German Medical Data Sciences: A Learning Healthcare System
U. Hübner et al. (Eds.)

doi:10.3233/978-1-61499-896-9-170

Prediction of Readmissions in the German DRG System Based on §21 Datasets

Alphons EGGERTH[a,b,1], Dieter HAYN[a], Sai VEERANKI[a,b], Jörg STIEG[c] and Günter SCHREIER[a,b]
[a] *AIT Austrian Institute of Technology*
[b] *Graz University of Technology*
[c] *Agfa HealthCare*

Abstract. Hospital readmissions receive increasing interest, since they are burdensome for patients and costly for healthcare providers. For the calculation of reimbursement fees, in Germany there is the German-Diagnosis Related Groups (G-DRG) system. For every hospital stay, data are collected as a so-called "case", as the basis for the subsequent reimbursement calculations ("§21 dataset"). Merging rules lead to a loss of information in §21 datasets. We applied machine learning to §21 datasets and evaluated the influence of case merging for the resulting accuracy of readmission risk prediction. Data from 478,966 cases were analysed by applying a random forest. Many cases with readmissions within 30 days had been merged and thus their prediction required additional data. Using 10-fold cross validation, the prediction for readmissions within 31 – 60 days showed no notable difference in the area under the ROC curves comparing unedited §21 datasets with §21 datasets with restored original cases. The achieved AUC values of 0.69 lie in a similar range as the values of comparable state-of-the-art models. We conclude that dealing with merged cases, i.e. adding data, is required for 30-day-readmission prediction, whereas un-merging brings no improvement for the readmission prediction of period beyond 30 days.

Keywords. Patient readmission, decision support techniques, diagnosis-related groups, hospital information systems, machine learning

1. Introduction

Readmissions, shortly after hospital discharge, are a common problem in most healthcare systems. Beside unpleasant circumstances for the patients and their families, high numbers of unplanned readmissions also come along with high costs for the healthcare system [1]. In USA, UK, Canada and New South Wales (Australia), the rate of 30-day unplanned readmissions lies between 7% and 16%, whereof many could probably be avoided [2].

In a systematic review from 2016, Zhou et al. [1] give an overview of models for 30-day readmission prediction that were published in 2011 – 2015. The models had an area under the receiver operating characteristic curve (AUROC) from 0.55 to 0.88. Many of them were risk score models, which had been derived from the results of machine learning models and allowed an estimation of the readmission risk by hand [1].

[1] Corresponding Author, Alphons Eggerth, AIT Austrian Institute of Technology, Reininghausstraße 13, 8020 Graz, Austria; E-mail: alphons.eggerth@ait.ac.at.

In a recent approach from 2018, Maali et al. [2] applied gradient boosted tree algorithms to data from New South Wales and created a risk score model, which had an AUROC of 0.71 for 30-day unplanned readmissions. They concluded, that their model was comparable to recent risk score models, which showed AUROCs from 0.68 to 0.75. Although Electronic Health Record (EHR) data were considered, most of the mentioned models relied on further data sources.

The German-Diagnosis Related Groups (G-DRG) system, used for structured reimbursement of comparable treatments, is supervised by the institute for reimbursement in hospitals (InEK) [3]. G-DRG codes are generated based on diagnoses (ICD-10-GM) and procedures (OPS codes) by a grouper software. Based on these data, the resulting reimbursement fees for each hospital stay are determined. By law, every year at the end of March, each German hospital has to deliver a dataset, which contains the calculated G-DRG reimbursement fees together with the underlying data, designated §21 datasets in the following. The contents of this dataset are standardized [3-5]. During G-DRG grouping, subsequent cases are merged in the following situations, leading to reduced reimbursement fees [6]:

- §2-1 FPV: If a patient is readmitted within the upper length-of-stay threshold within the same basis DRG (initial three characters of the DRG code).
- §2-2 FPV: If a patient is readmitted within 30 days within the same major diagnostic category (initial character of the DRG code), if the initial case fee was from the medical or other partition and the subsequent case fee is from the surgical partition.
- §2-3 FPV: If a patient is readmitted within the upper length-of-stay threshold because of hospital-caused complications.
- §3 FPV comprises additional rules for case merges with involved hospital transfers.

For merged cases in §21 datasets, the length of stay and the assignment of procedures, diagnoses, etc. to the separate cases gets lost. The aim of the present paper was to evaluate the influence of restoring merged original cases in §21 datasets for the accuracy of readmission predictions.

