

Machine-learning prediction of unplanned 30-day rehospitalization using the French hospital medico-administrative database

Franck Jaotombo, PhD^{a,b}, Vanessa Pauly, PhD^{a,c}, Pascal Auquier, MD, PhD^a, Veronica Orleans, PhD^c, Mohamed Boucekine, MD, PhD^a, Guillaume Fond, MD, PhD^a, Badih Ghattas, PhD^b, Laurent Boyer, MD, PhD^{a,c,*}

Abstract

Predicting unplanned rehospitalizations has traditionally employed logistic regression models. Machine learning (ML) methods have been introduced in health service research and may improve the prediction of health outcomes. The objective of this work was to develop a ML model to predict 30-day all-cause rehospitalizations based on the French hospital medico-administrative database.

This was a retrospective cohort study of all discharges in the year 2015 from acute-care inpatient hospitalizations in a tertiary-care university center comprising 4 French hospitals. The study endpoint was unplanned 30-day all-cause rehospitalization. Logistic regression (LR), classification and regression trees (CART), random forest (RF), gradient boosting (GB), and neural networks (NN) were applied to the collected data. The predictive performance of the models was evaluated using the H-measure and the area under the ROC curve (AUC).

Our analysis included 118,650 hospitalizations, of which 4127 (3.5%) led to rehospitalizations via emergency departments. The RF model was the most performant model according to the H-measure (0.29) and the AUC (0.79). The performances of the RF, GB and NN models (H-measures ranged from 0.18 to 0.29, AUC ranged from 0.74 to 0.79) were better than those of the LR model (H-measure=0.18, AUC=0.74); all *P* values < .001. In contrast, LR was superior to CART (H-measure=0.16, AUC=0.70), *P* < .0001.

The use of ML may be an alternative to regression models to predict health outcomes. The integration of ML, particularly the RF algorithm, in the prediction of unplanned rehospitalization may help health service providers target patients at high risk of rehospitalizations and propose effective interventions at the hospital level.

Abbreviations: AME = Aide Médicale d'Etat, APHM = Assistance Publique – Hôpitaux de Marseille, AUC = area under the curve, CART = classification and regression trees, CMU = Couverture Maladie Universelle, DT = decision tree, GB = gradient boosting, LOS = length of stay, LR = logistic regression, MDG = mean decrease in Gini, ML = machine learning, NN = neural networks, PMSI = Programme de Médicalisation des Systèmes d'Information, RF = random forest, ROC = receiving operating characteristic, VI = variable importance.

Keywords: health service research, machine learning, patient rehospitalization, prediction

1. Introduction

Reducing 30-day rehospitalizations is a priority of health care policies in Western countries.^[1,2] Unplanned rehospitalizations are common^[3,4] and costly,^[4,5] reflecting poor quality inpatient care,^[6–8] and poorly coordinated transitions between hospitals

and homes.^[9] Despite the growing literature on this issue, unplanned rehospitalizations are still poorly understood and controlled.^[3] We need to better identify patients at high risk of rehospitalization to improve the quality of care and reduce rehospitalizations and associated health care costs.^[10]

Editor: Phil Phan.

The authors have no funding and conflicts of interests to disclose.

Supplemental Digital Content is available for this article.

The data that support the findings of this study are available from a third party, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are available from the authors upon reasonable request and with permission of the third party.

^a Aix-Marseille University, EA 3279 - Public Health, Chronic Diseases and Quality of Life - Research Unit, La Timone Medical University, 27, boulevard Jean-Moulin,

^b Mathematics Institute of Marseille, Aix-Marseille University, Marseille, France, ^c Service d'Information Médicale, Public Health Department, La Conception Hospital, Assistance Publique - Hôpitaux de Marseille, 147 Boulevard Baille, Marseille, France.

* Correspondence: Laurent Boyer, Aix-Marseille University, Faculté de Médecine - Secteur Timone, EA 3279: CEReSS - Centre d'Etude et de Recherche sur les Services de Santé et la Qualité de vie, 27 Boulevard Jean Moulin, Marseille 13005, France (e-mail: laurent.boyer@ap-hm.fr).

Copyright © 2020 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Jaotombo F, Pauly V, Auquier P, Orleans V, Boucekine M, Fond G, Ghattas B, Boyer L. Machine-learning prediction of unplanned 30-day rehospitalization using the French hospital medico-administrative database. *Medicine* 2020;99:49(e22361).

