International Journal of Medical and Health Sciences
Volume 11 Issue 03 Sept 2025
ISSN- 2456 - 6063
DOI: https://doi.org/10.69980/gfs9k223

ANALYSIS OF HOSPITAL READMISSION RATES USING PREDICTIVE ANALYTICS FOR IMPROVED HEALTHCARE DELIVERY

Dr. Sandeep Kulkarni[1], Dr. Meenakshi Iyer[2], Dr. Vikram Singh Rathore[3], Dr. Pooja Chatterjee[4], Dr. Karthik Subramanian[5]

[1] Department of Health Informatics, Tata Institute of Social Sciences, Mumbai, Maharashtra, India
[2] Department of Biostatistics, National Institute of Mental Health and Neurosciences, Bengaluru, Karnataka, India
[3] Department of Internal Medicine, SMS Medical College, Jaipur, Rajasthan, India
[4] Department of Public Health, Indian Institute of Public Health, Gandhinagar, Gujarat, India
[5] Department of Hospital Administration, Christian Medical College, Vellore, Tamil Nadu, India
Email: sandeep.kulkarni@tiss.edu

Received:- 22/06/25	Revised:-24/07/25	Accepted:-26/08/25	Published:- 27/09/25

Abstract
Hospital readmission is a key measure of health care quality, safety and hospital performance. This study aimed to investigate hospital readmission rates using predictive analytics to determine which patients were at greater risk of readmission and help inform healthcare services. In this study used a structured hospital readmission dataset with 25,000 patients and 17 variables. Readmission was the dependent variable, with the independent variables being age, length of stay, number of laboratory tests, number of medications, number of outpatient visits, number of inpatient visits, number of emergency visits, medical speciality, diagnosis categories, glucose test, A1C test, medication change and diabetes medication status. There were no missing data or duplicate records in the dataset. Descriptive statistics revealed a readmission rate of 47.02% of patients, suggesting a high readmission rate. Machine learning models, such as Logistic Regression, Decision Tree and Random Forest were trained and tested based on accuracy, precision, recall, F1-score, and ROC-AUC. Random Forest emerged as the top model with the highest F1-score (0.5428) and Logistic Regression with the highest ROC-AUC (0.6434). Laboratory procedures, number of medications, time in hospital, age, diagnosis categories, medical speciality, and past inpatient visits were identified as important features in predicting readmission. The results indicate that predictive analytics can help hospitals identify patients at risk, plan discharge, organise follow-up care and allocate resources, which can lead to better patient-centred care.

Keywords: Hospital Readmission, Predictive Analytics, Machine Learning, Healthcare Delivery, Random Forest, Patient Risk.

1. Introduction

Readmission is a critical measure of health care quality, safety and hospital performance. Readmission rates are commonly used to indicate whether the patient's condition was not completely resolved, discharge planning was insufficient, or follow-up care and disease management were ineffective. In contemporary health care, avoiding preventable readmissions has become a significant priority as readmissions impose a greater clinical burden on hospitals and place a financial burden on health-care providers and patients. Burke stated that 30-day hospital readmission is a key performance indicator for hospitals, especially for conditions covered under hospital readmission reduction programs (Burke et al., 2017). Likewise, Dodson emphasised that patients with complex medical health and older adults are at particular risk of readmission after discharge (Dodson et al., 2019).

The increasing use of electronic health records and other structured data has opened the door to the use of predictive analytics in health care. Predictive analytics uses statistical and machine learning methods to flag patients at risk of readmission prior to or just after discharge. Choi et al. showed that recurrent neural networks can predict future clinical events from patient data, suggesting that more sophisticated machine learning approaches can be applied to health care (Choi et al., 2016). Churpek et al. also showed that machine learning algorithms can enhance the prediction of clinical deterioration compared with traditional regression models (Churpek et al., 2016). These findings suggest that prediction models can assist in early detection of risks and allow for intervention.

Machine learning is increasingly used in various clinical settings such as heart failure, pneumonia, acute kidney injury and all-cause hospital readmission prediction. Awan et al. highlighted the recent application of machine learning in heart failure management and its ability to assist decision-making (Awan et al., 2018). Huang et al. used machine learning techniques to predict 30-day hospital readmission in adults with pneumonia, demonstrating that machine learning models can help identify high-risk populations (Huang et al., 2022). Jamei et al. also showed the value of artificial neural networks for predicting all-cause 30-day hospital readmission (Jamei et al., 2017). These studies highlight the importance of predictive analytics for better managing health care and preventing unnecessary readmissions.

While technology is increasingly being used in healthcare, some hospitals still struggle to predict which patients are at risk of readmission. Clinical experience may not always reflect the interaction of patient demographics, clinical features, treatment intensity, medication complexity and prior hospital use. Davis et al. observed that to predict readmission, it is important to include relevant clinical features and machine-learned features that enhance model performance (S. Davis et al., 2022). But predictive models are not always optimally incorporated into hospital practice. Moreover, model stability is a critical concern because the predictive performance of models may deteriorate over time due to changes in the underlying population, clinical practice or data patterns (S. E. Davis et al., 2017). Thus, there is a need to develop and assess predictive models that can predict the risk of readmission using hospital data.

