

Unplanned 30-Day Readmissions: Hospital Data Warehouse Modelling.

Thibault Dhalluin (✉ thibault.dhalluin@univ-tours.fr)

Department of Medical Information, University Hospital of Tours, Tours, France. Medical school, University of Tours, EA 7505 EES, Tours France.

Guillaume Bouzillé

Univ Rennes, CHU Rennes Inserm, LTSI – UMR 1099, F-35000 Rennes, France

Aurélie Bannay

CHRU-Nancy, Université de Lorraine, Service d'évaluation et d'information médicales, F54000, France

Pierre Lemordant

Univ Rennes, CHU Rennes Inserm, LTSI – UMR 1099, F-35000 Rennes, France

Emmanuel Sylvestre

Univ Rennes, CHU Rennes Inserm, LTSI – UMR 1099, F-35000 Rennes, France

Leslie Grammatico-Guillon

Department of Medical Information, University Hospital of Tours, Tours, France. Medical school, University of Tours, EA 7505 EES, Tours France.

Marc Cuggia

Univ Rennes, CHU Rennes Inserm, LTSI – UMR 1099, F-35000 Rennes, France

Research Article

Keywords: Medical Informatics, Data Warehousing, Supervised Machine Learning, Patient Readmission/statistic and numerical data

Posted Date: March 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-274754/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Unplanned hospital readmissions are a major healthcare and economic burden. This study compared statistical methods and machine learning algorithms for predicting the risk of all-cause 30-day hospital readmission in two French academic hospitals.

Methods: The dataset included hospital stays selected from the clinical data warehouses (CDW) of the two hospitals (Rennes and Tours Academic Hospitals) using the criteria of the French national methodology to measure the 30-day readmission rate (i.e. ≥ 18 -year-old patients, geolocation, no iterative stays, and no hospitalization for palliative care). Then, the prediction performance of Logistic Regression, Naive Bayes, Gradient Boosting, Random Forest, and Neural Networks were compared separately for the two hospitals but using the same CDW data pre-processing for all algorithms. The area under the receiver operating characteristic curve (AUC) was calculated for the 30-day readmission prediction performance of each model as well as the time to train the algorithm.

Results: In total, 259,092 and 197,815 stays were included from the Rennes and Tours Academic Hospital CDWs, respectively, with readmission rates of 8.8% (Rennes) and 9.5% (Tours). The AUC of the regression models for the two hospitals ranged from 0.61 to 0.64, with computation times exceeding 18 hours. The AUC of the machine learning models ranged from 0.61 to 0.69 with computation times below 13 hours.

Conclusions: Better performance and shorter computation times are obtained with machine learning methods. It is still necessary to compare different algorithms to identify the most efficient model.

Background

The days following hospital discharge are at high risk of adverse events [1, 2], and preventing unexpected readmissions is an important issue. Moreover, unplanned readmissions represent a healthcare and economic burden. In France, the 30-day readmission rate varies between 12% and 14% [3]. In the United States, the 30-day readmission rate is about 20%, depending on the state, and the cost of these readmissions is estimated at 17.4 billion dollars [4].

The days following hospital discharge are a key point in the care pathway [1, 2], and must be articulated between caregivers. To comprehensively address this problem in the context of population ageing and the increasing prevalence of chronic diseases, the care pathway between hospital and primary care has to be taken into account. A systematic review highlighted that some 30-day readmission events are avoidable, but their proportion greatly varies, depending on the judgment criterion (a combination of diagnostic codes, adverse events, adverse drug reactions, and subjective criteria) used to define such avoidability [5].

Public health policies have been implemented to reduce the 30-day readmission rates in many countries, such as France, Germany, Denmark [6], and United States [7]. In France, the Agence Technique de l'Information sur l'Hospitalisation (ATIH; French national agency of medical information) has put in place an indicator of 30-day readmission, called RH30 [8], based on the French version of the Diagnosis-Related Group (F-DRG) system. However, it is also important to develop methods to anticipate unplanned readmissions, by identifying relevant clinical indicators, and by developing predictive models to integrate these indicators. The most widespread predictive models are based on classical statistical approaches that are widely used in medical research, such as logistic regression analysis, to assess potential improvements in medical care [9–11].

The progressive implementation of electronic health records (EHR) in healthcare facilities represents an unprecedented source of data that could be exploited to better apprehend the patient care pathway, both descriptively and analytically. The emergence of automatic learning methods allows using these data to anticipate the patient trajectories. However, to make EHR usable for multicentre studies, an interoperability effort is required to arrange data in clinical data warehouses (CDW) using common terminologies [12, 13]. Recently, machine learning methods have started to be tested (e.g., Gradient Boosting, Random Forest, Neural Networks) [9]. The contribution of each of these methods must be now assessed in similar conditions.

