Study Plan

Author: Arman Norouzi

2020-09-07, Arman.Norouzi@stud.ki.se

A Segmented Approach to Predicting Prediction Model Performance After Transfer Using Unlabelled Data

Introduction

Prediction models can be defined to be a statistical method that utilizes multiple predictor data which may be used to estimate a risk of a specific outcome in an individual. The outcome that is estimated could be a disease or a condition, referred to as a diagnostic model, or a specific event that will occur in the future, referred to as a prognostic model [1].

The implementation of prediction models within the field of medicine are strongly linked to the movement towards stratified medicine. Stratified medicine aims to aid the health care workers with their decision making depending on each individual's own profile of prediction factors [2]. The use of diagnostic models can help clinicians with reassuring the patients that their symptoms are not due to a serious disease, referral of patients to further testing or initiation of treatment. Prognostic models can help clinicians to help change patient lifestyles or with therapeutic decisions within a given period [3][4].

One of the more well known clinical prediction models is the Framingham Risk Score. The prediction model estimates the absolute risk of developing a cardiovascular disease within 10-years based on multiple patient predictor data [5]. To implement such a prediction model to a new setting that is different from where the model was developed, one has to perform an external validation of the model. This is due to the performance drop in internally validated prediction models when applied to a new set of individuals [6]. To validate a model that is

transferred to another setting, one has to gather both predictor and outcome data from the new setting.

Currently there are no methods for external validation without the use of outcome data when transferring a prediction model to another setting. This represents a substantial knowledge gap. Therefore, the aim of this study is to develop a method for predicting prediction model performance after transferring to a new setting while only utilizing predictor data.

Methods

Study design

The design of the study that we will conduct is a analytical registry study based on a public database made available by the retrospective analysis of a multinational observational study, "Combination of the National Early Warning Score (NEWS) and inflammatory biomarkers for early risk stratification in emergency department patients: results of a multinational, observational study" [7].

Participants

The patients enrolled in the database were all seeking emergency department (ED) care between March 2013 and October 2014 within three tertiary care centers located in USA (Clearwater Hospital), France (Hôspital de la Salpêtrière) and Switzerland (Kontonsspital Aarau). Inclusion criteria for participating in the registry was that an initial blood sample was taken. Paediatric and surgical patients were excluded.

Sample Size

The total number of participants available in the public database is 1303 which is the sample size arrived at in this study. In the original multinational observational study there were a total of 7132 participants registered. However only 1303 of them had fully registered NEWS parameters which is the reason for the sample size.

Variables

Model predictor

The predictors used in our study to estimate the outcome are respiratory rate (RR), con-

fusion, peripheral oxygen saturation (SpO2), systolic blood pressure (BPS), diastolic blood pressure (DPS), heart rate (HR), temperature (temp), midregional-proadrenomedullin (MR-proADM) and procalcitonin (PCT). Sex (male or female), discharge location, length of stay (LOS) and age were also registered.The methods used to measure predictor data such as HR, BPS, BPD, RR, Temp were not mentioned in the original study.

Model outcome

The primary outcome in our study is death within 30 days. The secondary outcome is whether or not the patients are referred to the intensive care unit (ICU) during their time in the hospital.

Missing data

To account for missing data we are using multivariate imputation by chained equations which is implemented in the R package mice [8]. The missing percentage of data will be equal to the number of imputations made. Qualitative variables and quantitative variables are imputed using logistic regression and predictive mean matching respectively.

Statistical methods

Dataset

In our study the dataset will consist of the publicly available register that has been previously mentioned.

Development sample and validation sample

The dataset will be split into a development sample and a validation sample with a test-train split function implemented in R. The development sample will represent the prediction model development setting, and the validation sample the transfer setting. To obtain stable coefficient estimates, the amount of events (where the outcome occurs) in the development setting is set to (x*10) were x is the amount of predictor parameters [9]. For the validation sample a minimum of 100 events will be included and 100 or more non-events [10].

Sequence of analysis

Analysis in this study is all performed in the programming language R [11]. The sequence

of analysis that will be performed to evaluate our method is model development, model validation, method testing and model comparison.

Model development

During model development a prediction model is developed by using the development sample to fit the model and test it. The accuracy of this model is calculated using logistic regression as implemented in R function glm. To compensate for the risk of overfitting the model a bootstrap procedure is used to estimate the linear shrinkage factor [12]. The shrinkage factor is then implemented to the prediction model coefficients .

Model validation

To examine the accuracy of the model in another setting the prediction model that has been developed in the previous stage will be used to estimate the outcome within the validation sample.

Method testing models

To evaluate the effect of our various methods of increasing model performance based solely on predictor data, we will calculate the accuracy of them predicting outcome in the validation sample. The methods that we will evaluate are:

Stacked propensity matching (13) other

Model comparison

Finally, during this step, the difference in accuracy in the development model and validation model is calculated, while the same is done between the development model and the different method testing models. The differences are then compared to each other to evaluate whether the method testing models increase performance in estimating outcome in the transfer setting.

Ethical considerations

Principle of autonomy

The data used in the original observational study has been approved by the institutional review boards of the three hospitals and has since been published for general use. Therefore the respect of autonomy is upheld.

Principle of beneficence

This study will attempt to act in the best interest of future analytical research and indirectly patients, in an attempt to find new ways of increasing external validity in prediction models based on only predictor data.

Principle of nonmaleficence

The database used in this study has already been depersonalized and publicized. Therefore no ways of identification leakage is possible. All analytical data Due to the data already being depersonalized and publicized there is no way for patient identification to be leaked. All analytical data collected in the study is not meant to harm intentionally or unintentionally.

Principle of Justice

Because this study is analytical, the principle of justice does not prevail. The data in the study will be treated equally.

Ethical permit

Due to this study being an analytical study based on a public database the need for an ethical permit is not required.

References

1. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod) the tripod statement. Circulation. 2015;131:211–9.

2. Steyerberg EW, Moons KG, Windt DA van der, Hayden JA, Perel P, Schroter S, et al. Prognosis research strategy (progress) 3: Prognostic model research. PLoS Med. 2013;10:e1001381.

3. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: What, why, and how? Bmj. 2009;338:b375.

4. Steyerberg EW. Clinical prediction models: A practical approach to development, validation, and updating. Springer; 2009.

5. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. American heart journal. 1991;121:293–8.

6. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98:691–8.

7. Eckart A, Hauser SI, Kutz A, Haubitz S, Hausfater P, Amin D, et al. Combination of the national early warning score (news) and inflammatory biomarkers for early risk stratification in emergency department patients: Results of a multinational, observational study. BMJ open. 2019;9:e024636.

8. Buuren S van, Groothuis-Oudshoorn K. Mice: Multivariate imputation by chained equations in r. Journal of statistical software. 2010;1–68.

9. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. Journal of clinical epidemiology. 1996;49:1373–9.

10. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. Journal of clinical epidemiology. 2005;58:475–83.

11. R-foundation. The r project for statistical computing [internet]. 2018 [cited 2020 2020-09-08]. Available from: Https://www.r-project.org/.

12. Steyerberg E, Eijkemans M, Habbema J. Application of shrinkage techniques in logistic regression analysis: A case study. Statistica Neerlandica. 2001;55:76–88.