better educate infectious disease trainees in empiric prescription. The datamart can be used to visually educate antibiotic prescribers on the dangers of antibiotic resistance. Further, the interactions of researchers on the datamart can be used to improve data literacy among

novices. In addition to medical professionals, the microbiology data mart can provide a concrete use-case to undergraduate data science students on designing for experts. After providing students with the scenario and base infrastructure of the data mart, they can be tasked with designing visualizations and interactions to amplify expertise. We are in the process of open-sourcing Icarus by testing interoperability with infectious disease researchers at NorthShore and Stanford.

The next step for this research would be to go from local institutional microbiology datamarts to a global data store, which pulls in information from multiple health centers and hospitals. As a first step, I hope to integrate the open-source critical care MIMIC dataset from MIT and then slowly expand to include information from collaborators at other institutions. The goal would be to develop a standard data model for microbiology data. Along with expanding to multiple locations, another aspect would be to incorporate unstructured sources. Our current approaches are constrained to representing structured data from microbiology lab results, but as is common knowledge, unstructured notes in EHR often contain critical information not captured by structured fields. Hence, it would be interesting to incorporate different phenotype signals from free-text notes to augment the microbiology datamart. Other sources include biomedical literature from which results of resistance studies can be extracted. Finally, social media is another data source. In a 2010 study, Scanfeld et al. [7] studied 1000 tweets relating to antibiotics. They found that antibiotic side-effects and negative reactions was the third most talked about topic (113/1000), while another 92 tweets discussed antibiotic resistance. Thus, there is potential to incorporate this information into the datamart to enable better public surveillance. While unstructured data has been used for building predictive models, more attention needs to be paid on making these models transparent, explainable and accessible to practitioners. The data need to represent the uncertainty of the models as well as provide actionable information. The latter involves providing causality information that can be modified to improve outcomes. Tools should provide diverse interactions that are not limited to syntactical changes but also allow the user to perform complex semantic updates.

The designation of antibiotic resistance as a global threat by WHO has led to many funding initiatives. The CDC's Antibiotic Resistance Initiative started with 160 million in 2016 has increased to 168 million in 2018. NIH's NIAID institute is specifically interested in antibiotic resistance. As a postdoc, I hope to apply for NIH's K-99 award.

A central microbiology datamart will allow the U.S. to track infections, antibiotic usage, and misuse. It will enable better antibiotic prescription and guide policies tailored to different institutions based on their local trends. Additionally, it can be used to audit institutions and physicians for incorrect prescriptions and misuse. The datamart will allow practitioners to transparently understand and interact with data and run "what-if" analysis as opposed to following decisions from black-box models. My experience in optimizing [8] and building interactive interfaces [9], along with working with infectious disease experts makes me the ideal candidate for this undertaking.

References

1. Blanchard S, Mailonline. The Ganges is teeming with 'astronomical' amounts of antibiotic-resistant bacteria. Daily Mail Online 2019. https://www.dailymail.co.uk/health/article-7824481/The-Ganges-teeming-astronomical-amounts-antibiotic-resistant-bacteria.html (accessed December 24, 2019).
2. Bartlett JG, Gilbert DN, Spellberg B. Seven ways to preserve the miracle of antibiotics. Clinical Infectious Diseases. 2013 May 15;56(10):1445-50.
3. Patterson ES, Dewart CM, Stevenson K, Mbodj A, Lustberg M, Hade EM, Hebert C. A mixed methods approach to tailoring evidence-based guidance for antibiotic stewardship to one medical system. In Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care 2018 Jun (Vol. 7, No. 1, pp. 224-231). Sage CA: Los Angeles, CA: SAGE Publications.
4. Rahman P, Hebert CL, Lai AM. Parsing complex microbiology data for secondary use. In AMIA 2016.
5. Rahman P, Hebert C, Nandi A. ICARUS: Minimizing human effort in iterative data completion. Proceedings of the VLDB Endowment. 2018 Sep 1;11(13):2263-76.
6. Rahman P, Chen J, Hebert C, Pancholi P, Lustberg M, Stevenson K, Nandi A. Exploratory Visualizations of Rules for Validation of Expert Decisions. IEEE VIS Workshop on Data Systems for Interactive Analysis. 2018.
7. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: Twitter and antibiotics. American journal of infection control. 2010 Apr 1;38(3):182-8.
8. Rahman P, Nandi A. Transformer: a database-driven approach to generating forms for constrained interaction. InProceedings of the 24th International Conference on Intelligent User Interfaces 2019 Mar 17 (pp. 485-496).
9. Rahman P, Jiang L, Nandi A. Evaluating interactive data systems. The VLDB Journal. 2019:1-28.