The goal of this study was to compare different approaches to predict the 30-day readmission risk using two CDWs of two French hospitals.

Methods

Study design

Rennes and Tours hospitals are academic hospitals with similar activity and organisation (Table 1). The CDW data model and technical design are the same in these two hospitals [13]. The same algorithms for selection criteria, readmission criteria, data management, and the same prediction algorithms were ran separately using the Rennes and Tours CDW data.

Table 1 – Number of hospital beds in Rennes and Tours academic hospitals in 2020.

Number of hospital beds	Rennes	Tours
Medicine	917	748
Surgery	464	537
Gynaecology/Obstetrics	107	126
Total	1488	1411

Patients with at least one hospital admission date in Rennes or Tours between 1 January 2013 and 31 December 2018 were included in the study. The two datasets were then divided into training set (admission date before 31 December 2016) and test set (admission date after 1 January 2017).

Rennes and Tours CDWs contain most of the EHR data, including clinical notes, drug prescriptions, laboratory test records, and claim data. The different stays are chained by patient in a de-identified EHR. All information about a patient's hospital stay is de-identified and organised in structured and unstructured data in the CDW.

Diagnoses and comorbidities were coded using the French version of the International Classification of Diseases, 10th Edition (ICD-10) [14], and grouped with the F-DRG [15]. Medical and surgical procedures were also coded according to the French classification of clinical procedures (CCAM) [16]. Drug prescriptions were coded using the Anatomical, Therapeutic and Chemical (ATC) codes in the Rennes CDW. Drug data were not yet available in the Tours CDW at the time of the analysis. Hospital departments and laboratory data are currently coded according to a local thesaurus.

Inclusion criteria

To ensure the result comparability with the French national indicator for 30-day readmission (RH-30), the selection criteria were those of the ATIH RH30 national methodology [8]: ≥ 18 -year-old patients who received obstetrical/gynaecological, surgical, or medical care.

Non-inclusion criteria

Patients without geolocation in mainland France and patients in palliative care settings were excluded. Hospital stays with a different entry mode than from home, and iterative stays were also excluded (chemotherapy/radiotherapy sessions, transplant context, renal haemodialysis sessions, cataract surgery). Readmissions within 30 days after an iterative stay were not considered as unplanned readmissions.

Unplanned readmission definition

An unplanned readmission was defined as a hospital stay within 30 days after the end of the index stay. The index stay was defined as a hospital stay with a discharge to home and not preceded by another hospitalisation within 30 days before the index admission date. Stays corresponding to the inclusion and exclusion criteria and with at least one index stay were included. The aim was to predict readmission within 30 days of the index stay.

Data extraction

Covariates extracted from the two CDWs [13] were: age, sex, length of stay, number of previous hospitalisations, illness severity (F-DRG classification), major diagnostic categories, medical diagnoses and comorbidities (ICD-10), medical and/or surgical procedures (CCAM classification), hospital department, and available laboratory data.

ICD-10 and CCAM codes were grouped by the three first characters indicating the diagnostic category and the procedure and organ, respectively. Laboratory data were coded as binary variables, and were considered as abnormal if at least one value during the stay was outside the normal range.

Demographic data, aggregated at the city and district levels when the municipality had more than 10,000 inhabitants [17], were merged with the patient using the geometric map background and the corresponding geolocation:

- Median pre-tax household income per year
- Part of ≥ 15 -year-old population, education level
- Unemployment rate for the 15 to 64-year-old population
- Socio-professional categories for the 15 to 64-year-old population

Data Processing

Missing data were imputed using the K Nearest Neighbor method (taking the mean of the five nearest neighbors); numerical values were considered Missing At Random (MAR). Features with less than 1% of events were removed to avoid unpredictable and inexplicable responses from prediction models related to rare events. For the logistic regression models, only significant covariates (p -value ≤ 0.05 by univariate analysis) were retained and then, a multivariate model with step-by-step selection of variables was performed to obtain the most parsimonious model to maximise the area under the Receiver Operating Characteristic (ROC) curve (AUC). The algorithms used for readmission prediction were the most frequently described in the literature: logistic regression, Random Forest, Gradient Boosting, Naive Bayes, and Neural Networks [9, 18–21].

The main outcome was the AUC, and the secondary outcomes were sensitivity, specificity, positive and negative predictive value of the cut-point closest to the top-left corner of the ROC space. The calculation time was evaluated after the data management step and was from the start of the feature selection to the end of model fitting.