The first aim of this research is to investigate the pattern of hospital readmissions using patient-level clinical and hospital use data. The second objective is to determine the key variables associated with readmission, including number of days spent in hospital, number of laboratory tests, medication load, diagnostic categories, and history of hospital visits. The third objective is to build models to classify patients into risk groups for readmission. Other recent research has demonstrated that interpretable machine learning can assist health care providers in understanding the important predictors of readmission and support decision-making (Gao et al., 2023). Likewise, Islam et al. highlighted that clinical information systems can enhance healthcare processes if data are successfully used for decision-making (Islam et al., 2018).

This research is important as it uses predictive analytics to analyse hospital readmission data to enhance healthcare. Predicting readmission risk can help hospitals prioritize post-discharge care, improve discharge processes and optimise resource allocation. Predictive models can also help clinicians use data to inform decisions and prevent readmissions. The research adds to the evidence that machine learning and analysis of clinical data can help improve patient outcomes, hospital efficiency and guide evidence-based decision-making in health care.

2. Methodology

2.1 Research Design

This is an applied quantitative research design involving predictive analytics to study the incidence of hospital readmission and to determine the key factors that predict readmission. The study uses a supervised machine learning approach because the primary objective is to develop a model to predict the likelihood of patient readmission following discharge from hospital. The outcome variable is the readmission status of the patient (readmitted or not readmitted). The independent variables include patient demographics, clinical, diagnosis and hospital utilization factors. This is an appropriate approach as the study can not only describe the data but also build predictive models that can inform decision-making in health care.

2.2 Data Source and Description

The dataset used in this study is the hospital readmission data of around 25,000 patients. The data contains information about the age of the patient, length of hospital stay, number of lab procedures, number of medications, number of outpatient visits, emergency visits, inpatient visits, diagnostic category, glucose serum test, A1C test, insulin status, medication change, diabetes medication and readmitted. The outcome variable to be predicted is readmitted, which is a binary variable (yes, no) indicating if a patient is readmitted to the hospital. The dataset is appropriate for the current study as it includes both clinical and treatment variables that are known to be associated with hospital readmission (DD, 2023).

2.3 Variables Used in the Study

Hospital readmission is the dependent variable in this study. This is a dichotomous variable that categorises patients into two groups: readmitted and not readmitted. Age, hospital length of stay, number of laboratory procedures, number of procedures, number of medications, number of diagnoses, number of outpatient visits, number of emergency visits, number of inpatient visits, primary diagnosis, glucose serum test, A1C test, insulin, medication change, and diabetes medication status are the independent variables. These variables were chosen as they reflect important factors of patient's health status, hospital resource use, intensity of treatment and disease management.

2.4 Data Preprocessing

To enhance the quality of the data and make it more suitable for predictive modelling, pre-processing was performed. The data were first checked for missing values, duplicate records, inconsistent values and irrelevant columns. Appropriate methods were used to deal with missing values, depending on the nature and extent of the missing data. In cases of few missing values, the records could be dropped, and in cases of categorical missing values, they could be encoded as a new category if the missing values represented a clinically significant absence of information. Categorical variables such as diagnosis category, glucose serum test result, A1C result, insulin status, medication change, diabetes medication status, and readmission status were encoded in a numerical format. Numerical variables, including age, time in hospital, number of lab tests, number of medications, number of outpatient visits, number of emergency visits, number of inpatient visits and number of diagnoses were examined for their distribution and possible outliers. Such data preparation allowed the data to be used for statistical and machine learning analysis.

2.5 Exploratory Data Analysis

Data exploration was performed to gain insight into the data structure and characteristics before predictive modeling. The distribution of the readmission variable was analysed to understand the distribution of patients who were readmitted and those who were not. Descriptive summaries were calculated for the quantitative variables: age, length of stay, number of medications, number of lab tests, number of diagnoses and previous hospital visits. Cross-tabulation, frequency analysis and visualizations were used to examine the association between readmission and important predictors. This phase assisted in the identification of important patterns, such as whether patients with longer hospital stays, higher medication count, more diagnoses and previous hospital visits were more likely to be readmitted. Exploratory data analysis also informed the choice of variables to include in the predictive model.

2.6 Predictive Model Development

In the predictive modeling stage, supervised machine learning models were used to classify patients according to their risk of readmission. The dependent variable is binary in nature, so classification models were suitable for the study. Logistic regression was applied as a baseline model as it is interpretable and can be used to understand the effect of predictor variables on the likelihood of readmission. Decision tree classification was used to model non-linear effects and provide a set of rules for assessing patient readmission risk. Random forest classification was also applied because it is an ensemble of decision trees that usually increases the accuracy of predictions by preventing overfitting. Using several models enabled us to compare their performance and find the best model for predicting hospital readmission.

