

Machine learning-assisted selection of antibiotic prescription

Xavier Didelot^{1,*} and Koen B. Pouwels²

¹ School of Life Sciences and Department of Statistics, University of Warwick, UK

² Health Economics Research Centre, Nuffield Department of Population Health,
University of Oxford, UK

* xavier.didelot@warwick.ac.uk

Machine learning models can use information on patients and their clinical history to help physicians select the antibiotics most likely to successfully treat urinary infections despite growing levels of resistance.

Recent years have seen a worrying increase in antibiotic resistance levels of many bacterial infections. Antibiotic resistance not only makes it difficult to treat bacterial infections, but also decreases the effectiveness of antibiotic prophylaxis needed for safe surgeries, organ transplantation and cancer treatment¹. There is an urgent need for new effective antibiotics, however, the antibiotic development pipeline is dry. Without government intervention the research for new antibiotics is rarely profitable, which led most major pharmaceutical companies to leave the field¹. It is therefore crucial to use the antibiotics we have at our disposal in an optimised way, to avoid the risks of both treatment failure and increasing resistance levels further². In this issue of *Nature Medicine*, Yelin et al.³ describe a strategy for combating drug resistance caused by mismatched antibiotic prescriptions in urinary tract infections (UTIs).

Antibiotic treatment for various bacterial infections, such as UTIs, is typically started empirically without knowing to which antibiotics the bacteria causing the infection are susceptible. UTIs are one of the most common infections encountered in primary care⁴, and account for 29%-66% of antibiotic prescriptions in care home settings⁵. The bacteria causing UTIs are often carried asymptomatically in the human body and are therefore frequently exposed to antibiotics, including those taken to treat other infections⁶. Consequently, bacteria causing UTIs are frequently resistant to various commonly used antibiotics⁷. Physicians of UTI patients are therefore routinely faced with several difficult questions. Which antibiotic is most likely to cure the patient? How can we preserve the usefulness of antibiotics in the long term? Should published guidelines be followed, or is there something special about the patient that justifies a different course of action?

Yelin et al.³ analyse data from more than 700,000 UTI cases occurring between 2007 and 2017 in Israel, including information on both the patients' demographics, clinical history, and previous history of antibiotic use. The resistance of cultured bacterial pathogens against six antibiotics frequently prescribed against UTIs was found to be associated with several demographic factors, especially the age of the patient, their

gender and whether or not they reside in a retirement home. Many patients contracted multiple UTIs during the ten-year study period, in which case the resistance profiles were often similar between their UTIs, suggesting either relapse or reinfection from the same source depending on the time span. The history of antibiotic use was also found to be associated with resistance of the bacterial pathogen, with the strongest association being between use and resistance of the same antibiotic, but other less expected associations exist too, as has been previously noted⁶. There are various mechanisms by which use of one antibiotic may select for resistance against another unrelated antibiotic^{3,6,8}.

Beyond the importance for our understanding of the drivers of antimicrobial resistance within hosts and at the population scale^{8,9}, the study by Yelin et al.³ has the potential for the application to personalised medicine. The authors applied a machine learning method, called gradient boosting decision trees, to derive an algorithm which takes as input all available information on demographic factors, previous infection history and antibiotic use, and returns as output a prediction for the resistance profile of a new infection. In machine learning, algorithms are trained to perform complex tasks by recognising patterns in high dimensional large datasets. In contrast to more traditional statistical regression techniques, no explicit relationships have to be assumed between input and output parameters. Because part of the recognised patterns might be due to random variation and specific to the training dataset, it is crucial to evaluate the performance of the trained model in a test dataset (Figure 1). In this study, in order to fairly assess the accuracy of the predictions, the model was trained using only the first nine years of data, so that the remaining tenth year could be used for independent testing. This benchmark demonstrated that the model had a strong capacity to predict resistance to specific antibiotics. Considering the choice between the six drugs frequently prescribed against UTIs, it was found that in 8.5% of cases the physicians had prescribed an inappropriate drug, that is one to which the infection was resistant. This is only slightly better than the 10% of inappropriate prescriptions that would

happen if the drug was chosen at random. On the other hand, by using the predictive computer model, the proportion of inappropriate prescription was reduced to 5%.

The study by Yelin et al.³ paves the way for the use of machine learning-assisted decision in the prescription of antibiotics against UTIs in a way that minimises the risk of treatment failure. However, the model would have to be retrained separately based on relevant data for application to other populations and to keep up with current trends in resistance. The availability of such training data may be limited in some settings, for example in patient populations for which doctors do not routinely send samples for antibiotic susceptibility testing. In addition, detailed data about patients' history of antibiotic use and previous test results are not always available to the prescribing physicians. It would be interesting to see in future work whether a model with less detailed input data requirements could also improve selection of antibiotics that are effective against UTIs. When detailed data are available for most patients, similar strategies could be used for the treatment of other bacterial infections, as was recently proposed for bloodstream infections in a hospital setting¹⁰. Accurate prediction of resistance against different antibiotics is directly beneficial from the point of view of the patients since it avoids treatment failures. It could also have additional long terms benefits, for example enabling the use of more targeted antibiotics, reducing the need to use multiple antibiotics to cure the same infection and lowering the risk of onward transmission^{2,11}. All these factors would contribute to a reduction in the selective advantage of resistant pathogens, and since resistance often has a significant fitness cost, this could lead to a reduction of the overall resistance levels and an easing of the global threat posed by antibiotic resistance¹².