As a secondary objective, the model explainability was assessed by identifying the covariates considered to be important by the different algorithms. This importance was assessed according to criteria adapted to each algorithm: Odds Ratio (OR) for logistic regression, relative influence for Gradient Boosting, Gini index for Random Forest, and Garson algorithm for Neural Networks. The most relevant covariates were compared between hospitals.

Data handling and pre-processing were performed on R Studio, version 3.6.0 [22].

Ethics and Consent

The clinical data warehouses have been authorized by the Commission Nationale de l'Informatique et des Libertés (CNIL national commission for information technology and civil liberties), requiring individual and collective patient information. Patients have a right to oppose the reuse of their data.

Results

At Rennes Academic Hospital, among the 410,833 hospital stays with enough available data, 259,092 (63%) were included (193,906 for the training set, and 65,186 for the test set). The 30-day readmission rate for the included stays was 8.8%. At Tours Academic Hospital, among the 574,306 hospital stays with enough available data, 197,815 (34%) were included (156,204 for the training set, and 41,611 for the test set). The 30-day readmission rate for the included stays was 9.5% (Figure 1).

Comparison of patients with and without unplanned 30-day readmission showed that patients with unexpected readmission were older (+3.8 years in Rennes and +6.9 years in Tours), had longer mean length of stay (+1.4 days in Rennes and +0.9 days in Tours) and higher mean number of previous stays (2.6 and 2.4 in Rennes and Tours, respectively, *versus* 1.8). The socio-demographic data did not differ between patients with and without unplanned 30-day readmission (Table 2).

Table 2 – Patients' characteristics

Hospital	Rennes		Tours	
	Yes	No	Yes	No
n	22,890	236,202	18,824	178,991
Age (mean (SD))	60.30 (19.34)	56.46 (20.67)	66.08 (18.40)	59.22 (20.25)
Sex, Male (%)	13,352 (58.3)	11,6320 (49.2)	10,327 (54.9)	89,323 (49.9)
Length of stay (days) (mean (SD))	6.18 (10.47)	4.81 (8.14)	4.68 (8.75)	3.76 (8.45)
Number of previous stays (mean (SD))	2.58 (3.27)	1.82 (1.85)	2.37 (2.97)	1.82 (1.79)
>15-year-old population with higher education degree (mean (SD))	0.30 (0.13)	0.30 (0.13)	0.25 (0.11)	0.25 (0.11)
>15-year-old population with secondary school diploma (mean (SD))	0.17 (0.03)	0.17 (0.03)	0.16 (0.03)	0.16 (0.03)
>15-year-old population with vocational diploma (mean (SD))	0.25 (0.07)	0.25 (0.07)	0.27 (0.06)	0.27 (0.06)
>15-year-old population without any secondary school diploma (mean (SD))	0.28 (0.09)	0.28 (0.09)	0.32 (0.09)	0.31 (0.09)
15-64-year-old population with managerial occupations (mean (SD))	0.17 (0.11)	0.17 (0.11)	0.14 (0.09)	0.14 (0.09)
15-64-year-old population with craftsman, shopkeeper jobs (mean (SD))	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)
15-64-year-old population with blue collar jobs (mean (SD))	0.23 (0.10)	0.23 (0.10)	0.23 (0.09)	0.23 (0.09)
15-64-year-old population with farm jobs (mean (SD))	0.02 (0.04)	0.02 (0.04)	0.02 (0.04)	0.02 (0.04)
15-64-year-old population without employment (mean (SD))	0.12 (0.06)	0.12 (0.06)	0.13 (0.07)	0.13 (0.07)
Median tax income per household (mean (SD))	20,654.08 (3,274.24)	20,643.80 (3,210.95)	20,355.08 (3,312.09)	20,347.54 (3,275.59)

Analysis of the models' performance for readmission prediction gave AUC values of 0.61 and 0.64 for logistic regression, 0.68 and 0.68 for Gradient Boosting, 0.69 and 0.67 for Random Forest, 0.66 and 0.68 for Naive Bayes, and 0.61 and 0.65 for Neural Networks, using the datasets from Rennes and Tours, respectively (Figure 2). Other performance indicators are summarised in Tables 3 and 4.

Concerning the model explainability, the covariates selected by the stepwise regression models were all different, except for the 'musculoskeletal and connective tissue diseases and injuries' covariate. With Gradient Boosting, among the 20 covariates with the greatest relative influence, seven covariates were shared by both hospitals: age, number of previous stays, ICD-10 R18 ascites, red blood cell count, haemoglobin, urea, and C-reactive protein. The Random Forest model prioritized numerical covariates in both hospitals: age, severity level, length of stay, number of previous stays, and sociodemographic variables. Neural Networks did not have any common covariate between hospitals (additional file 1 shows the most important variables for the different tested models with the CDW data from Rennes and Tours hospitals).