2.7 Training and Testing Procedure

To assess the performance of the predictive models, the dataset was split into training and testing data. We used the standard 70:30 split for training and testing the models, with 70 percent of the data set used for training and 30 percent for testing. The training data was used to build the models and learn the associations between patient features and readmission. The models were then tested on the testing dataset to assess their performance on new data. This approach minimised the risk of overfitting and allowed a better assessment of model performance.

2.8 Model Evaluation
We used classification metrics to assess the performance of the predictive models. Accuracy was used to assess the overall accuracy of the predictions. Precision was used to measure the proportion of patients predicted to be readmitted that were actually readmitted. Recall (also called sensitivity) was used to assess the proportion of patients predicted as readmitted who were actually readmitted. The F1-score was used as a harmonic mean of precision and recall. Furthermore, we considered the ROC-AUC to measure the model's ability to separate readmitted and non-readmitted patients. These performance measures are crucial in medical research as the ability to accurately predict which patients are at risk of readmission can help hospitals to identify and reduce preventable readmissions.

2.9 Feature Importance and Interpretation
Once the models were trained, we performed feature importance analysis to determine the most influential predictors of hospital readmission. For tree-based models (decision tree and random forest), feature importance scores were used to identify the variables that were most influential in the classification process. Predictors such as prior hospital visits, length of stay, number of medications prescribed, number of diagnoses, number of laboratory tests and diabetes treatment variables were thought to have a strong impact on the risk of readmission. This was an important step because predictive analytics in health care should not only be predictive, but also provide actionable insights to inform clinical and administrative decisions.

2.10 Ethical Considerations
The data analysed in this study were anonymised secondary data. The analysis did not involve any personal identification of patients. Research was undertaken only for educational purposes. Health data are inherently sensitive and issues of confidentiality, privacy and data use were taken into account. The findings from the predictive analysis were interpreted with caution to inform better healthcare practices and not to label or discriminate against patients. The goal of the study was to assist in discharge planning, identify patients at risk of poor outcomes and ultimately improve the delivery of health care through evidence-based practice.

2.11 Methodological Summary
Overall, the approach of this study was to use a structured hospital readmission dataset to build predictive models to identify patients at risk of readmission. Data cleaning, preprocessing, and descriptive and visual analysis were conducted. Machine learning algorithms such as logistic regression, decision tree, and random forest were used to predict readmission. The models were evaluated based on accuracy, precision, recall, F1-score, and ROC-AUC. The methodology is fit for the title of the research because it involves the analysis of hospital readmissions, predictive modelling, and offers insights that can potentially inform better healthcare services.

3. Results
3.1 Descriptive Statistics
There were 25,000 hospital patient records and 17 variables in the data set. There were no missing values or duplicate records, so the data set was suitable for analysis. The dependent variable was readmitted, and the independent variables were age, days in hospital, laboratory procedures, medical procedures, medications, number of previous outpatient visits, number of previous inpatient visits, number of previous emergency visits, medical speciality, diagnosis categories, glucose test result, A1C test result, medication change and diabetes medication status.

Descriptive analysis indicated that the average length of hospital stay was 4.45 days (1-14 days). The average number of laboratory procedures was 43.24 and the average number of medications was 16.25. The average values for the number of outpatient, inpatient, and emergency visits were relatively low, but the maximum values indicated that some patients had multiple hospital visits. The complete descriptive statistics are in Table 1.

Table 1. Descriptive Statistics of Numerical Variables

Variable	N	Mean	Standard Deviation	Minimum	Median	Maximum
Time in hospital	25,000	4.45	3.00	1	4	14
Number of lab procedures	25,000	43.24	19.82	1	44	113
Number of procedures	25,000	1.35	1.72	0	1	6
Number of medications	25,000	16.25	8.06	1	15	79
Number of outpatient visits	25,000	0.37	1.20	0	0	33
Number of inpatient visits	25,000	0.62	1.18	0	0	15
Number of emergency visits	25,000	0.19	0.89	0	0	64

The distribution of the readmission status revealed 13,246 (52.98%) patients were not readmitted and 11,754 (47.02%) patients were readmitted. This suggests that the data was fairly well distributed between the two classes, and could be used to model classification problems. The distribution of readmission status is presented in Table 2 and Figure 1.

Table 2. Distribution of Hospital Readmission Status

Readmission Status	Frequency	Percentage
No	13,246	52.98%
Yes	11,754	47.02%

3.2 EDA Findings

Data exploration revealed that hospital readmission was a frequent outcome in our data, with almost half of patients being readmitted. The large readmission group suggests that the prediction models are needed to identify the patients who are most at risk. The readmission pattern shown in Figure 1 demonstrates that readmitted and non-readmitted patients were well represented in our data.