2. Methods

Our §21 datasets comprised 478,966 cases recorded in the time from 2013-2017 at five independent hospitals in Germany, including one university hospital. For each hospital, five files were provided (see Table 1). Data protection was assured through contracts between the participating hospitals and Agfa HealthCare GmbH on the one hand side, and on the other hand side by a corresponding contract between Agfa HealthCare GmbH and AIT, which received de-identified data in the framework of their collaboration in the funded HIS-PREMO project (see acknowledgment).

For the analysis of 31-60-day readmissions, §21 datasets and §21 datasets with restored original cases were used. To restore the merged original cases, we received additional information. However, the additional information did not contain all the fields of the §21 dataset. In table 1, the fields of the csv files were grouped according to their availability for the underlying original cases. Fields like the “year of birth” remained unaffected and could be copied for all separate original cases. Features, which varied for each separate original case, like the “date of admission” could not be restored and had to be extracted from the hospital information systems (HIS) as additional information. We rejected some fields of the §21 dataset for our analyses, as no additional information was given to restore them.

We used our Predictive Analytics Toolset for Healthcare (PATH) for data pre-processing, feature engineering, modelling and evaluation. PATH is a Matlab (The MathWorks, Nattick, US) based toolset, which has already been applied to various predictive modelling scenarios (see [7, 8]). PATH allowed aggregating the data to one row for each admission by various grouping and linking functions.

Time-based features were calculated with the given dates (e.g. length-of-stay = date of discharge minus date of admission). Some features were extracted from particular digits of encoded fields. Hospital specific encodings of the additional files were converted. Codes, which comprised information that would be unavailable in a prospective situation were not considered. Charlson's comorbidity score [9] was calculated from ICD codes. ICD and OPS codes were disassembled to obtain the underlying classification hierarchy. Only inpatient stays were considered.

For the machine learning task, the "fitensemble" Matlab function was used to apply a bagged ensemble of 25 regression trees. [10]

Table 1. Overview of the fields of the various csv files, which were available for separate original cases. Keys, targets and constants are crossed out. For the separation of case merges "unaffected" features could be copied, "HIS required" features needed additional data from the HIS and for "rejected" features, no additional data from the HIS was available. ENTGELTE.csv files were available too, but not considered.

Name	Content	unaffected	HIS required	rejected (no HIS data)
FALL	cases	hospital (ID), ~~fee category~~, ~~pseudonymised case ID~~, insurance (ID), year of birth, month of birth (if age < 1 year), sex, ~~merged case (yes/no)~~, ~~reason of case merge~~, age in years at admission, ~~pseudonymised patient ID~~	date of admission, reason of admission, date of discharge, reason of discharge, age in days at admission (if age < 1 year), hours of mechanical ventilation	type of admission, admission weight (if age < 1 year), number of intermittent dialyses, starting date of preadmission treatment, number of preadmission treatment days, end date of post-discharge treatment, number of post-discharge treatment days, transfer hospital (ID)
ICD	diagnoses		type of diagnosis, ICD version, ICD code, localisation, secondary code, localisation for secondary code	
OPS	procedures		OPS version, OPS code, localisation, OPS date	
FAB	hospital departments		hospital department, hospital department admission date, hospital department discharge date	

3. Results

The analysis of the time to readmission for §21 datasets and §21 datasets with restored original cases showed a significant number of merged cases (13%) within the first 30 days (see Figure 1). For readmissions within 31 – 60 days, very few original cases were merged and thus, this time span was chosen as the basis for our further analysis.

For 31-60-day readmissions, we compared the prediction accuracy of a model based on §21 datasets and a model based on §21 datasets with restored original cases. Both models were based on "unaffected" and "HIS required" features only (see Table 1). For both analyses, the AUROC was 0.69 (see Figure 2) and at the chosen cut-off, the sensitivity was 0.64 and the specificity was 0.64.

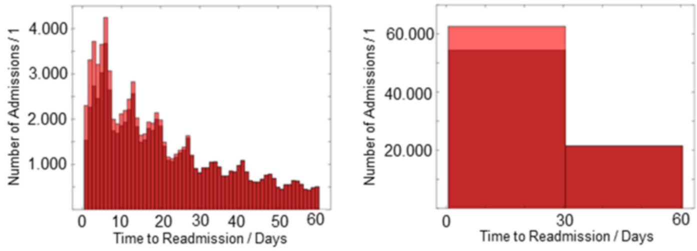


Figure 1. Comparison of days to readmissions after discharge for §21 datasets (dark red) and §21 datasets with restored original cases (light red). *Left*: bin size of 1 day, *Right*: bin size of 30 days, only very few case merges lie within 31 – 60 days to readmission.

