Temporal Patterns of Primary Care Utilization as Predictors for ICU Admission

Benjamin Post[1,2], Stephen Brett[1,3], Aldo Faisal [1,2,4]
[1] UKRI Center in AI for Healthcare [2] Dep of Bioengineering & Dept. of Computing, [3] Dept of Surgery & Cancer, Imperial College London; [4] Institute of Artificial & Human Intelligence, University of Bayreuth

Background. Primary care electronic healthcare records have been widely used to predict hospital admissions, where most work has focused on using features such as patient demographics, medical history and physiological parameters (Hippisley-Cox 2013, Donnan 2008, Rahimian 2018). However, these data are sparse by nature, as they can only be collected when patients interact with primary care services. Here we establish a novel feature, namely the temporal *pattern* of healthcare access itself. We illustrate this with an exemplary analysis on temporal *patterns* of primary care utilization prior to an intensive care unit (ICU) admission on data from the country of Wales.

Methods. We performed a retrospective observational study using the Secure Anonymised Information Linkage (SAIL) Databank (Ford 2009, Jones 2014, Lyons 2009), which contains billions of person-based health records and covers over 80% of the population of the country of Wales. We identified adult patients with an unplanned ICU admission in the year 2018 with at least 10 primary care interactions in the preceding 2 years. The timings of the primary care events (General Practice [GP] visits, lab tests, etc) prior to an ICU admission were extracted. An equally sized control group (*not* admitted to ICU in 2018) was sampled with matching age and gender and ten sequential primary care events were extracted at random. The date differences between sequential visits were calculated and for each patient the interaction interval sequence (e.g. 5 days, 33 days, 26 days,…) was computed. The structures of these sequences were then analyzed using clustering machine learning methods to interpret the intervisit patterns. Classifiers were then built to assess the power of these patterns to predict future ICU admission. This study was approved by the SAIL independent Information Governance Review Panel (IGRP) (ref 1323).

Results. 1.69 million patients met inclusion criteria for sampling, with a population ICU admission rate of 0.325% in our study window. 54.3% of patients admitted to ICU were male (median age 65 yrs [IQR 23]). 4986 ICU admissions met final inclusion criteria. We found that ICU admissions had characteristic temporal patterns of pre-admission primary care use and were robust (e.g. across variation in cluster numbers) and distinct from those of our control group. Typically, the diversity of over 85% of patients was captured by only *3* temporal interaction clusters. An example analysis is shown below in Figure 1. In general we found a single cluster - that explained over 60% of all ICU admitted patients - consisted of a roughly weekly periodic pattern of primary care use. Other relevant clusters reflected longer periodic patterns (e.g. 2 weekly, monthly) and contracting temporal intervals (possibly reflecting escalating healthcare needs). These temporal patterns outperformed established predictors, when incorporated into classifiers (logistic regression, gradient boosting classifier and neural network) for the prediction of ICU admission. We found that e.g. age & gender together have an AUC of around 0.49, however, using the temporal pattern alone (ignoring other features), we achieve a substantially higher prediction performance (AUC 0.65). The best classifier (AUC 0.72) was trained by combining the temporal clusters with 6 established features (age, sex, ethnicity, Welsh Index of Multiple Deprivation (WIMD), total chronic medical diagnoses, and total current prescriptions).

Conclusion. We showed that prediction of critical deterioration from sparse primary care data can be enriched by using temporal patterns of healthcare access. We found characteristic patterns by using interpretable methods for analysis and showed how data accessible at the administrative level can be rendered useful for clinical prediction.