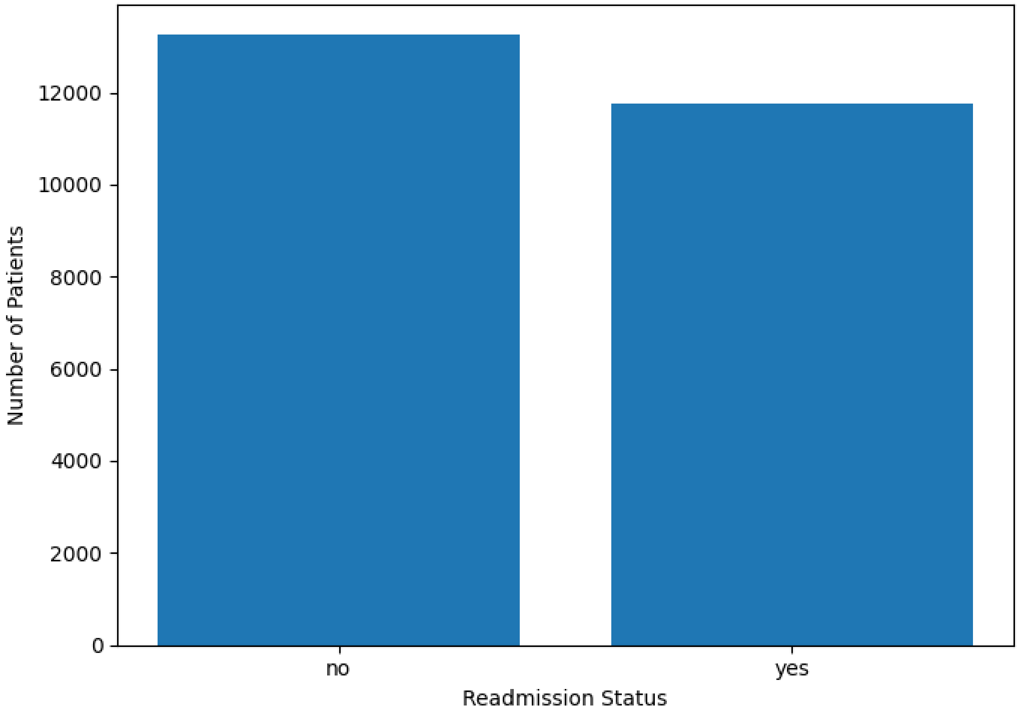

Figure 1. Distribution of Hospital Readmission Status

The quantitative variables demonstrated substantial differences in treatment and health-care usage. The average time in hospital is 4.45 days, the average number of laboratory procedures is 43.24 procedures, and the average number of medications is 16.25, as shown in Table 1. These numbers suggest that patients were being evaluated for many different clinical parameters and medications while in hospital.

Comparing the average clinical and utilization variables by whether patients were readmitted indicated that readmitted patients were associated with greater clinical and healthcare utilization. The variables of time in hospital, number of laboratory procedures, number of medications, previous inpatient visits, outpatient visits and emergency visits were helpful to compare patients with and without readmission. This is shown in Figure 2.