Received: 17 October 2019 / Received in final form: 18 July 2020 / Accepted: 26 August 2020

<http://dx.doi.org/10.1097/MD.00000000000022361>

In a recent work,^[11] we developed an easy-to-use predictive rehospitalization risk score of unplanned 30-day all-cause rehospitalization using a logistic regression (LR) model based on 13 variables from the French hospital medico-administrative database (Programme de Médicalisation des Systèmes d'Information - PMSI). This predictive rehospitalization risk score yielded better discriminatory properties than the LACE index score^[12] (*c*-statistic = 0.74 vs 0.66, respectively). The LACE index score is one of the most widely used predictive tools in the world and the current instrument recommended by the French Health Authority. Despite this improvement, this new score presented moderate discriminative ability and needs to be more accurate. The other prediction models in the literature present similar properties, with *c*-statistics of approximately 0.70 (e.g., hospital score = 0.72^[13]). The common point among the prior work is to use traditional statistical methods such as logistic regression (LR) models. Recently, machine learning (ML) methods have been introduced in health service research and have shown a better level of prediction than traditional statistical approaches in several domains.^[14–21] ML methods offer key benefits over traditional statistical approaches because they account for nonlinear relationships between the outcome and the predictors and yield more stable predictions.^[22] ML methods account for interactions between predictors which relaxes the homogeneity assumption that there are no interactions among predictors. To our knowledge, ML methods have rarely been applied to improve the prediction of all-cause rehospitalization. A recent study^[23] developed models using an ML approach to predict 30-day all-cause rehospitalization in patients hospitalized for heart failure but without prediction improvement when compared to LR models (*c*-statistics < 0.61). Another recent study^[24] reported that automated ML better predicted readmissions than commonly used readmission scores in 3 US hospitals (*n* = 16,649).

Thus, the objective of this work was to compare the predictive performance of traditional logistic and ML models to predict 30-day all-cause rehospitalizations on a large population-based study from the French hospital medico-administrative database, based on the following 2 criteria: the area under the receiving operating characteristic curve and the H-measure. For this purpose, we selected the best ML methods: random forest (RF), neural networks (NN) and gradient boosting (GB),^[25] which we compared with 2 reference methods: LR and classification and regression trees (CART) methods.

2. Methods

2.1. Study design

This was a retrospective cohort study of all acute-care inpatient hospitalization cases discharged from January 1 to December 31, 2015, from the largest university health center in south France (*Assistance Publique – Hôpitaux de Marseille*, APHM). All data were collected from the French Hospital database (PMSI - *Programme de Médicalisation des Systèmes d'Information*).^[26] The PMSI is the French medico-administrative database for all hospitalizations based on diagnosis-related groups that we could group into significant diagnostic categories. Research on such retrospective data are excluded from the framework of the French Law Number 2012–300 of March 5, 2012 relating to the research involving human participants, as modified by the Order Number 2016–800 of June 16, 2016. Neither the French competent authority (Agence Nationale de Sécurité du Médica-

ment et des Produits de Santé, ANSM) approval nor the French ethics committee (Comités de Protection des Personnes, CPP) approval is required in this context.

2.2. Study setting and inclusion criteria

The APHM is a public tertiary-care center comprising 4 hospitals (La Timone, La Conception, Sainte-Marguerite, and North) with 3400 beds and 2000 physicians. Approximately 300,000 hospitalizations are recorded every year at the APHM, involving approximately 210,000 patients. All acute-care hospitalizations were included in this study. We excluded hospitalizations in the ambulatory care unit (i.e., ambulatory surgery, radiotherapy, dialysis, chemotherapy, and transfusions) as well as in-hospital mortalities.

2.3. Study outcome

The study outcome was unplanned 30-day all-cause rehospitalization (a binary variable where positive rehospitalization is coded *y* = 1), defined as any cause of readmission via emergency departments in any acute care wards within 30 days of discharge. To calculate this outcome, a unique and individual PMSI identifying variable was used to track rehospitalizations, 30 days following discharge. No more than 1 rehospitalization for each discharge was taken into account. Readmission via the emergency department was employed to identify unplanned rehospitalizations.^[27]

2.4. Collected data

The dataset collected from the PMSI used 29 predictor variables based on a previous work:^[11]

- sociodemographic characteristics: age, gender, state-funded medical assistance (Aide Médicale d'Etat, AME) (i.e., health coverage for undocumented migrants), and free universal health care (Couverture Maladie Universelle, CMU) (i.e., universal health coverage for those not covered by employment/business-based schemes);
- clinical characteristics: category of disease based on the 10th revision of the International Statistical Classification of Diseases, disease severity (no or low severity, moderate – high severity or not determined for short hospitalizations) based on an algorithm issued from the PMSI and 17 comorbidities from the Charlson comorbidity index^[28] (supplementary file 4, <http://links.lww.com/MD/F248>);
- hospitalization characteristics: patient origin (home or other hospital institution), hospitalization via emergency departments, LOS, destination after hospital discharge (home or transfer to other hospital institution), and hospitalization via emergency departments in the previous 6 months.

