Heart of England NHS Foundation Trust

WARWICK THE UNIVERSITY OF WARWICK

Towards intelligent guidelines: can machine learning guide empirical antibiotic choice?

Ed Moran[1], Esther Robinson[1], Chris Green[1], Matt Keeling[2], Ben Collyer[2]

Heartlands Hospital[1], Birmingham and Zeeman Institute[2], University of Warwick

INTRODUCTION

- Every hospital produces empirical antibiotic treatment guidelines. These indicate the agent that should be given for a specific condition until such time as an organism is isolated from a specimen (e.g. urine, blood). Such guidelines are largely didactic documents.
- Given the clear association between antibiotic use and the development of resistance within populations it is critical that a balance is struck between risking potentially ineffective, or unnecessarily broad, empirical treatment.
- The modern hospital has access to enormous amounts of information about each patient, including previous admissions, microbiology, diagnoses, electronic prescribing data and basic demographics.
- We hypothesised that this information could be used to generate an individualised antibiotic resistance risk assessment for individuals requiring antibiotics. This would direct prescribers to narrower spectrum agents for those at low risk of resistance, and restrict the use of broad agents to those at proven high risk.
- Such a system could be integrated into a hospital EPR system, seamlessly producing specific "intelligent" recommendations at the point of prescribing electronically.
- As a proof of principle we assessed the accuracy of a machine learning algorithm (MLA) trained in predicting antibiotic resistance in *E. coli, K. pneumoniae*, and *Ps. aeruginosa* isolated from patients within 48 hours of admission.
- We calculated length of stay and mortality for patients treated with effective and ineffective initial therapy in an attempt to guide the extent to which the MLA could accept the risk of ineffective empirical treatment.

METHODS

- We extracted basic demographics, previous admissions dates, and electronic prescribing data for all patients from whom *E. coli, K. pnemumoniae*, or *Ps. aeruginosa* were isolated over a 7 year period. NHS data was anonymised before sharing with academic partners.
- This represented 27350 patient admissions and antibiotic initiations with available associated resistance testing information. 80% of this data was used to train an XGBoost machine-learning algorithm. XGBoost is a popular open-source Machine Learning library that uses ensembles of classification decision-trees to generate predictive models (figure 1) and has been utilised by multiple winning teams competing in machine learning competitions [1].
- We compared the performance of medical staff and the trained algorithm in appropriate selection of co-amoxiclav, piperacillin-tazobactam and carbapenem therapy. This represented 2411 empirical antibiotic initiations.

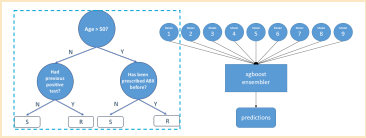

Figure 1. ABOVE: XGBoost model structure. Individual classification tree models (e.g. dashed box) are combined to make predictions.
Figure 2. RIGHT: Receiver operating characteristic (ROC) curve for algorithmic predictions of piperacillin-tazobactam resistance.

RESULTS

Who is best at selecting appropriate empirical antibiotic therapy at admission? Medical staff or the machine learning algorithm?

- For every patient the MLA produces a number between 0 (susceptible) and 1 (resistant). This is determined by patient's risk factors and what the MLA has learnt from the test data regarding their influence on resistance.
- The "bravery" of the MLA's choice of antibiotic can be varied by adjusting the threshold at which it labels a patient as likely to grow sensitive or resistant organisms.
- E.g. if the patient's value was 0.4 and we set a threshold for predicting they had a resistant organism at 0.5, the MLA would consider the patient as likely to have a sensitive organism.
- Settling the "bravery threshold" low means it will be likely to "over"-treat people with sensitive organisms, setting it high means it is likely to "under"-treat people with resistant organisms.
- Optimum reduction in over-prescribing with no significant rise in under-prescribing compared to medical staff is achieved with a threshold of around 0.35 (figure 3).

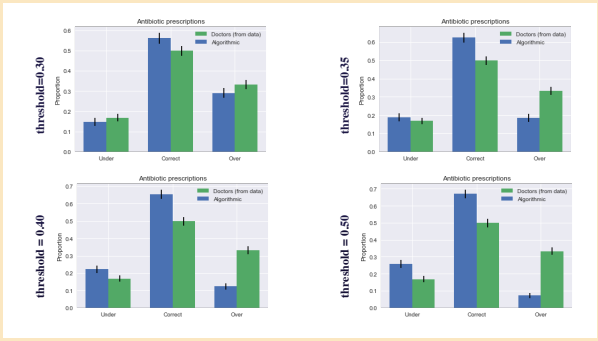

Figure 3. Medical staff (green) compared to the MLA (blue) at different "bravery thresholds"

Does under-prescribing in the first 96 hours matter?