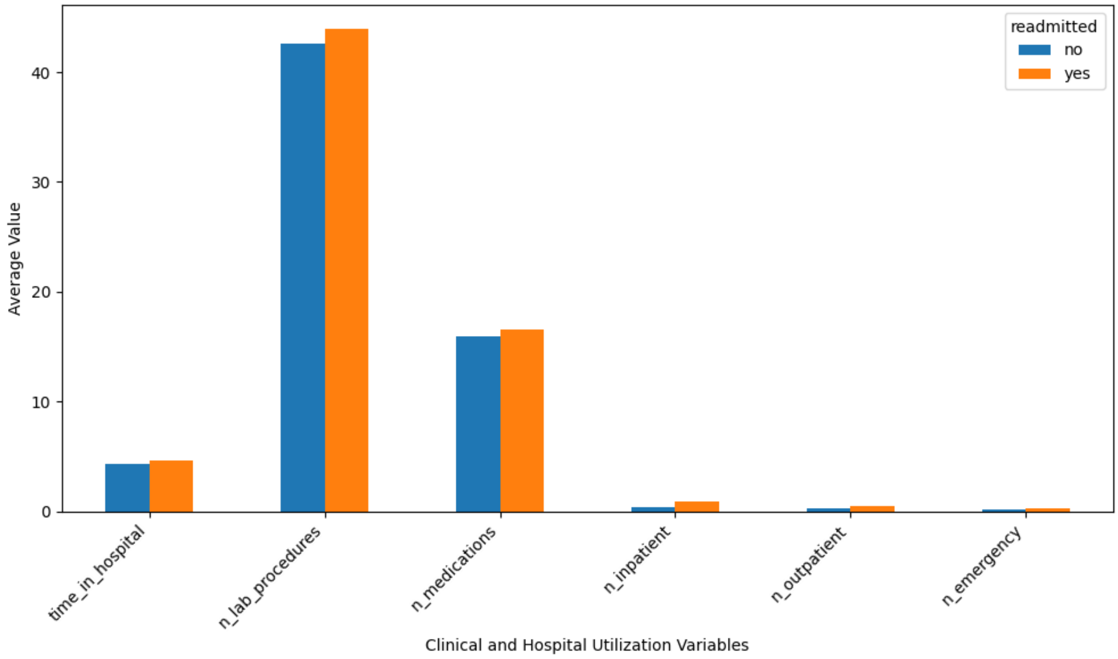

Figure 2. Average Clinical Characteristics by Readmission Status

In summary, the results of the EDA show that hospital readmission is associated with a complex of care intensity, patient status at discharge, and previous health care use. This provided the rationale for using data mining to identify patients at risk of readmission.

3.3 Model Performance

We developed and tested three supervised machine learning algorithms: Logistic Regression, Decision Tree, and Random Forest. The performance metrics used were accuracy, precision, recall, F1-score and ROC-AUC. The results of the comparison are shown in Table 3.

Table 3. Predictive Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.6060	0.6284	0.3962	0.4860	0.6434
Decision Tree	0.5409	0.5117	0.5159	0.5138	0.5395
Random Forest	0.5991	0.5851	0.5062	0.5428	0.6283

As shown in Table 3, Logistic Regression had the highest accuracy (0.6060) and ROC-AUC (0.6434). But it had a low recall (0.3962), which indicates that it was not extremely efficient in detecting patients who were readmitted. The Decision Tree model had a recall of 0.5159, but the lowest accuracy and ROC-AUC.
The Random Forest model achieved an accuracy of 0.5991, precision of 0.5851, recall of 0.5062, F1-score of 0.5428, and ROC-AUC of 0.6283. While Logistic Regression achieved the highest ROC-AUC, Random Forest achieved the highest F1-score, which means it achieved the best balance between precision and recall. Hence, the Random Forest model was considered the best model for this investigation.

3.4 Feature Importance

We used the Random Forest model to conduct feature importance analysis, which measures the importance of the predictors of hospital readmission. This analysis revealed that number of laboratory procedures was the most important predictor, followed by number of medications and time in hospital. This suggests that the complexity of diagnosis, medication load and time in hospital were key factors in predicting hospital readmission. The most important predictors are listed in Table 4 and displayed in Figure 3.

Table 4. Top Ten Important Predictors of Hospital Readmission

Rank	Predictor	Importance Score
1	Number of lab procedures	0.1743
2	Number of medications	0.1447

3	Time in hospital	0.0983
4	Age	0.0725
5	Primary diagnosis	0.0698
6	Medical specialty	0.0676
7	Third diagnosis	0.0669
8	Second diagnosis	0.0652
9	Number of procedures	0.0606
10	Number of inpatient visits	0.0558

The results in Table 4 indicate that clinical and utilization variables affected readmission. The importance of the number of laboratory procedures and number of medications suggests that patients with more laboratory tests and medications may have more complicated disease. Likewise, the importance of time in hospital suggests that the duration of hospital stay may be related to disease severity. The presence of diagnostic variables (primary, secondary and tertiary diagnoses) among the top predictors also indicates that disease profile is a critical factor in predicting readmission.