2.5. Statistical models

Five distinct types of predictive models were fitted to the data: LR considered as the reference, CART, RF, GB, and 1 hidden-layer NN. These models have been explained elsewhere in detail^[29]; a brief summary is presented here.

LR is a linear model of the exponential family such that $\ln\left(\frac{\pi}{1-\pi}\right) = w^T x$, where $\pi = P(y = 1|x)$ and *w* is the weight vector to be estimated from the data.

CART^[30] is a binary decision tree (DT) method that involves segmenting the predictor space into a number of simple regions. CART can be applied to both regression and classification problems, as in our study. A DT is constructed through an iterative process by applying a binary splitting rule. For each variable x_j in the data, a rule of the form $x_j < a$ ($a \in \mathbb{R}$ is a threshold) is used to split the initial set of observations (denoted t_0 , the root of the tree) into 2 subsets t_l and t_r (the sibling nodes). Each observation falling in those regions is then predicted by the highest frequency class. The best split is defined as the one minimizing a loss function (i.e., the Gini index). Once the best split has been defined, the same process is applied to the 2 nodes t_l and t_r and repeated until a predefined minimum number of observations is reached. Then, a pruning algorithm can be used to search for an optimal tree, given a penalty criterion (*complexity parameter*) applied to the objective function. A DT can be represented graphically and thus can be directly interpretable, given its simple structure.

RF^[31] is an ensemble learning method based on aggregating *n* trees similar to the ones constructed with CART, each one grown from a bootstrap sample of the original data set. Each tree in the forest uses only a random subset of *m* predictors at each node. The trees are not pruned. Each value predicted by RF is the average of the values predicted by the *n* trees.

GB^[32] is also an ensemble learning method based on DT but does not involve bootstrap sampling. Given a loss function (i.e., squared error for regression) and a weak learner (i.e., regression trees), the GB algorithm seeks to find an additive model that minimizes the loss function. It is initialized with the best guess of the response (i.e., the mean of the response in regression), then the gradient (i.e., residual) is calculated, and a model is then fit to the residuals to minimize the loss function. The current model thus obtained is added to the previous model, adjusted by a *shrinkage* parameter, and the procedure continues for a user-specified number of iterations, leading to a *n*.trees total number of trees, a tree depth equal to *interaction.depth* and a given minimum number of observations in the trees terminal nodes, *n*.minobsinnode.

NN^[33] are nonlinear statistical models for regression or classification. They are structured in layers of “neurons” where the input layer is made of the predictor variables, the output layer contains as many neurons as there are classes (2 in our study), and one to many intermediate (*size*) layers of “weights” called hidden layers. Each neuron is a linear combination of the neurons of the previous layer, to which is applied an activation function, typically the sigmoid function: $g(x) = \frac{1}{1 + \exp(-x)}$. The weights are the parameters of the model and they are estimated through a back-propagation algorithm called *gradient descent*. The loss function used is the cross-entropy to which a *decay* penalty is applied.

2.6. Statistical analyses

The statistical unit of the data was hospitalization. Descriptive analyses for the sociodemographic, clinical, and hospitalization data were expressed as frequencies and percentages. Chi-squared tests were employed to compare sociodemographic, clinical, and hospitalization data between unplanned 30-day all-cause rehospitalized ($y=1$) and nonrehospitalized patients ($y=0$).

To train and evaluate the different models (i.e., LR, CART, RF, NN, and GB), the dataset was split into a 70% training sample and a 30% test sample, stratified on the outcome variable. On the

training set, we performed a 5-fold cross validation repeated 5 times to tune the hyperparameters. We kept the optimal hyperparameter values for which the loss was minimum. The tuning process and the values of the optimal hyperparameters are presented in supplementary file 2, <http://links.lww.com/MD/F246>. On the test set, we assessed the performance of each model using the optimal hyperparameters. We randomly split the test set in 2 parts: 70% of the sample as a training set and 30% of the sample as a test set. This procedure was repeated 100 times and we computed the average of H-measure and AUC for each model. Since we evaluate different classification rules and the outcome distribution is unbalanced, we used the H-measure, which has the advantage of being classifier-independent and is relevant for heavily unbalanced datasets.^[34] The area under the receiving operating characteristic (ROC) curve (AUC) was also used because it is threshold independent and is a widely used measure. The H-measure and the AUC of each prediction model were compared using a paired *t* test.

Finally, we presented variable importance (VI) (i.e., the most important discriminators between classes) for LR and the optimal prediction model (i.e., RF). VI for the LR is given by the reduction in the deviance each variable brings to the null model. For the RF algorithm, VI is calculated by the mean decrease in Gini (MDG) over all the *m* trees for each variable. We applied a corrected feature importance measure to consider categorical variables with a large number of categories which can bias RF models.^[35] The changes in Gini are aggregated for each variable and normalized.^[31] A high value of the aggregate of the changes indicates great variable importance. All analyses were implemented with R (version 3.5.0) using the caret R (version 6.0.80), hmeasure (version 1.0) and pROC packages (version 1.12.1).

