Derivation of Expert Consensus Rules for Missing Antimicrobial Susceptibility Data

Protiva Rahman, BS, Erinn M. Hade, PhD, Arnab Nandi, PhD, Preeti Pancholi, PhD, Mark Lustberg, MD, PhD, Kurt Stevenson, MD, MPH, Courtney Hebert, MD, MS
The Ohio State University, Columbus, OH

Background

In order to develop predictive models for empiric antibiotic prescribing, we required input from experts in infectious diseases to provide unavailable susceptibility data for a dataset with nearly 75,000 missing cells. Sparse and missing susceptibility data is a common problem when using clinical microbiology data. This is, in part, because microbiology labs report only susceptibility information for the subset of antibiotics commonly used to treat the organism that grows in culture. When a clinician is deciding on antibiotic coverage based on culture results, they often need to apply their domain knowledge of antibiotic susceptibilities, as opposed to only the reported microbiology results. This knowledge-based, and sometimes subjective, inference of coverage is a common activity when treating patients. Our challenge is in translating these *"rules*" to a dataset with several thousand cultures, without introducing undue bias due to subjectivity. This poster discusses our experience achieving consensus among experts in filling in these data. Consensus techniques such as Delphi[2] often require multiple iterations and do not prioritize reduction in expert effort, which is essential for complex datasets. Three experts on the study team used ICARUS[1], a newly developed data completion tool to fill in the unreported data through iterative rules, resulting in three sets of rules and three completed datasets. Our objective was to arrive at (1) a consensus dataset with filled in values and (2) a set of consensus rules.

Method

Consolidating rules is a complex task because: different rules may have been used by our experts to fill in the same set of cells, the rules are often conditional on previous rules, and even when two rules fill in the same result set, they can have semantically different meanings. We use a three-phase process to address these challenges. Phase 1 - Identify conflicts and generate unanimous agreement rules: (a) For those cells where all three experts agree on the value for the cell, we generate rules through a modified decision tree algorithm that splits based on semantic information such as antibiotic and organism classification; (b) Compare result sets of three experts to identify a set of conflicting rules (i.e., rules that fill in opposite values for a subset of common cells). Phase 2 - Consensus Meeting: Four experts, three of whom had interacted with the data using ICARUS, met to accept/reject rules generated in Phase 1 and to discuss conflicts. Accepted rules were applied and the Phase 1 algorithm was repeated to generate rules for remaining missing cells, which were again validated through experts. Phase 3 - Validate rules: The final set of rules were applied and then validated through an antibiogram, which was visualized by plotting resistance against antibiotic.

Results

Phase 1(a) filled in 52,569 (70%), Phase 1(b) filled in 16,647 (22%), while rule generation after consensus meeting in Phase 2 filled in 4,519 (6%). Our process thus filled in 98% of the missing cells. All rules were manually validated to be semantically correct. By using this method, we were able to focus our discussion at the consensus meeting to a set number of items.

Conclusion

Our method was developed to increase the rigor and reproducibility of a major data cleaning task of our study. Automatic input consolidation reduced the number of expert meetings required and focused conflict resolution discussions to a few items. The finalized rules can be shared for result replication, without the need to share data.

Acknowledgments

Research reported in this publication was supported by the NIAID of the NIH under Award Number R01AI116975.

References

1. Rahman P, Hebert C, and Nandi A. Icarus: Minimizing human effort in iterative data completion. In Submission
2. Dalkey N, Brown B, Cochran S. The Delphi method, III: Use of self-ratings to improve group estimates. RAND CORP SANTA MONICA CALIF; 1969 Nov