33rd ECCMID Copenhagen, Denmark 15–18 April 2023
The congress of ESCMID

O0140
Machine learning based white blood count estimation for individualised antimicrobial cessation

11. Other

11d. Digital health and infection (incl. AI, data mining, informatics)

#O0140

Accepted abstract

W. Bolton [1, 2, 3], B. Hernandez [1], R. Wilson [1, 4], D. Antcliffe [5, 6], P. Georgiou [1, 7], A. Holmes [1, 4, 8], T. Rawson [1, 4]

[1]Centre for Antimicrobial Optimisation, Imperial College London - London (United Kingdom), [2]AI4Health Centre for Doctoral Training, Imperial College London - London (United Kingdom), [3]Department of Computing, Imperial College London - London (United Kingdom), [4]National Institute for Health Research, Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, Imperial College London - London (United Kingdom), [5]Department of Critical Care, Imperial College Healthcare NHS Trust - London (United Kingdom), [6]Faculty of Medicine, Imperial College London - London (United Kingdom), [7]Centre for Bio-inspired Technology, Department of Electrical and Electronic Engineering, Imperial College London - London (United Kingdom), [8]Department of Infectious Diseases, Imperial College London - London (United Kingdom)

Background

It is challenging on an individual patient level to determine when cessation of antimicrobial treatment is most appropriate. Stopping too early can result in ineffective treatment while unnecessary exposure causes side effects and drives antimicrobial resistance (AMR). Total white blood count (WBC) can be used as a proxy for treatment response to assist with cessation decision-making. Based on our previous research we utilised a recurrent neural network autoencoder and a synthetic control-based approach to estimate patients' WBC for any given day, under the alternative scenarios of stopping versus continuing antibiotic treatment (Figure 1).

Methods

The MIMIC-IV database was used and filtered based on intravenous antibiotic treatment (<8 days) in the Intensive Care Unit (ICU). 22 input features including lab test results, vitals and assessment scales were extracted based on prevalence and clinical advice. Features were normalised, aggregated by day and missing values highlighted or forward filled. Data was split into training, validation and testing sets for model development and evaluation. Control days defined as days where the real outcome under antibiotic treatment is known, allows the

reliability of estimations to be tested. While on impact days where the real outcome is unknown, the effect of alternative antibiotic treatment can be assessed.

Results

7,867 unique ICU stays were identified, the mean WBC was 10.88 with an interquartile range of 5.50. The standard autoencoder achieved a root mean square error (RMSE) of 3.33 for WBC prediction on the unseen test set. For test dataset synthetic WBC estimations an embedding was created for every day of each patient's stay (Figure 2). Results on control days show estimations are reliable with a RMSE of 3.24 and 3.44 for stopping and continuing respectively. As expected, the mean delta on control days was minimal at <0.01 for stopping and 0.02 for continuing. On impact days however, statistically significant differences were observed with 0.71 and -0.31 for stopping and continuing respectively (Table 1).

Conclusions

Results demonstrate the machine learning model can estimate patient outcomes under alternative antibiotic treatment and indicate that early cessation can decrease WBC. Future research will discern the ability of such a tool to influence antimicrobial decision-making.

Synthetic estimation methodology

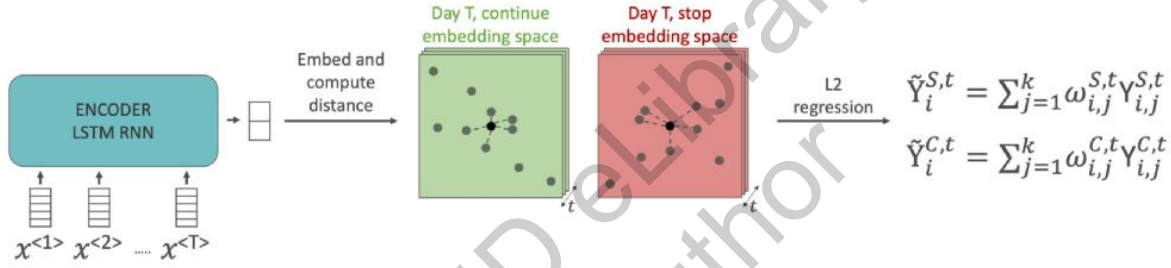

Figure 1: Adapted from Bolton et al. (2022). To create synthetic estimations under antibiotic temporality, an embedding is created for every day of each patient's stay within the dataset; embedding spaces are partitioned temporally and based on if the patient stopped or continued antibiotics on that particular day. The closest k neighbours are selected as donors from each embedding space and L2 regression returns weights that minimise the square difference between the patient and the donors. Stop and continue synthetic estimations are created as the weighted average of donors' outcomes.