3. Results

3.1. Rates of unplanned 30-day all-cause rehospitalization

A total of 289,358 hospitalizations (112,662 patients) were recorded in the year 2015 at this French University Hospital. After excluding mortalities and hospitalizations for ambulatory surgery, radiotherapy, and dialysis, 118,650 hospitalizations (82,862 patients) were included. The most common diseases were digestive disease, nervous system conditions, and cardiovascular and pulmonary diseases. In total, 4127 (3294 patients) (3.5%) hospitalizations resulted in rehospitalizations via emergency departments 30 days after discharge. Rehospitalization rates according to sociodemographic, clinical, and hospitalization characteristics are presented in supplementary file 1, <http://links.lww.com/MD/F245>.

3.2. Predictive model performance

The predictive performance of each model is presented in Table 1, and the comparison of each models H-measure and AUC is presented in Table 2. The RF model was the most performant model with the highest H-measure (0.290) and AUC (0.794), superior to all the other models (all *P* values <.0001). The performance of the RF, GB, and NN models (H-measures ranged from 0.184 to 0.290, AUC ranged from 0.741 to 0.794) was superior to that of the LR model (H-measure=0.184, AUC=0.740); all *P* values <.0001. In contrast, LR was superior to CART (H-measure=0.162, AUC=0.707), *P*<.0001.

Table 1
Predictive performance for LR, CART, RF, GB, and NN models.

	H (95%CI)	AUC (95%CI)
LR	0.1838 (0.1822;0.1854)	0.7398 (0.7387;0.7408)
CART	0.1551 (0.1536;0.1566)	0.7010 (0.6999;0.7021)
RF	0.3653 (0.3630;0.3675)	0.7688 (0.7675;0.7701)
GB	0.2193 (0.2175;0.2210)	0.7626 (0.7615;0.7636)
NN	0.1846 (0.1830;0.1862)	0.7408 (0.7397;0.7418)

95%CI = 95% confidence interval, AUC = area under the ROC curve, CART = classification and regression trees, GB = gradient boosting, H = H-measure, LR = logistic regression, NN = neural networks, RF = random forest.

From the optimal cut-point estimated for RF model, the specificity was high (0.99) and the sensitivity was low (0.18).

3.3. Variable importance

The variable importance is presented for the RF and LR models in Figures 1 and 2. The 7 most important variables are identical (with slightly difference in ranking) and their contributions to reducing the deviance are comparable: “at least one previous hospitalization via emergency departments 6 months before”, “category of disease”, “hospitalization via emergency departments”, “length of stay”, “age”, “severity”, and “type of hospital stay”.

The variable importance of the other models is presented in supplementary file 3, <http://links.lww.com/MD/F247>.

4. Discussion

In this large sample of acute care inpatients (82,862 patients and 118,650 hospitalizations), ML methods (i.e., RF, GB, and NN), except for CART, are superior to LR for predicting 30-day all-cause rehospitalizations. To date, the majority of studies have focused on particular conditions, for example, patients with specific diagnoses.^[36] This finding confirms the importance of ML models in predicting rehospitalization, despite previous contradictory results on this subject.^[23] RF achieves the best

Table 2
Comparison of AUC and H-measures of ML models to LR and RF (paired *t* tests).

Ref. Model: LR	Index	Statistic	<i>P</i> value
H <i>t</i> tests	H-GB	−93.84	<.0001
	H-NN	−9.15	<.0001
	H-CART	67.00	<.0001
	H-RF	−194.00	<.0001
AUC <i>t</i> tests	AUC-GBM	−97.76	<.0001
	AUC-NN	−29.67	<.0001
	AUC-CART	122.28	<.0001
	AUC-RF	−50.34	<.0001
H <i>t</i> tests	H-GB	166.61	<.0001
	H-NN	196.01	<.0001
	H-CART	200.53	<.0001
	H-LR	194.00	<.0001
AUC <i>t</i> tests	AUC-GB	11.97	<.0001
	AUC-NN	49.22	<.0001
	AUC-CART	106.25	<.0001
	AUC-LR	50.34	<.0001

CART = classification and regression trees, GB = gradient boosting, LR = logistic regression, NN = neural networks, RF = random forest.

performance among all models according to the H-measure and the AUC. This result is consistent with recent studies reporting that RF is a relevant and accurate method for predicting health outcomes,^[37–40] although some studies report no improvement in ML models compared to LR.^[23]

RF is an easy-to-understand method providing an original variables importance index that helps identify the top-ranked variables associated with 30-day all-cause rehospitalizations.^[31] This property of RF should be highlighted regarding the traditional trade-off between accuracy and interpretability in statistical modeling.^[41] Contrary to LR, ML models (e.g., RF, GB, NN) are considered to be black boxes because there is not always a clear interpretable connection between outcomes and predictors. However, there has been a tremendous amount of work in developing ways to explain black box models. Variable importance is one of them. In our study, 2 important findings should be highlighted.