Table 3 - Rennes

Algorithm	TP	TN	FP	FN	Se	Sp	PPV	NPV	AUC	TimeTrain (s)
Naive Bayes	4741	21725	37774	946	0.83	0.37	0.11	0.96	0.66	541.479
Gradient Boosting	3537	37752	21747	2150	0.62	0.63	0.14	0.95	0.68	4368.794
Logistic Regression	1573	50911	8588	4114	0.28	0.86	0.15	0.93	0.61	92223.848
Neural Networks	3080	37327	22172	2607	0.54	0.63	0.12	0.93	0.61	1844.773
Random Forest	3574	38998	20501	2113	0.63	0.66	0.15	0.95	0.69	44758.750
TP True Positive; TN True Negative; FP False Positive; FN False Negative; Se Sensitivity; Sp Specificity; PPV Positive Predictive Value; NPV Negative Predictive Value; AUC Area Under receiver operating characteristic Curve										

Table 4 - Tours

Algorithm	TP	TN	FP	FN	Se	Sp	PPV	NPV	AUC	TimeTrain (s)
Naive Bayes	2286	24526	13450	1349	0.63	0.65	0.15	0.95	0.68	213.514
Gradient Boosting	2475	21601	16375	1160	0.68	0.57	0.13	0.95	0.68	1473.248
Logistic Regression	1888	26499	11477	1747	0.52	0.70	0.14	0.94	0.64	66236.542
Neural Networks	1862	27130	10846	1773	0.51	0.71	0.15	0.94	0.65	690.124
Random Forest	2322	23264	14712	1313	0.64	0.61	0.14	0.95	0.67	28427.632
TP True Positive; TN True Negative; FP False Positive; FN False Negative; Se Sensitivity; Sp Specificity; PPV Positive Predictive Value; NPV Negative Predictive Value; AUC Area Under receiver operating characteristic Curve										

Discussion

This is the first 30-day readmission predictive modelling study performed using data from two CDWs in France. This study was carried out using a methodology as close as possible to the definition of the French national indicator of unplanned 30-day readmission and the explanatory variables suggested by the French national agency of medical information.

The population health of the neighbouring areas, the local health care network, the hospital management, the computerization of services, and the data integration advancement in CDWs are factors that might influence the results of the models. The choice of inclusion criteria and the focus on all-cause 30-day readmissions influenced the performance of the predictive models. The selection of covariates and the algorithm used are end-of-process parameters that depend on the previous steps.

The healthcare organization is very similar at the Rennes and Tours academic hospitals, both in terms of positioning within the local healthcare network and in terms of patient discharge policy, emergency activities, unplanned care rate, complex clinical situations, and university hospital activities providing innovative care. The main difference concerns cancer management that is part of the care activities of Tours hospital, whereas it is performed in a separate cancer centre in Rennes. This difference explains the higher number of excluded stays for palliative care and iterative stays at Tours compared with Rennes. Data integration using the same CDW technology reduces variability. However, the two CDWs are at different stages of development in terms of flow development and data integration, and therefore drug data was not yet available in the Tours CDW. Running the same algorithms with the same inclusion criteria also allowed assessing the consequences of differences in data quality on algorithm tuning.

The readmission rates found for Rennes and Tours academic hospitals were lower than previously reported; however, explaining factors need to be taken into account. First, the selection of stays was carried out according to a methodology that excluded <18-year-old people, rehabilitation stays, iterative stays, and palliative care, thus eliminating stays with often more frequent 30-day readmissions. Moreover, as early readmissions on the first day of discharge were merged with the previous stay due to the claim data rules, it was not possible to identify them in the structured data. This caused an underestimation of the readmission rate and a loss of information for prediction modelling. As the two CDWs contained only data for a specific hospital, it was not possible to identify potentially unexpected readmissions in other hospitals or to identify alternate care facilities [23]. In agreement, the performance (AUC value) of our prediction models was lower than the mean AUC of 0.71 found in the literature [9] (i.e. studies on 30-day hospital readmission prediction models, in peer-reviewed journals and in repositories available on GitHub [24]).