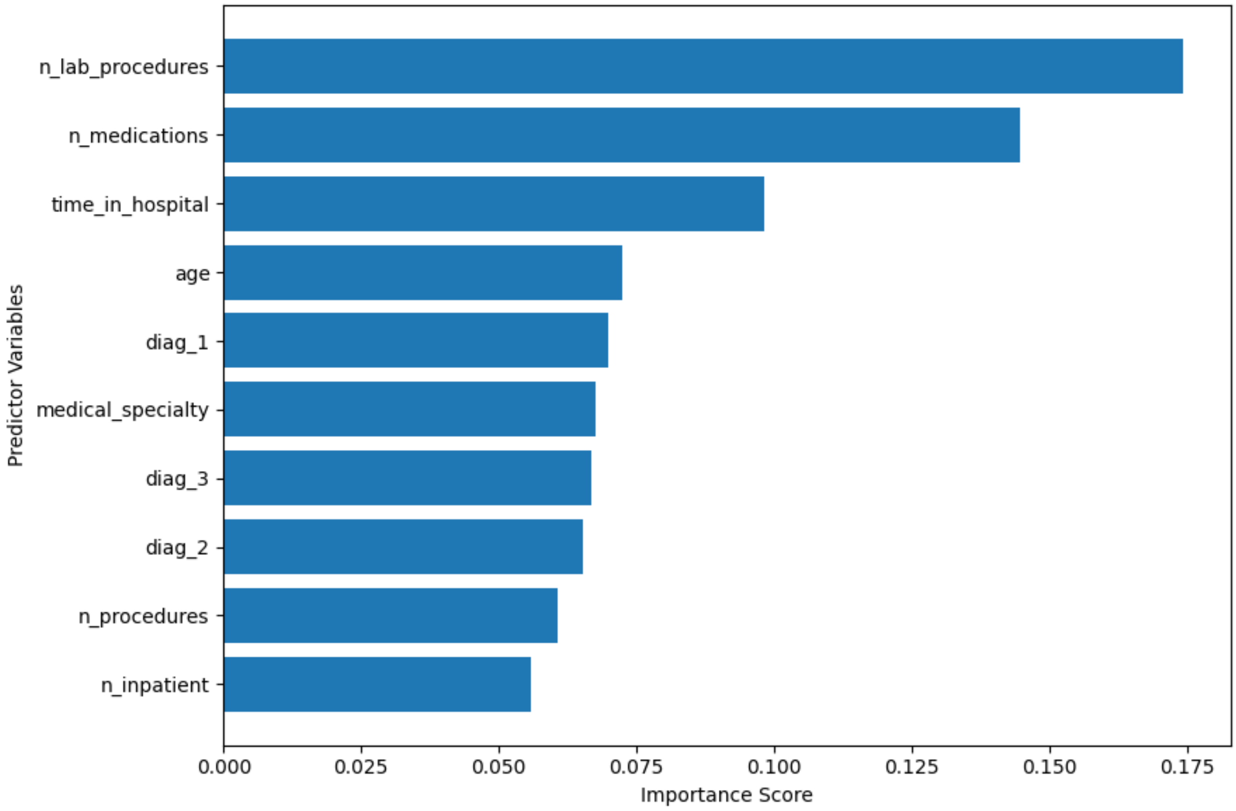

Figure 3. Top Predictors of Hospital Readmission Based on Random Forest Feature Importance

3.5 Visual Summary

The visual analysis complemented the statistical and predictive analyses of the study. The readmission distribution plot indicated that the study sample was well balanced, with more non-readmitted patients. This confirmed the numerical distribution in Table 2, in which 52.98% of patients were not readmitted and 47.02% were readmitted.

The clinical characteristics plot compared the mean values of some numerical variables by readmission status. This plot helped to compare readmitted and non-readmitted patients in terms of time in hospital, number of laboratory procedures, number of medications, number of inpatient visits, number of outpatient visits, and number of emergency visits. This is helpful because it can show the clinical and utilization variables that may lead to readmission. The feature importance plot indicated the number of laboratory procedures, number of medications, and time in hospital were the top predictors in the Random Forest model. This finding indicates that patients with more diagnostic procedures, more medications, and longer length of stay may warrant more attention after discharge. The presence of age, diagnosis variables, medical speciality, and prior hospital visits among the top predictors also confirms that the risk of readmission is part of patient complexity and history of hospital visits.

In summary, the figures showed that hospital readmission can be predicted using the current dataset, but the moderate performance measures indicate that future research should consider including other patient-related variables like patient's socioeconomic status, discharge type, insurance, comorbidity index, and follow-up arrangements to enhance prediction accuracy. The visual results also confirm that predictive analytics is not a theoretical concept, but a practical approach to identifying risk factors and improving health care through improved discharge and post-discharge planning.

4. Discussion

This study suggests that hospital readmission is driven by a combination of severity of illness, intensity of treatment, patient diagnosis categories and prior health care use. Random Forest was found to be the best model with F1-score, which indicates that ensemble-based machine learning approaches can be used to achieve a balance between readmitted and non-readmitted groups. The top predictors were number of laboratory procedures, number of medications, time in hospital, age, diagnosis categories, medical speciality, number of procedures and previous inpatient visits.

The significance of laboratory procedures may reflect that patients who need more diagnostic tests may have complex or uncertain diseasesn(Nguyen et al., 2018). Likewise, more medications may be needed to treat multiple diseases, or complex therapies. A longer length of stay may reflect serious illness, slow recovery or the need for more intensive care. Demographic factors such as age and diagnosis also predicted readmission, which suggests that patient characteristics and disease severity are key factors in the risk of readmission. Our results align with previous studies that have demonstrated that clinical factors, diagnosis, and electronic health record data can be used to predict hospital readmission risk (Liu et al., 2020; Matheny et al., 2021).

The findings of this study have significant implications for health care as they demonstrate that predictive analytics can assist in the early identification of patients at risk of readmission. By predicting the risk of readmission pre-discharge, hospitals can offer more care planning, medication reconciliation and patient education, and follow up after discharge. This may help avoid readmissions and ensure better continuity of care.

Predictive analytics can also help hospitals optimise their resources. Hospital staff may target patients with a high risk of readmission for discharge planning, case management, telephone follow-up, home care or referral to a specialist. This can help avoid the hospital burden and improve health outcomes. There has been growing awareness of the benefits of artificial intelligence and machine learning in diagnosis, prediction, treatment and decision making in health care (Jiang et al., 2017; Marafino et al., 2021).