LSTM, Long Short-Term Memory; RNN, Recurrent Neural Network

Principal component analysis embeddings

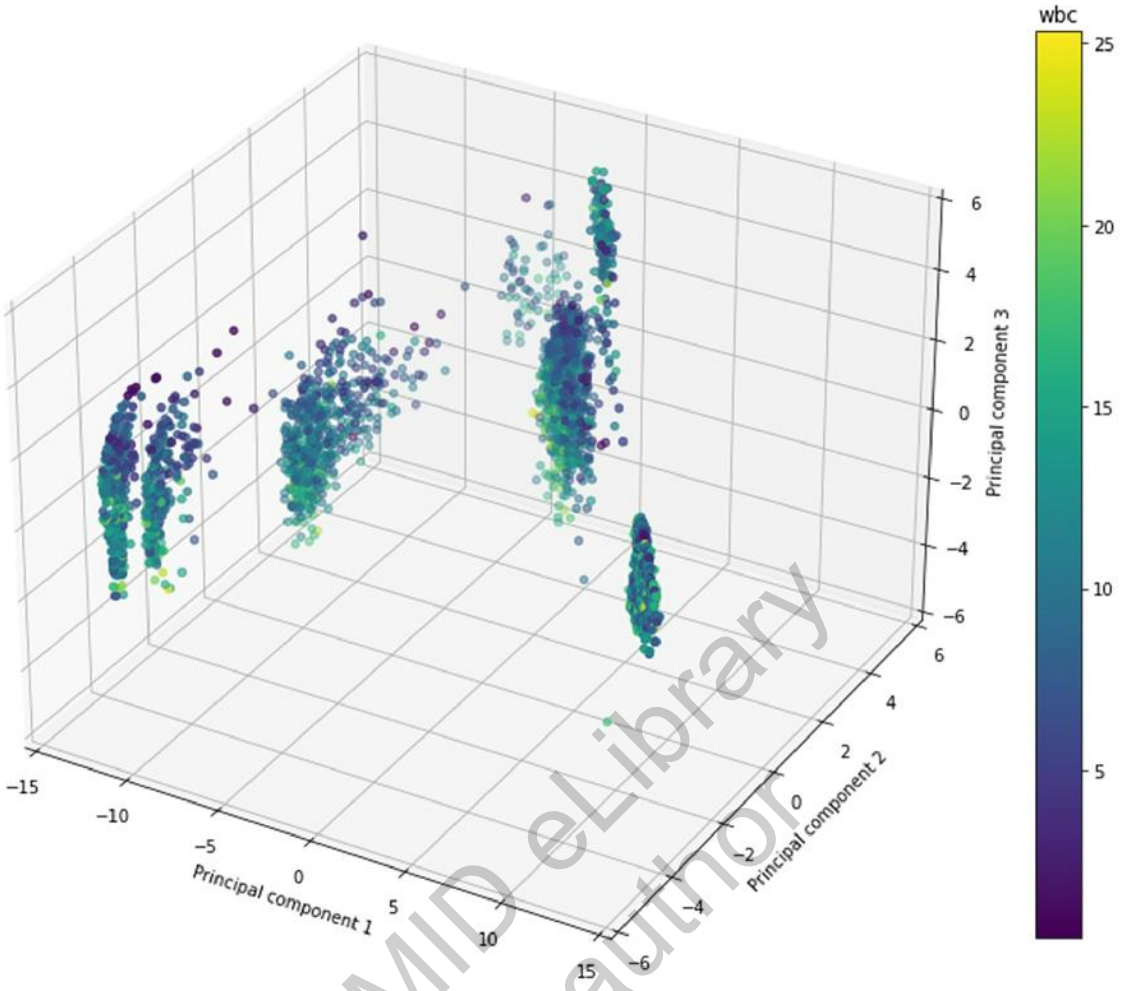

Figure 2: All embeddings within the unseen test set visualized through principal component analysis. Six clear clusters and differentiation by wbc can be observed.

WBC, White Blood Count

Synthetic estimation results

SCENARIO	DAY(S)	WBC			
		Mean delta (days, p-value)	MAPE	MAE	RMSE
STOP	IMPACT	0.71*, <0.01	0.34	2.99	3.89
	CONTROL	0.00, 0.06	0.32	2.50	3.24
CONTINUE	IMPACT	-0.31*, <0.01	0.33	2.53	3.18
	CONTROL	0.02*, 0.01	0.32	2.70	3.44

Table 1: Synthetic wbc estimation results on the unseen test set for stop and continue scenarios.

* Statistical significance with alpha set at 0.05. WBC, White Blood Count; MAPE, Mean Absolute Percentage Error; MAE, Mean Absolute Error; RMSE, Root Mean Squared Error