First, the 7 most important variables are identical (with slightly difference in ranking) and their contributions to reducing the deviance are comparable between RF and LR. This homogeneity of findings between the 2 methods is reassuring for the interpretation of results by health care providers. Hospitalization via Emergency Departments and previous hospitalization via emergency departments 6 months before are generally associated with higher readmission in previous works.^[42] Older adults are also described as at higher risk of readmission in previous studies.^[4,5] Concerning the category of disease, medical-psychiatric comorbidity was highly related to rehospitalizations, confirming previous studies on this complex population.^[43,44] This finding justifies the identification of hospitalized patients with psychiatric conditions to better address their behavioral needs. The length of hospital stay was inconstantly associated with higher readmission in previous works.^[45–47] French hospitals are under pressure to save on costs, and reducing LOS is strongly advocated. Future studies should thus explore the consequences of this health policy in the French context, particularly its impact on rehospitalization and, more generally, on quality of care.

Second, there are more variables of importance above a threshold of 10% in RF (7 variables: state funded medical assistance, gender, destination on discharge, congestive heart failure, chronic pulmonary disease, dementia, free universal health care, and malignancy) than in LR (only 1 variable: dementia). This suggests that RF is better able to identify discriminating variables than LR, including clinical and socioeconomic variables. For example, socioeconomic status (i.e., state funded medical assistance and free universal health care in our study) was associated with rehospitalization in our study, confirming recent findings on social risk (poverty, disability, housing instability, residence in a disadvantaged neighborhood) and rehospitalization.^[48] Interestingly, previous studies also reported gender inequalities^[49] and risk associated with congestive heart failure,^[4] chronic pulmonary disease,^[50] dementia.^[51]

Despite our findings in favor of RF and ML methods, 2 issues must be considered in future work: a moderate improvement, especially for the AUC, between ML and LR, and the use of data at discharge.

Our study included a relatively small set of variables (29 variables), relevant for classical statistical methods based on standard parametric models but suboptimal for ML methods in some respects. Several additional pieces of information could be

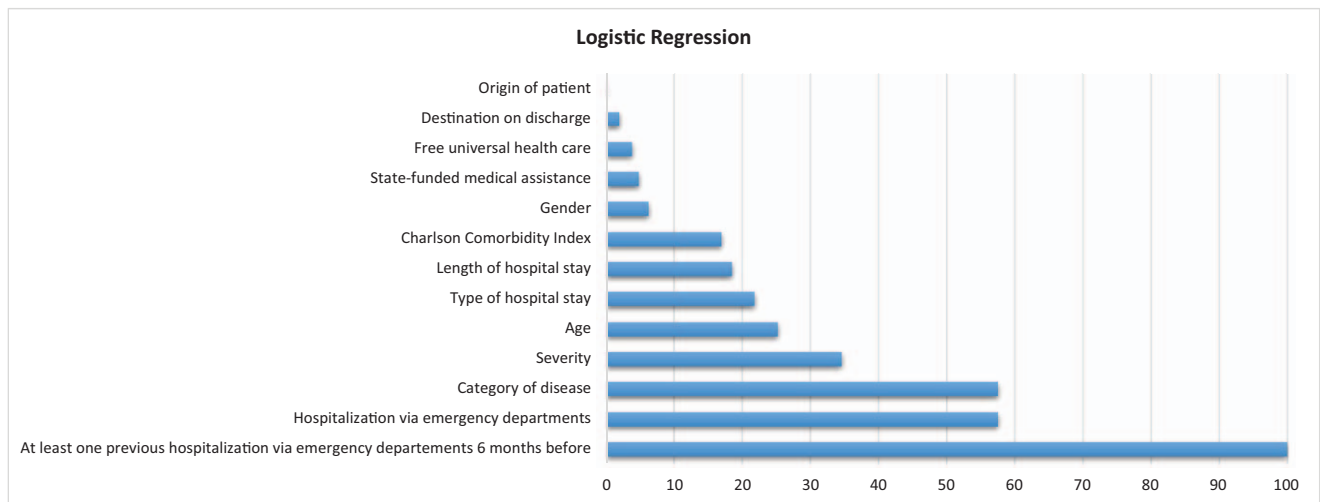


Figure 1. Importance of variables in LR model.

relevant to predict rehospitalization, including structured (e.g., socioeconomic status, drugs, and self-reported functional status) and unstructured (e.g., clinical notes from physicians, nurses, and other professionals) data available in electronic medical records. These data could improve prediction by offering richer medical information than those found in the only medico-administrative databases. Previous studies reported that the performance of ML methods could be improved by taking into account a larger number of variables.^[52] Future studies should include all data available in electronic medical records.