The results of this research can be used in hospital decision support systems. A model that predicts a patient's risk of readmission based on age, length of stay, medication, laboratory test, diagnosis category, and prior hospital admission can warn clinicians of a patient's likelihood of readmission. This can assist clinicians in making informed decisions about discharge, scheduling appointments, adjusting medication doses and educating patients about their condition.

The model can also be applied to monitor patients in real-time, using electronic health records. As patient data is updated during the hospital stay, the system can update the patient's risk of readmission. This would enable earlier patient intervention and more tailored patient care. But while prediction performance is important for the successful use of artificial intelligence in clinical practice, it should also be timely, interpretable, and trusted by users (Kelly et al., 2019; Peng et al., 2023).

This study has certain limitations. First, it did not consider socio-economic factors such as income, education, occupation, family support, housing and follow-up care. These variables may affect the patient's decision to return to the hospital. Second, the dataset did not include hospital characteristics such as hospital size, staffing, discharge policies, treatment cost, insurance and access to follow-up care. These variables may prevent the models from explaining the data well.

Third, the dataset was cross-sectional and did not include longitudinal data. Post-discharge factors such as medication compliance, disease progression, follow-up appointments and home care may influence the risk of readmission. Fourth, the models had moderate rather than extremely high performance scores. This indicates that hospital readmission is a complex outcome that cannot be fully predicted by the variables considered. Such limitations have been observed in other studies of prediction in healthcare where the performance of the models is highly dependent on the nature of the data - such as its completeness and clinical validity (Adhiya et al., 2024; Ebied & Cooper-DeHoff, 2018).

Future studies should use multi-hospital datasets with larger sample sizes and greater diversity to better generalise the performance of prediction models. Data from multiple hospitals could be used to compare different health-care facilities, patient populations, and treatment regimes. Future research should also consider socio-economic, behavioural, insurance and post-discharge care factors to better understand the risk of readmission. Future studies can also incorporate deep learning methods to enhance prediction accuracy. Neural networks, recurrent models and medical code embedding approaches can model complex data and may be valuable for predicting readmission. Future studies should also include real-time clinical data from electronic health records, including laboratory trends, medication changes, discharge summaries, and follow-up records. Finally, future research should consider interpretability, fairness and clinical feasibility to ensure predictive analytics can be safely used for better clinical care.

5. Conclusion

The purpose of this study was to use predictive analytics on hospital readmission rates to inform health care. Data from 25,000 patients containing demographic, clinical, diagnostic, treatment and hospital utilization data was used. This analysis demonstrated that hospital readmission was a serious issue, with 47.02% of patients being readmitted. This high readmission rate suggests the need for a mechanism to identify patients who may need more care following discharge. The results of predictive modeling indicated that machine learning methods may be helpful in classifying patients based on risk of readmission. Random Forest was the best-performing model in terms of F1-score amongst the three models, which indicates that the model performed best in terms of the balance between precision and recall. The feature importance analysis revealed that the number of laboratory procedures, number of medications, time in hospital, age, diagnosis categories, medical speciality, number of procedures, and previous hospital visits are key factors associated with readmission. The results indicate that patients with increased disease complexity, number of medications, time spent in hospital and previous hospital stays are at higher risk of readmission. Hence, predictive modeling can assist health-care providers to identify patients at risk, optimise discharge planning, improve post-discharge care and resource allocation. The models had moderate performance, but the findings demonstrate the benefits of using data to inform decision-making to prevent avoidable readmissions and improve patient-centred care.