One of the main issues mentioned in previous works is the lack of reproducibility. Our work overcomes this difficulty by performing the same prediction analyses using data from two similar CDWs with the same definition of target population, the same data management, and the same algorithms. Predictive models from studies in the United States often included as covariates the patients' resource conditions, health insurance coverage, ethnicity, and job status, which are available in American health care facilities, and have a significant impact on the probability of readmission. In France, the social context also is a determinant factor for 30-day readmissions prediction, but fewer data are available to assess the patients' social situation. Therefore, aggregated data were included at the level of cities and districts to try to improve the predictive models' performances. However, aggregated data can introduce an ecological bias, due to the lack of individual specification of this information and the correlations between social and geographical inequalities. For both cities, socio-demographic characteristics were not significantly different between readmitted and non-readmitted patients.

The long computation times for the logistic regression models showed the inadequacy of classical statistic methods of stepwise covariate selection, in addition to their limited performance and reproducibility. Conversely, methods adapted to large datasets gave better prediction performances with shorter computation times. The covariates were selected using a data-driven method by choosing the information that improved the model performance. The purpose of this study was not to interpret the association between covariates and unplanned 30-day readmission. ICD-10 and CCAM codes were grouped in hierarchical codes to reduce dimensionality and to have a relevant clinical explicability. The importance of covariates differed among algorithms, for reasons specific to the mathematical logic of each model. For instance, Random Forest placed more importance on numerical features.

The objective of this study was to compare the performance of different algorithms. As the models do not have a specific use yet, there was no reason to focus on sensitivity (to screen for stays at risk of readmission), and specificity (to target specifically stays with readmission). During the development and usage of care indicators, their usefulness and the medical facts interpreted must be constantly questioned. The results obtained with these prediction models could be used to target some patients at

high-risk of unplanned 30-day readmission. However, their current performance is not sufficient for a wider use. Furthermore, explaining the reasons for a potential 30-day readmission understood by the physician is still a challenge for the ease of medical decision making in the current clinical practice [25].

In the future, an improvement in the models' performance or the identification of events that provide a more operational response for decision making at discharge are necessary. In addition, implementing CDWs with a similar data model might allow validating prediction algorithms using data from different hospitals and assessing the replicability of results.

Abbreviations

ATIH Agence Technique de l'Information sur l'Hospitalisation

RH30 National indicator of 30-day readmission

F-DRG French version of the Diagnosis-Related Group

EHR Electronic Health Record

CDW Clinical Data Warehouse

ICD-10 International Classification of Diseases, 10th Edition

CCAM French classification of clinical procedures

ATC Anatomical, Therapeutic and Chemical

ROC Receiver Operating Characteristic

AUC Area Under the ROC Curve

Declarations

Ethics and Consent

The clinical data warehouses have been authorized by the CNIL (Commission Nationale de l'Informatique et des Libertés), requiring individual and collective patient information.

Patients have a right to oppose the reuse of their data.

- Tours Clinical Data Warehouse Registration n° m862699436z
- Tours treatment registration n° 2021_012
- Rennes Clinical Data Warehouse Authorization CNIL le February 27, 2020 (deliberation n°2020-028)

Availability of data and materials statement

The data used for this study are real-life medical data from Rennes and Tours academic hospitals. These data are integrated and de-identified in clinical data warehouses.

The datasets generated and/or analysed during the current study are not publicly available due professional secrecy but are available from the corresponding author on reasonable request.

Competing interests

None declared

Funding

None

Authors' contributions

Thibault Dhalluin and Guillaume Bouzillé - data analytics and wrote the main manuscript.

Aurélié Bannay, Emmanuel Sylvestre, Pierre Lemordant, Leslie Grammatico-Guillon and Marc Cuggia reviewed the manuscript.

References

1. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The Incidence and Severity of Adverse Events Affecting Patients after Discharge from the Hospital. *Annals of Internal Medicine*. 2003;138:161–7.
2. Schnipper JL, Kirwin JL, Cotugno MC, Wahlstrom SA, Brown BA, Tarvin E, et al. Role of Pharmacist Counseling in Preventing Adverse Drug Events After Hospitalization. *Arch Intern Med*. 2006;166:565–71.
3. Gusmano M, Rodwin VG, Weisz D, Cottenet J, Quantin C. A Comparative Analysis of Hospital Readmissions in France and the US. *Journal of Comparative Policy Analysis: Research and Practice*. 2016;18:195–209.
4. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med*. 2009;360:1418–28.
5. van Walraven C, Bennett C, Jennings A, Austin PC, Forster AJ. Proportion of hospital readmissions deemed avoidable: a systematic review. *CMAJ*. 2011;183:E391–402.
6. Kristensen SR, Bech M, Quentin W. A roadmap for comparing readmission policies with application to Denmark, England, Germany and the United States. *Health Policy*. 2015;119:264–73.