As for the majority of predictive risk scores, our study was based on data at discharge, while predictive risk scores should ideally give information early enough during hospitalization to trigger care intervention.^[53] To date, instruments based on discharge data have been proven to lead to models with better performance^[53,54] than models based solely on admission data. An important perspective would be to implement real-time predictive rehospitalization risk scores during hospitalization, updated for all new available data, and then propose early alerts

for high risk of rehospitalization. A recent study reported that ML methods can be used in real-time predictions using routinely collected clinical data exclusively, without the need for any manual processing.^[55] Another recent study trained and tested a neural network model to predict the risk of patients rehospitalization within 30 days of their discharge based on real-time data from EHR, and thus applicable at the time discharge from hospital.^[56]

Our findings must be interpreted in the context of our study's limitations. Despite the large overall sample size of this multihospital study, our findings may not be applicable to all French hospitals, particularly general hospitals where patients have potentially different characteristics from those of university hospitals. In addition, the 4 university hospitals included in our study were located in only one geographical area, and social and healthcare geographical characteristics (e.g., poverty, density of physicians, number of beds, and private hospitals) are known to influence the risk of rehospitalization.^[53,57] Future studies should thus be conducted in different categories of hospitals and in

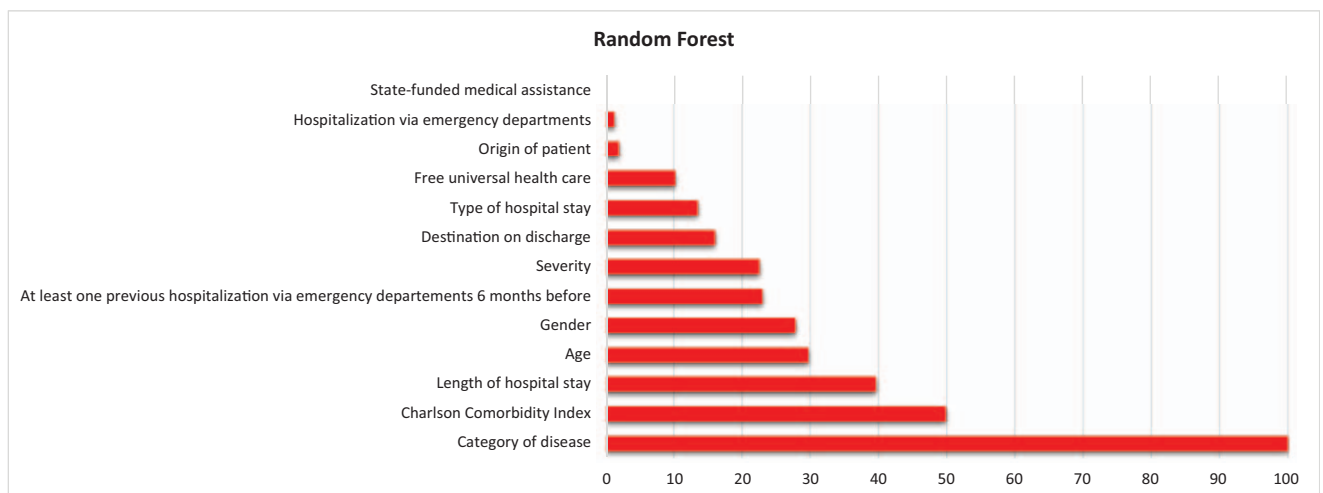


Figure 2. Importance of variables in RF model.

several geographical areas to confirm the properties and importance of our predictive risk score. Our model does not factor in deaths outside the hospital because we do not account for this information in our database. Other studies with available data on outpatient events are needed to investigate to what extent this could impact our predictive risk score using a competing risk model as an example. We excluded ambulatory surgery from the analyses. This specific topic should be studied in the French context, strongly marked by pressures for reducing length of stay. Lastly, the caret R package offers the possibility of using other statistical models that could be studied in future work (e.g., Multi-Layer Perceptron Neural Network, Support Vector Machine, Bayesian Network).

5. Conclusion

The use of ML may be an alternative to regression models to predict health outcomes. The integration of ML, particularly the RF algorithm, in the prediction of unplanned rehospitalization, may help health service providers target patients at high risk of rehospitalizations and propose effective interventions at the hospital level.

Author contributions

F Jaotombo and L Boyer wrote the first draft of the manuscript. V Pauly and V Orleans carried out the selection process. F Jaotombo and B Ghattas carried out the statistical analyses. All authors have reviewed the final manuscript.

