Mortality prediction using machine learning on laboratory results of medical ED patients

Start aspect

Prediction of mortality risk is one of the most important yet difficult tasks for a doctor who works at the emergency department (ED). Currently, even the most commonly used tools that help the doctor, ancillary investigations and risk models, are insufficient to reliably assess the individual mortality risk. Reliable prediction can help tailor the management of the patient, to intensify treatment and monitoring or to de-escalate and choose palliative treatment, when appropriate.
In this study, we will evaluate the clinical impact of a newly developed prediction model, based on machine learning (ML). The model has been built in our ED and uses laboratory values that have been ordered within the first two hours of the ED stay. The model does not require additional information and is therefore feasible for use in the ED. The discriminatory value is very high with an AUC in predicting 31-day mortality of 0.94 (95%CI: 0.94-0.95). It is currently unclear how ML based models with such high discriminatory value influence clinical practice. This current study is designed to investigate the additional value for clinical practice of this ML based model. If this model is successful in our ED, a multicentre study will be designed and executed.

Background/relevance

One of the most important challenges a doctor who works at the ED has to face is to assess the risk of mortality of the patients. Based on this assessment, many important decisions are made, such as admitting a patient to the intensive care unit (ICU) when a patient is considered highly at risk of dying.

There are several ways of assessing mortality risk. Intuition and laboratory values, such as lactate, are rather reliable tools to differentiate patients who will live from patients who will die. Clinical risk scores, like CURB-65, which is developed for community acqured pneumonia, or APACHE II, which is developed for ICU patients, are examples of disease specific and general risk scores, respectively. On average, commonly used tests and scores have fair to good discriminatory value with AUCs ranging between 0.65 and 0.85, which suffices when applied to groups of patients. However, for the individual ED patient, the discriminatory value is insufficient to guide important decisions.

Currently, based on ML techniques, new models are being developed to predict mortality more precisely. These techniques learn from patterns in data and are capable of integrating all available data of an individual patient. The main advantage of ML is that it is unbiased and has the ability to be updated when new information becomes available.

Recently, using historical laboratory data of over 60,000 samples of our ED patients, we developed a model that yielded an AUC in predicting 31-day mortality of 0.944 (95%CI: 0.935-0.951). All laboratory tests that are taken within the first two hours of the ED visit were included in this ML model. This highly accurate mortality prediction raises the question how this information influences clinical practice in the ED.

Therefore, we set up this prospective study in the ED, in which we aim to investigate the clinical impact of real-time and accurate mortality prediction. We will use ML to categorize the ED patients into 3 risk groups (low, intermediate, high risk) and we will analyze differences in patient characteristics (age, comorbidity, vital signs, diagnoses, commonly used risk scores), treatment decisions (policy restrictions, (ICU) admission) and outcome (length of hospital stay, 31-day mortality) per category. We will further compare ML prediction with the intuition of the ED doctor and investigate how often the doctor is surprised and changes the treatment plan after seeing the ML prediction.

Hypothesis: Mortality prediction based on ML outperforms the judgement of the ED doctor and accurately categorizes internal medicine ED patients in low, intermediate and high risk. ML will surprise ED doctors and change their treatment plan often.