Competing financial interests

The authors declare no competing interests.

References

1. Roope, *et al.* The challenge of antimicrobial resistance: what economics can contribute. *Science* **364**, eaau4679 (2019).
2. Whittles, L. K., White, P. J. & Didelot, X. Estimating the fitness benefit and cost of cefixime resistance in *Neisseria gonorrhoeae* to inform prescription policy: A modelling study. *PLoS Med.* **14**, e1002416 (2017).
3. Yelin, I. *et al.* Personal clinical history predicts antibiotic resistance in urinary tract infections. *Nat. Med.* (2019).
4. Dolk FCK, *et al.* Antibiotics in primary care in England: which antibiotics are prescribed and for which conditions? *J. Antimicrob. Chemother.* **73** (Suppl 2), ii2-ii10 (2018).
5. Ahmed H, *et al.* Incidence and antibiotic prescribing for clinically diagnosed urinary tract infection in older adults in UK primary care, 2004-2014. *PLOS ONE* **13**, e0190521 (2018).
6. Pouwels, K. B. *et al.* Association between use of different antibiotics and trimethoprim resistance: Going beyond the obvious crude association. *J. Antimicrob. Chemother.* **73**, 1700–1707 (2018).
7. Flores-Mireles, A. L., Walker, J. N., Caparon, M. & Hultgren, S. J. Urinary tract infections: Epidemiology, mechanisms of infection and treatment options. *Nat. Rev. Microbiol.* **13**, 269–284 (2015).
8. Didelot, X., Walker, A. S., Peto, T. E., Crook, D. W. & Wilson, D. J. Within-host evolution of bacterial pathogens. *Nat. Rev. Microbiol.* **14**, 150–162 (2016).
9. Holmes, A. H. *et al.* Understanding the mechanisms and drivers of antimicrobial resistance. *Lancet* **387**, 176–187 (2016).
10. Oonsivilai M, Mo Y, Turner P, Cooper BS. Using machine learning to guide targeted and locally-tailored empiric antibiotic prescribing in a children's hospital in Cambodia. *Wellcome Open Res.* **3**, 131 (2018).
11. McAdams D, *et al.* Resistance diagnostics as a public health tool to combat

- antibiotic resistance: a model-based evaluation. *PLOS Biol.* **17**, e3000250 (2019).
12. Andersson, D. I. & Hughes, D. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat. Rev. Microbiol.* **8**, 260–271 (2010).

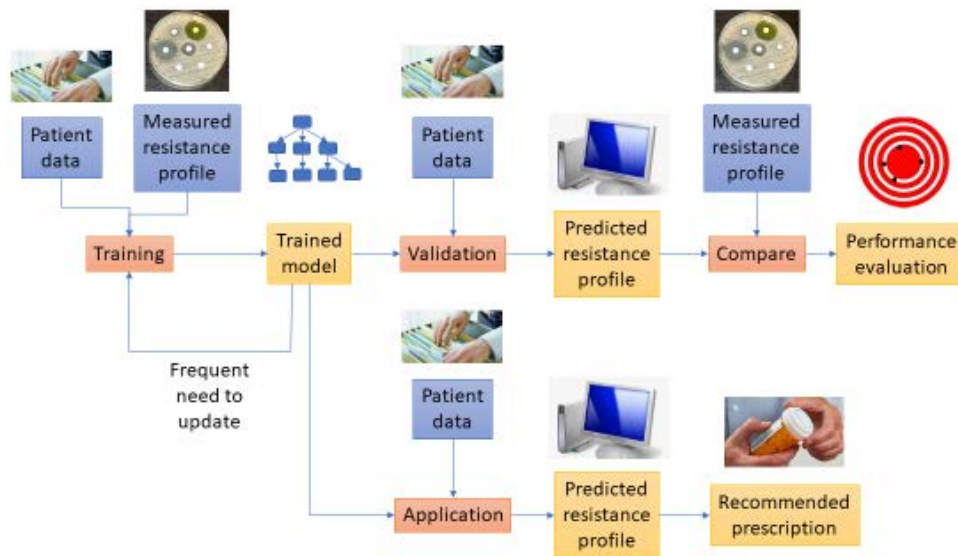


Figure 1. The machine learning method is first trained based on a large number of patient records and antibiotic resistance measurements. The resulting trained model is able to predict resistance profiles based on patient data, which in turn is used to select the antibiotic prescription most likely to succeed in curing the infection. The performance of this trained model is evaluated based on a separate set of patients, to allow for comparison between the predicted and observed resistance measurements.