References

- [1] Boutwell AE, Johnson MB, Rutherford P, et al. An early look at a four-state initiative to reduce avoidable hospital readmissions. *Health Aff (Millwood)* 2011;307:1272–80.
- [2] HAS. Haute Autorité de Santé. Note méthodologique et de synthèse documentaire «Sortie d'hospitalisation supérieure à 24 heures—Établissement d'une check-list». In Available from: http://www.has-santefr/portail/upload/docs/application/pdf/2015-05/note_documentaire_check-list_sortie_hospitalisation_webpdf.pdf. 2015
- [3] Gusmano M, Rodwin V, Weisz D, Cottenet J, Quantin C. Comparison of rehospitalization rates in France and the United States. *J Health Serv Res Policy* 2015;201:18–25.
- [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] Friedman B, Basu J. The rate and cost of hospital readmissions for preventable conditions. *Med Care Res Rev* 2004;612:225–40.
- [6] Ashton CM, Kuykendall DH, Johnson ML, Wray NP, Wu L. The association between the quality of inpatient care and early readmission. *Ann Intern Med* 1995;1226:415–21.
- [7] Balla U, Malnick S, Schattner A. Early readmissions to the department of medicine as a screening tool for monitoring quality of care problems. *Medicine (Baltimore)* 2008;875:294–300.
- [8] Francois P, Bertrand D, Beden C, Fauconnier J, Olive F. Early readmission as an indicator of hospital quality of care. *Rev Epidemiol Sante Publique* 2001;492:183–92.
- [9] Coleman EA, Parry C, Chalmers S, Min SJ. The care transitions intervention: results of a randomized controlled trial. *Arch Intern Med* 2006;16617:1822–8.
- [10] Leppin AL, Gionfriddo MR, Kessler M, et al. Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. *JAMA Intern Med* 2014;1747:1095–107.
- [11] Pauly V, Mendizabal H, Gentile S, et al. Predictive risk score for unplanned 30-day rehospitalizations in the French universal health care system based on a medico-administrative database. *PLoS One* 2019;14:e0210714.
- [12] van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *CMAJ* 2010;1826:551–7.
- [13] Donze JD, Williams MV, Robinson EJ, et al. International Validity of the HOSPITAL Score to Predict 30-Day Potentially Avoidable Hospital Readmissions. *JAMA Intern Med* 2016;1764:496–502.
- [14] Acion L, Kelmansky D, van der Laan M, et al. Use of a machine learning framework to predict substance use disorder treatment success. *PLoS One* 2017;124:e0175383.
- [15] Ahn JM, Kim S, Ahn KS, et al. A deep learning model for the detection of both advanced and early glaucoma using fundus photography. *PLoS One* 2018;1311:e0207982.
- [16] Chekroud AM, Zotti RJ, Shehzad Z, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry* 2016;33:243–50.
- [17] Gholipour C, Rahim F, Fakhree A, Ziapour B. Using an Artificial Neural Networks (ANNs) model for prediction of Intensive Care Unit (ICU) outcome and length of stay at hospital in traumatic patients. *J Clin Diagn Res* 2015;94:OC19–23.
- [18] Kim SJ, Cho KJ, Oh S. Development of machine learning models for diagnosis of glaucoma. *PLoS One* 2017;125:e0177726.
- [19] Kuo PJ, Wu SC, Chien PC, et al. Derivation and validation of different machine-learning models in mortality prediction of trauma in motorcycle riders: a cross-sectional retrospective study in southern Taiwan. *BMJ Open* 2018;81:e018252.
- [20] LaFaro RJ, Pothula S, Kubal KP, et al. Neural network prediction of ICU length of stay following cardiac surgery based on pre-incision variables. *PLoS One* 2015;1012:e0145395.
- [21] Stylianou N, Akbarov A, Kontopantelis E, Buchan I, Dunn KW. Mortality risk prediction in burn injury: comparison of logistic regression with machine learning approaches. *Burns* 2015;415:925–34.
- [22] Kuhn M, Johnson K. *Applied Predictive Modeling*. 2013;Springer, 26.
- [23] Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol* 2017;22:204–9.
- [24] Morgan DJ, Bame B, Zimand P, et al. Assessment of machine learning vs standard prediction rules for predicting hospital readmissions. *JAMA Netw Open* 2019;23:e190348.
- [25] Fernandez-Delgado M, Cernadas E, Barro S. Do we need hundreds of classifiers to solve real world classification problems? *J Mach Learn Res* 2014;15:3133–81.
- [26] Boudemaghe T, Belhadj I. Data resource profile: the French National Uniform Hospital discharge data set database (PMSI). *Int J Epidemiol* 2017;462:392–1392.
- [27] Bottle A, Aylin P, Majeed A. Identifying patients at high risk of emergency hospital admissions: a logistic regression analysis. *J R Soc Med* 2006;998:406–14.
- [28] Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;4311:1130–9.
- [29] Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. Springer Series in Statistics. New York: Springer; 2017. 764 pages.
- [30] Breiman L. *Classification and Regression Trees*. 1st ed. Wadsworth: International Group; 1984. 368 pages.
- [31] Breiman L. Random forests. *Mach Learn* 2001;45:5–32.
- [32] Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals Statist* 2001;1189–232.
- [33] Arbib MA. *The Handbook of Brain Theory and Neural Networks*. 2003; A Bradford Book, MIT Press, 1344 pages.
- [34] He H, Garcia EA. Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 2009;21:1263–84.
- [35] Altmann A, Tolosi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. *Bioinformatics* 2010;2610:1340–7.
- [36] Garcia-Arce A, Rico F, Zayas-Castro JL. Comparison of machine learning algorithms for the prediction of preventable hospital readmissions. *J Health Qual* 2018;403:129–38.
- [37] Hsieh MH, Hsieh MJ, Chen CM, Hsieh CC, Chao CM, Lai CC. Comparison of machine learning models for the prediction of mortality of patients with unplanned extubation in intensive care units. *Sci Rep* 2018;81:17116.
- [38] Taylor RA, Pare JR, Venkatesh AK, et al. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach. *Acad Emerg Med* 2016;233:269–78.