References

1. Adhiya, J., Barghi, B., & Azadeh-Fard, N. (2024). Predicting the risk of hospital readmissions using a machine learning approach: A case study on patients undergoing skin procedures. *Frontiers in Artificial Intelligence*, *6*, 1213378.
2. Awan, S. E., Sohel, F., Sanfilippo, F. M., Bennamoun, M., & Dwivedi, G. (2018). Machine learning in heart failure: Ready for prime time. *Current Opinion in Cardiology*, *33*(2), 190–195.
3. Burke, R. E., Schnipper, J. L., Williams, M. V., Robinson, E. J., Vasilevskis, E. E., Kripalani, S., Metlay, J. P., Fletcher, G. S., Auerbach, A. D., & Donzé, J. D. (2017). The HOSPITAL score predicts potentially preventable 30-day readmissions in conditions targeted by the hospital readmissions reduction program. *Medical Care*, *55*(3), 285–290.
4. Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F., & Sun, J. (2016). Doctor ai: Predicting clinical events via recurrent neural networks. *Machine Learning for Healthcare Conference*, 301–318. http://proceedings.mlr.press/v56/Choi16
5. Churpek, M. M., Yuen, T. C., Winslow, C., Meltzer, D. O., Kattan, M. W., & Edelson, D. P. (2016). Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Critical Care Medicine*, *44*(2), 368–374.
6. Davis, S. E., Lasko, T. A., Chen, G., Siew, E. D., & Matheny, M. E. (2017). Calibration drift in regression and machine learning models for acute kidney injury. *Journal of the American Medical Informatics Association*, *24*(6), 1052–1061.
7. Davis, S., Zhang, J., Lee, I., Rezaei, M., Greiner, R., McAlister, F. A., & Padwal, R. (2022). Effective hospital readmission prediction models using machine-learned features. *BMC Health Services Research*, *22*(1), 1415. https://doi.org/10.1186/s12913-022-08748-y
8. DD. (2023). *Predicting Hospital Readmissions*. https://www.kaggle.com/datasets/dubradave/hospital-readmissions
9. Dodson, J. A., Hajduk, A. M., Murphy, T. E., Geda, M., Krumholz, H. M., Tsang, S., Nanna, M. G., Tinetti, M. E., Goldstein, D., Forman, D. E., Alexander, K. P., Gill, T. M., & Chaudhry, S. I. (2019). Thirty-Day Readmission Risk Model for Older Adults Hospitalized With Acute Myocardial Infarction: The SILVER-AMI Study. *Circulation: Cardiovascular Quality and Outcomes*, *12*(5), e005320. https://doi.org/10.1161/CIRCOUTCOMES.118.005320
10. Ebied, A. M., & Cooper-DeHoff, R. M. (2018). 2017 is banner year for drug approvals by the Food and Drug Administration. *The American Journal of Medicine*, *131*(9), 1025–1033.
11. Gao, X., Alam, S., Shi, P., Dexter, F., & Kong, N. (2023). Interpretable machine learning models for hospital readmission prediction: A two-step extracted regression tree approach. *BMC Medical Informatics and Decision Making*, *23*(1), 104. https://doi.org/10.1186/s12911-023-02193-5

12. Huang, Y., Talwar, A., Lin, Y., & Aparasu, R. R. (2022). Machine learning methods to predict 30-day hospital readmission outcome among US adults with pneumonia: Analysis of the national readmission database. *BMC Medical Informatics and Decision Making*, *22*(1), 288. https://doi.org/10.1186/s12911-022-01995-3
13. Islam, Md. M., Poly, T. N., & Li, Y.-C. (2018). Recent Advancement of Clinical Information Systems: Opportunities and Challenges. *Yearbook of Medical Informatics*, *27*(01), 083–090. https://doi.org/10.1055/s-0038-1667075
14. Jamei, M., Nisnevich, A., Wetchler, E., Sudat, S., & Liu, E. (2017). Predicting all-cause risk of 30-day hospital readmission using artificial neural networks. *PloS One*, *12*(7), e0181173.
15. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, *2*(4). https://svn.bmj.com/Content/2/4/230.abstract
16. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*(1), 195. https://doi.org/10.1186/s12916-019-1426-2
17. Liu, W., Stansbury, C., Singh, K., Ryan, A. M., Sukul, D., Mahmoudi, E., Waljee, A., Zhu, J., & Nallamothu, B. K. (2020). Predicting 30-day hospital readmissions using artificial neural networks with medical code embedding. *PloS One*, *15*(4), e0221606.
18. Marafino, B. J., Escobar, G. J., Baiocchi, M. T., Liu, V. X., Plimier, C. C., & Schuler, A. (2021). Evaluation of an intervention targeted with predictive analytics to prevent readmissions in an integrated health system: Observational study. *Bmj*, *374*. https://www.bmj.com/content/374/bmj.n1747.abstract
19. Matheny, M. E., Ricket, I., Goodrich, C. A., Shah, R. U., Stabler, M. E., Perkins, A. M., Dorn, C., Denton, J., Bray, B. E., & Gouripeddi, R. (2021). Development of electronic health record–based prediction models for 30-day readmission risk among patients hospitalized for acute myocardial infarction. *JAMA Network Open*, *4*(1), e2035782.
20. Nguyen, O. K., Makam, A. N., Clark, C., Zhang, S., Das, S. R., & Halm, E. A. (2018). Predicting 30-Day Hospital Readmissions in Acute Myocardial Infarction: The AMI “READMITS” (Renal Function, Elevated Brain Natriuretic Peptide, Age, Diabetes Mellitus , Nonmale Sex , Intervention with Timely Percutaneous Coronary Intervention, and Low Systolic Blood Pressure) Score. *Journal of the American Heart Association*, *7*(8), e008882. https://doi.org/10.1161/JAHA.118.008882
21. Peng, H. T., Siddiqui, M. M., Rhind, S. G., Zhang, J., Da Luz, L. T., & Beckett, A. (2023). Artificial intelligence and machine learning for hemorrhagic trauma care. *Military Medical Research*, *10*(1), 6. https://doi.org/10.1186/s40779-023-00444-0