7. Hospital Readmissions Reduction Program (HRRP) | CMS.
<https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program>. Accessed 26 Jun 2020.
8. DGOS. Les indicateurs de réhospitalisation et de coordination. Ministère des Solidarités et de la Santé. 2020. <https://solidarites-sante.gouv.fr/soins-et-maladies/qualite-des-soins-et-pratiques/qualite/les-indicateurs/article/les-indicateurs-de-rehospitalisation-et-de-coordination>. Accessed 26 Jun 2020.
9. Artetxe A, Beristain A, Graña M. Predictive models for hospital readmission risk: A systematic review of methods. *Computer Methods and Programs in Biomedicine*. 2018;164:49–64.
10. Zhou H, Della PR, Roberts P, Goh L, Dhaliwal SS. Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review. *BMJ Open*. 2016;6:e011060.
11. Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, et al. Risk Prediction Models for Hospital Readmission: A Systematic Review. *JAMA*. 2011;306:1688–98.
12. Bouzillé G, Sylvestre E, Campillo-Gimenez B, Renault E, Ledieu T, Delamarre D, et al. An Integrated Workflow For Secondary Use of Patient Data for Clinical Research. *Stud Health Technol Inform*. 2015;216:913.
13. Madec J, Bouzillé G, Riou C, van Hille P, Merour C, Artigny M-L, et al. eHOP Clinical Data Warehouse: From a Prototype to the Creation of an Inter-Regional Clinical Data Centers Network. *Studies in health technology and informatics*. 2019;264:1536–7.
14. CIM-10 FR 2018 à usage PMSI | Publication ATIH. <https://www.atih.sante.fr/cim-10-fr-2018-usage-pmsi>. Accessed 13 Jul 2020.
15. Manuel des GHM - Version définitive 2019 | Publication ATIH. <https://www.atih.sante.fr/manuel-des-ghm-version-definitive-2019>. Accessed 13 Jul 2020.
16. CCAM en ligne - Consultation par chapitre. <https://www.ameli.fr/accueil-de-la-ccam/trouver-un-acte/consultation-par-chapitre.php>. Accessed 26 Jun 2020.
17. Géoservices | Accéder au téléchargement des données libres IGN.
<https://geoservices.ign.fr/documentation/diffusion/telechargement-donnees-libres.html>. Accessed 26 Jun 2020.
18. Ripley B, Venables W. nnet: Feed-Forward Neural Networks and Multinomial Log-Linear Models. 2020. <https://CRAN.R-project.org/package=nnet>. Accessed 17 Nov 2020.
19. Cutler F original by LB and A, Wiener R port by AL and M. randomForest: Breiman and Cutler's Random Forests for Classification and Regression. 2018. <https://CRAN.R-project.org/package=randomForest>. Accessed 17 Nov 2020.
20. Greenwell B, Boehmke B, Cunningham J, Developers (<https://github.com/gbm-developers>) GBM. gbm: Generalized Boosted Regression Models. 2020. <https://CRAN.R-project.org/package=gbm>. Accessed 17 Nov 2020.
21. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F, C++-code) C-CC (libsvm, et al. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien.

2020. <https://CRAN.R-project.org/package=e1071>. Accessed 17 Nov 2020.

22. R Core Team, R Foundation for Statistical Computing. R: A Language and Environment for Statistical Computing. 2019. <https://www.r-project.org/>. Accessed 29 Jun 2020.

23. Hameed T, Bukhari S. Predicting 30-days All-cause Hospital Readmissions Considering Discharge-to-alternate-care-facilities. 2020. p. 864–73. <https://www.scitepress.org/Link.aspx?doi=10.5220/0009385608640873>. Accessed 26 Jun 2020.

24. Readmission. GitHub. <https://github.com/search?q=Readmission>. Accessed 29 Jun 2020.

25. Tanzer MW, Heil EM. Why Majority of Readmission Risk Assessment Tools Fail in Practice. In: 2013 IEEE International Conference on Healthcare Informatics. 2013. p. 567–9.