- Patients were prescribed a number of different antibiotics empirically at the point of admission. These were refined once culture and sensitivity information became available (figure 4).
- The level of risk one is willing to accept in this initial decision – whether that be as an individual prescriber or in designing a MLA – requires knowledge of the consequences of empirical under-prescribing.
- Length of stay and mortality rates were calculated for those patients who received effective and potentially ineffective initial treatment (figure 5).
- Ineffective therapy in the initial period before culture results was associated with an increased length of stay and death rate. However, this effect was probably confounded by the increased likelihood of older age and frailty in those patients at risk of resistant organisms.
- This is evidenced by the similar length of stay and death rates in people treated with piperacillin-tazobactam with organisms sensitive to this agent. Further work is needed to control for these factors.

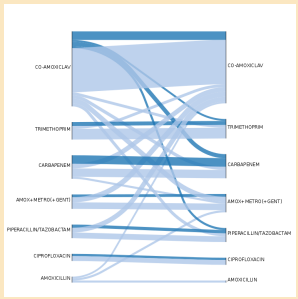

Figure 4. Primary antibiotic prescribed in first 96 (left) and that received after 96h (right).
Data includes all 3 organisms isolated in any specimen (urine, blood, sputum). Light blue=co-amoxiclav sensitive organism, dark blue=resistant

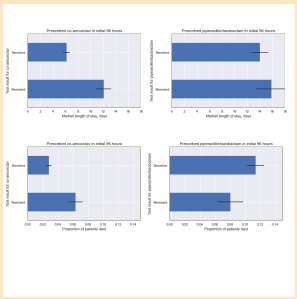

Figure 5. Length of stay and mortality related to likely efficacy of the main antibiotic received in the first 96h of admission

DISCUSSION

- Medical staff selection of empirical antibiotics was appropriate approximately 50% of occasions. 33% of initiations were excessively broad and 17% too narrow.
- A machine learning algorithm trained with data representing 2411 patient antibiotic initiations performed well when tested on the whole data set.
- A threshold of 0.35 produced no significant increase in under-prescribing decisions and increased appropriate treatment by 27%. Use of unnecessarily broad spectrum agents dropped by 44%.
- Outcome data indicated that ineffective initial treatment is potentially associated with adverse outcomes. However confounding factors which are themselves associated with an increased risk of resistant organisms would need to be controlled for prior to setting a "bravery level" for a MLA.
- There were a number of other limitations to this study. We restricted the data to emergency admissions, we were unaware of antibiotics that might have been prescribed on paper in A&E or by GPs, we restricted MLA training data to initiations with co-amoxiclav, piperacillin-tazobactam or carbapenems, a number of "over-treatment" decisions may have been due to penicillin allergy (the Trust's current default being a carbapenem).

CONCLUSIONS

- Machine learning techniques show considerable promise in the area of empirical antibiotic selection. They would allow more nuanced antibiotic selection, individualised according to the multiple risk factors of a specific patient.
- The technique can be easily and quickly applied to data held by different hospitals, creating a MLA suitable for each specific institution.
- Our study indicates that this could lead to reduced use of broad-spectrum agents. Improved antibiotic stewardship would likely lead to reduced patient complications (e.g. C. diff diarrhoea), costs, and a reduction in the later emergence of resistance.
- Further studies are needed in the following areas: replicating this proof of principle in different institutions, understanding the consequences of ineffective initial therapy so that an appropriate "bravery index" can be selected, and developing an MLA with more comprehensive datasets suitable for use in a clinical trial.
- Ultimately MLAs should sit behind hospital electronic prescribing systems, guiding appropriate antibiotic choice in line with a patient's "real time" individual risk factor information.

ACKNOWLEDGEMENTS

- Ben Collyer is an INTEGRATE AMR Early Career Research Fellow, and would like to acknowledge support and funding from EPSRC's 'Bridging the Gaps' initiative.
- NHS data was shared only after anonymisation and under a Privacy Impact Assessment.

REFERENCES

1. Distributed Machine Learning Community (2017). *Machine Learning Challenge Winning Solutions*, retrieved from https://github.com/dmlc/xgboost/tree/master/demo#machine-learning-challenge-winning-solutions