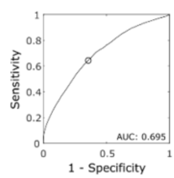
4. Discussion

Our predictive models have been derived from retrospective training datasets. In the future, they could be applied in a prospective way to unseen data to predict future outcomes. For the learning phase, case merging needed to be reversed to re-construct all separate hospital stays and, thus, obtain a comprehensive training set from §21 datasets. In clinical routine use, predictions would rely on information, collected into a data warehouse in near real-time from the hospital information system and other related source systems.

Merged cases did not contain all details of the underlying cases (admission and discharge dates; assignment of diagnoses, procedures, etc. to separate cases). We found that most of the case merges of §21 datasets were done within 30 days after discharge, although they can occur within up to 144 days. Within 31-60 days after discharge, only very few cases were merged and thus, this time window was chosen for the comparison of the prediction accuracy between models based on §21 datasets and based on §21 datasets with restored original cases.

a)

	Truth (-)	Truth (+)
Prediction (-)	221,196	6,818
Prediction (+)	122,905	12,291



b)

	Truth (-)	Truth (+)
Prediction (-)	224,968	6,870
Prediction (+)	126,531	12,251

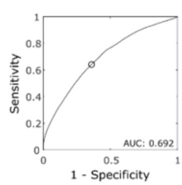


Figure 2. Results of the 31-60 day readmission prediction analyses: a) for §21 datasets, b) for §21 datasets with restored original cases, *left:* Confusion Matrix, (+) … readmission within 31-60 days, (-) … no readmission within 31-60 days, *right:* ROC curve with marked cut-off (circle). Both analyses resulted in a Sensitivity = 0.64, a Specificity = 0.64 and an AUROC = 0.69.

Our analysis showed, that for short-term hospital readmissions within 30 days, the restoration of original cases is required, whereas no improvement of the prediction accuracy could be observed for 31-60-day readmissions. Our resulting AUROC of 0.69 was comparable to recent state-of-the-art readmission prediction models.

Once our model has been validated, it will be a valuable tool, which could be used in real-time at the point of discharge to decide, whether a patient is ready to leave the hospital whether additional treatment/monitoring may be necessary or whether additional measures (rehab, etc.) are indicated.

5. Conclusion

Case merges in §21 datasets limit their utility for research in general. Our approach shows that this is the case in particular for 30-day-readmission predictions, whereas they do not significantly impact the readmission prediction within 31 – 60 days.

6. Conflict of Interest

To the authors' knowledge, no conflicts of interest were given.

7. Acknowledgement

This work was partly funded by the Austrian Research Promotion agency (www.ffg.at – B1: "Hospital Information Systems – Predictive Modeling", HIS-PREMO, 853264)

References

[1] H. Zhou, P.R. Della, P. Roberts, et al, Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review., *BMJ open* **6** (2016).
[2] Y. Maali, O. Perez-Concha, E. Coiera, et al., Predicting 7-day, 30-day and 60-day all-cause unplanned readmission: a case study of a Sydney hospital., *BMC medical informatics and decision making* **18** (2018).
[3] IneK, http://www.g-drg.de/, Access: 2018-02-06
[4] DIMDI, https://www.dimdi.de/static/de/klassi/icd-10-gm/anwendung/zweck/g-drg/, Access: 2018-02-06
[5] R. Busse, A. Geissler, A. Aaviksoo, et al., *Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals*, McGraw-Hill Education, UK, 2011.
[6] Vereinbarung zum Fallpauschalensystem für Krankenhäuser für das Jahr 2017 (FPV 2017).
[7] Y. Xie, G. Schreier, M. Hoy, et al., Analyzing health insurance claims on different timescales to predict days in hospital., *J Biomed Inform* **60** (2016).
[8] D. Hayn, K. Kreiner, H. Ebner, et al., Development of Multivariable Models to Predict and Benchmark Transfusion in Elective Surgery Supporting Patient Blood Management, *Applied Clinical Informatics* **8** (2017), 617-631.
[9] H. Quan, B. Li, C.M. Couris, et al., Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries., *American journal of epidemiology* **173** (2011), 676-682.
[10] MathWorks, https://de.mathworks.com/help/stats/ensemble-algorithms.html, Access: 2018-02-07