Figures

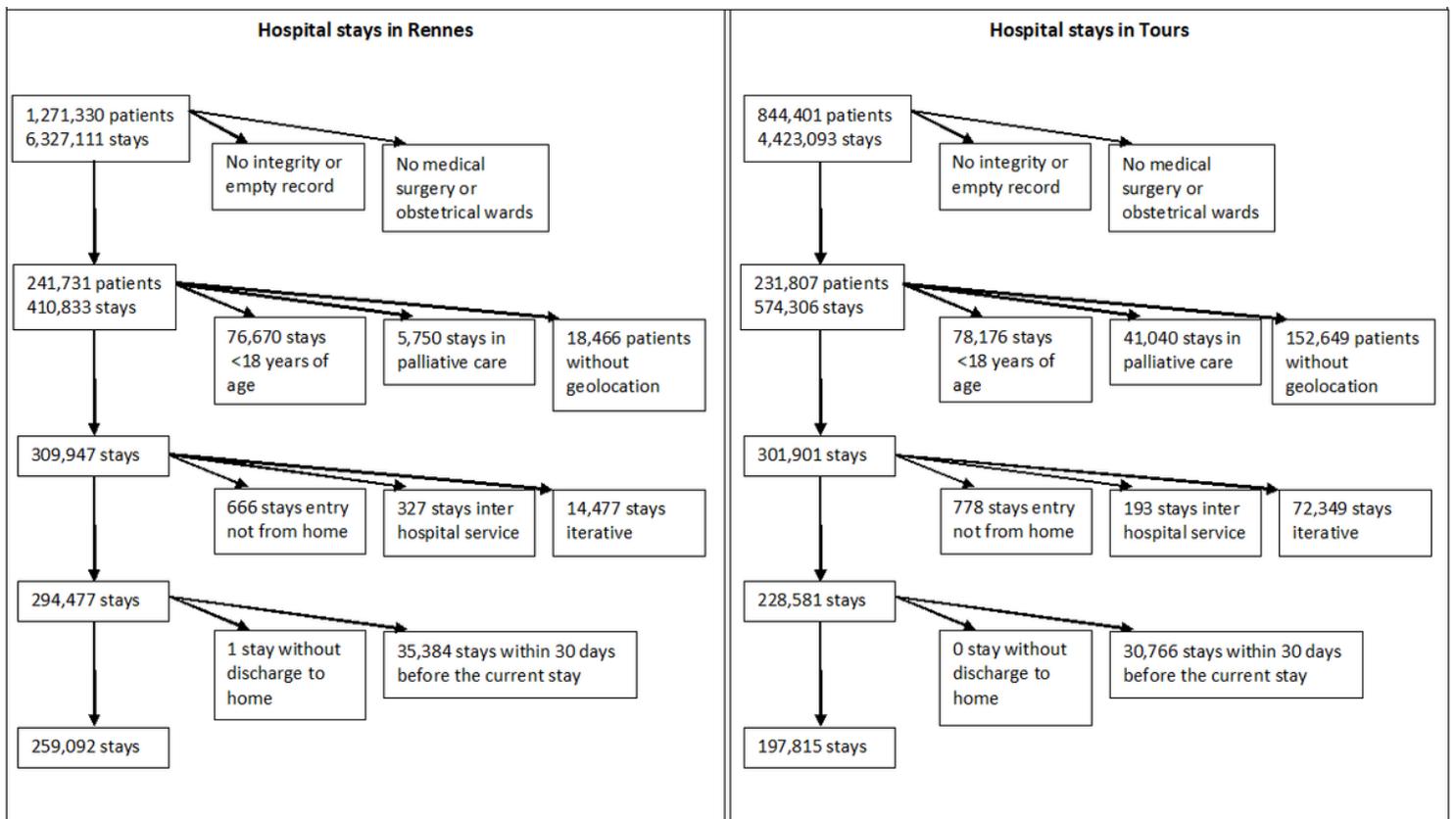


Figure 1

Flow chart of hospital stay selection

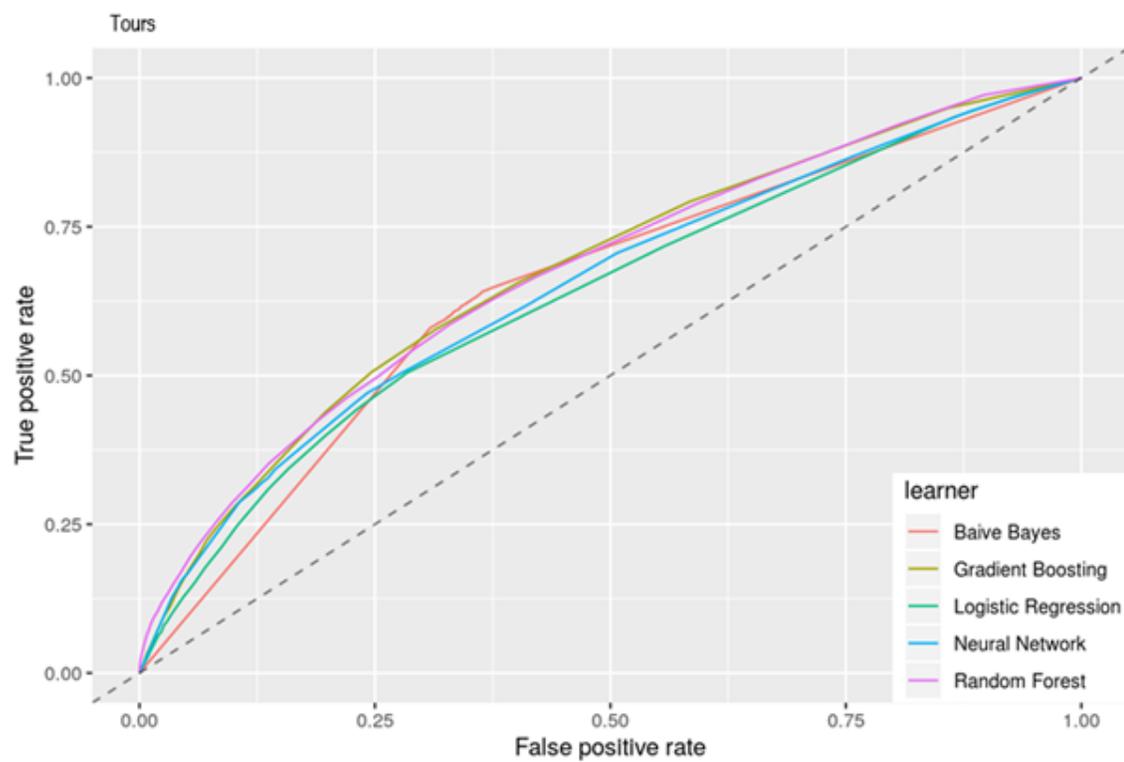
