Improving Clinical Early Warning Systems by Learning Physiological Trajectories

Allan Pang [1,2,3], Alwyn Kotzé A [1], Geoff Hall [1,3], Owen Johnson [3], Marc de Kamps [3,4]
[1] Leeds NHS Teaching Hospitals Foundation Trust, Leeds, UK [2] Academic Department of Military Anaesthesia and Critical Care, Royal Centre for Defence Medicine, UK [3] UKRI CDT for AI in Medical Care and Diagnosis, University of Leeds, UK
[4] Alan Turing Institute, London, UK

Background.

Early Warning Scoring (EWS) systems use physiological data to highlight patients at risk of impending in-hospital death. Current EWS systems in clinical practice use rules-based thresholds based on aggregated scores from physiological data to identify patients at risk of death. Due to their simplicity and scalability, these systems are ubiquitous in inpatient care, with the National Early Warning Score (NEWS2) being the standard of care within the UK [1]. Other rules-based thresholds EWS systems include the Electronic Cardiac Arrest Triage (eCART) score, used in the USA [2]. A systematic review of EWS systems has brought the predictive value of these systems into doubt due to methodological flaws in the development and validation, leading to over-optimistic estimates of model performance [3]. The current EWS modelling framework has two limitations. They decompose physiological signals into a single value, negating relational information between covariates. Additionally, they consider the data at a single time point and cannot consider previous clinical states or trajectories. While assessments are repeated at multiple time points, current EWS systems treat each observation independently. We propose that multi-variate time-series modeling techniques could address these limitations and improve the prediction of impending death.

Methods.

We use the MIMIC-IV dataset to demonstrate the utility of trajectory in predicting death [4]. We trained an LSTM model to learn trajectories from the previous 48 hours of physiological observations, restricting the model to covariates available within NEWS2 and eCART. We compare the discriminative ability of the LSTM model against NEWS2 and eCART to identify patients at risk of impending death using the metrics recall, F1-score, AUROC & AUPRC. As part of the model evaluation, we conducted 10-fold stratified cross-validation and qualitative analysis on the sensitivity of the LSTM model EHR data imputation patterns.

Results.

We present the mean values with 95% confidence intervals of the models NEWS2, eCART, and LSTM, respectively. The LSTM model has statistically significant performance advantages in F1-score (0.3162(±0.0070), 0.2687(±0.0043), **0.3391** (±0.0093)), AUROC (0.6423 (±0.0073), 0.6069 (±0.0074), **0.7399** (±0.0119)), and AUPRC (0.1783 (±0.0041), 0.1542 (±0.0026), **0.337** (±0.0212)). eCART has the highest recall, followed by the LSTM (0.5811 (±0.0135), **0.8197** (±0.0108), 0.6964 (±0.0235)). This suggests learning the multi-variate trajectories of physiological signals improves the prediction of impending death. In an EWS system, the clinical benefit of such a model would allow clinicians to correctly identify more at-risk patients without increasing the false alarm rate within the MIMIC-IV population. Time-series data may let us consider the patterns of EHR data inputs, which provide a further context of the patient.

Conclusion.

Using a model architecture that considers physiological data sequentially, we have found that physiological trajectories give further context and improve prediction. Our results demonstrate that using a multi-variate time-series approach using only routinely collected physiological data significantly improves the prediction of death compared to the current EWS systems in widespread clinical use. This system could be implemented as part of an EWS to highlight patients at risk of impending death. We have demonstrated the utility of time-series modeling techniques that allow us to consider patient trajectory through physiological signals and the data pattern within an EHR. These can improve the prediction of impending death and may offer further insights for other downstream prediction tasks. We propose that the development of any future EWS system should incorporate physiological time-series.

References

1. National Institute for Health and Care Excellence. *Acutely Ill Adults in Hospital: Recognising and Responding to Deterioration. NICE Clinical Guideline CG50*.; 2007
2. Churpek MM, Yuen TC, Winslow C, et al. Multicenter development and validation of a risk stratification tool for ward patients. *Am J Respir Crit Care Med*. 2014;190(6):649-655. doi:10.1164/RCCM.201406-1022OC/SUPPL_FILE/DISCLOSURES.PDF
3. Gerry S, Bonnici T, Birks J, et al. Early warning scores for detecting deterioration in adult hospital patients: Systematic review and critical appraisal of methodology. *The BMJ*. 2020;369(3). doi:10.1136/BMJ.M1501
4. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Roger M. MIMIC-IV. Published online 2022.