- [39] Ambale-Venkatesh B, Yang X, Wu CO, et al. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. *Circ Res* 2017;1219:1092–101.
- [40] Artetxe A, Beristain A, Grana M. Predictive models for hospital readmission risk: a systematic review of methods. *Comput Methods Programs Biomed* 2018;164:49–64.
- [41] Nanayakkara S, Fogarty S, Tremeer M, et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. *PLoS Med* 2018;1511:e1002709.
- [42] Brennan JJ, Chan TC, Killeen JP, Castillo EM. Inpatient readmissions and emergency department visits within 30 days of a hospital admission. *West J Emerg Med* 2015;167:1025–9.
- [43] Jansen L, van Schijndel M, van Waarde J, van Busschbach J. Health-economic outcomes in hospital patients with medical-psychiatric comorbidity: a systematic review and meta-analysis. *PLoS One* 2018;133:e0194029.
- [44] Boaz TL, Becker MA, Andel R, McCutchan N. Rehospitalization risk factors for psychiatric treatment among elderly Medicaid beneficiaries following hospitalization for a physical health condition. *Aging Ment Health* 2017;213:297–303.
- [45] Bueno H, Ross JS, Wang Y, et al. Trends in length of stay and short-term outcomes among Medicare patients hospitalized for heart failure. *JAMA* 2010;30321:2141–7.
- [46] Kaboli PJ, Go JT, Hockenberry J, et al. Associations between reduced hospital length of stay and 30-day readmission rate and mortality: 14-year experience in 129 Veterans Affairs hospitals. *Ann Intern Med* 2012;15712:837–45.
- [47] Sud M, Yu B, Wijeyesundera HC, et al. Associations between short or long length of stay and 30-day readmission and mortality in hospitalized patients with heart failure. *JACC Heart Fail* 2017;58:578–88.
- [48] Joynt Maddox KE, Reidhead M, Hu J, et al. Adjusting for social risk factors impacts performance and penalties in the hospital readmissions reduction program. *Health Serv Res* 2019;542:327–36.
- [49] Gonzalez JR, Fernandez E, Moreno V, et al. Sex differences in hospital readmission among colorectal cancer patients. *J Epidemiol Community Health* 2005;596:506–11.
- [50] Buhr RG, Jackson NJ, Dubinett SM, et al. Factors associated with differential readmission diagnoses following acute exacerbations of chronic obstructive pulmonary disease. *J Hosp Med* 2020;152:e1–9.
- [51] Pickens S, Naik AD, Catic A, Kunik ME. Dementia and hospital readmission rates: a systematic review. *Dement Geriatr Cogn Dis Extra* 2017;73:346–53.
- [52] Couronne R, Probst P, Boulesteix AL. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics* 2018;191:270.
- [53] Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011;30615:1688–98.
- [54] Nguyen OK, Makam AN, Clark C, et al. Predicting all-cause readmissions using electronic health record data from the entire hospitalization: model development and comparison. *J Hosp Med* 2016;117:473–80.
- [55] Meyer A, Zverinski D, Pfahringer B, et al. Machine learning for real-time prediction of complications in critical care: a retrospective study. *Lancet Respir Med* 2018.
- [56] Jamei M, Nisnevich A, Wetchler E, Sudar S, Liu E. Predicting all-cause risk of 30-day hospital readmission using artificial neural networks. *PLoS One* 2017;127:e0181173.
- [57] Hernandez AF, Greiner MA, Fonarow GC, et al. Relationship between early physician follow-up and 30-day readmission among Medicare beneficiaries hospitalized for heart failure. *JAMA* 2010;30317:1716–22.