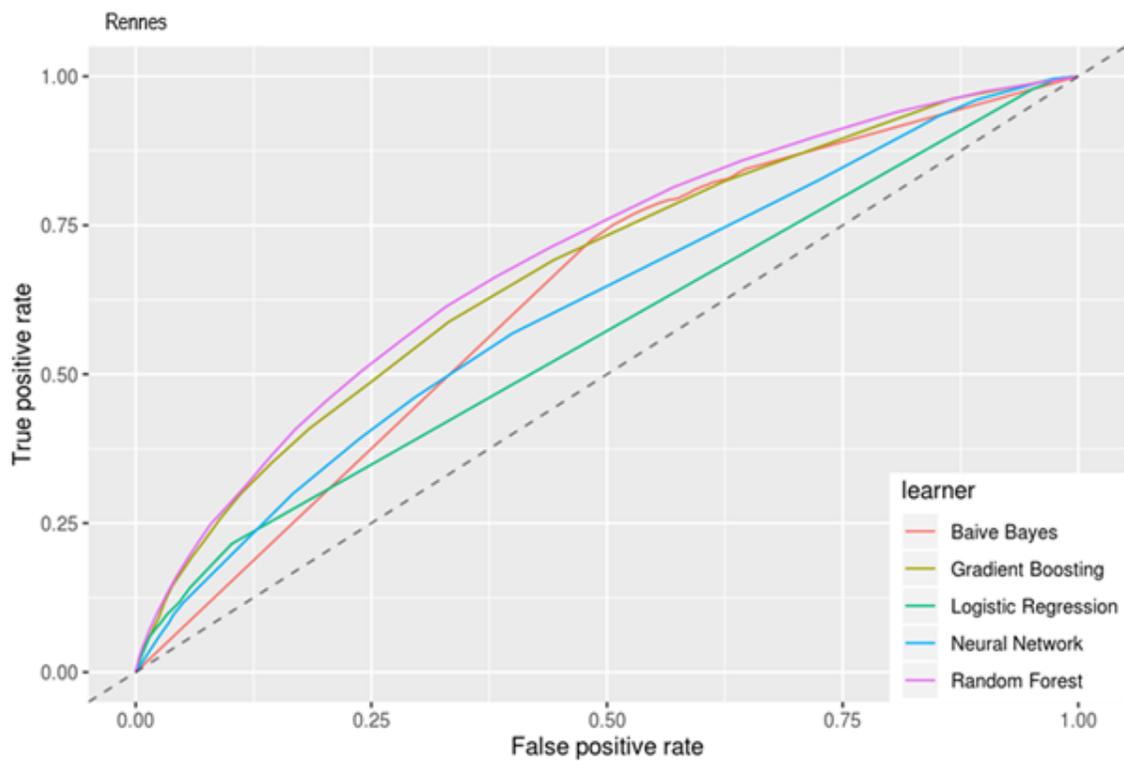


Figure 2

ROC curve of Rennes and Tours prediction models

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Annexe0203.docx](#)