International Journal of Technology Research and Management
ISSN (Online): 2348-9006
Vol 11 Issue 12 December 2024

Study of various technique for Predicting Disease outcomes and Patient Readmission Rates

Aman Soni[1], Prof. Rakesh Shivhare[2]
Research Scholar[1], HOD[2]
Department of Computer Science & Engineering, Radharaman Engineering College, Bhopal, India[1, 2]
amansoni2696@gmail.com[1]

Abstract: *Predicting disease outcomes and patient readmission rates is a critical area of research in healthcare, aiming to improve patient care, reduce hospital burdens, and optimize resource allocation. This survey paper provides a comprehensive analysis of the various techniques employed for disease outcome prediction and readmission rate estimation. The study reviews a range of machine learning, deep learning and statistical methods, highlighting their strengths, limitations, and applications in diverse clinical scenarios. Key methods discussed include regression models, decision trees, support vector machines, ensemble learning and neural networks, alongside emerging approaches such as explainable AI and federated learning. The paper also examines data preprocessing techniques, feature selection methods, and evaluation metrics, emphasizing their role in enhancing predictive accuracy and reliability. Additionally, the survey explores challenges associated with data quality, class imbalance and interpretability, proposing potential solutions and areas for future research. By synthesizing findings from recent studies, this paper aims to provide a roadmap for researchers and practitioners seeking to advance predictive modeling in healthcare. The insights gained from this survey contribute to the development of robust and scalable models, ultimately supporting evidence-based clinical decision-making and improving patient outcomes.*

Keywords: *Machine Learning, Electronic Health Record, AI, NLP*.

1. INTRODUCTION

Healthcare predictive analytics has emerged as a powerful tool for transforming patient care and hospital operations. With the growing adoption of digital technologies in healthcare, vast amounts of data are generated daily from electronic health records (EHRs), wearable devices, medical imaging and other sources. These datasets offer an unprecedented opportunity to uncover patterns and trends that can support decision-making, improve patient outcomes, and enhance healthcare system efficiency. Among the cutting-edge approaches enabling this transformation is machine learning (ML), a branch of artificial intelligence that leverages algorithms to identify patterns, make predictions, and derive insights from complex data [1].

The application of ML in healthcare predictive analytics holds the potential to revolutionize key areas such as disease diagnosis, treatment planning, and resource optimization. By analyzing structured data, including demographics, laboratory results, and vital signs, as well as unstructured data like clinical notes and imaging, ML models can predict patient outcomes with remarkable accuracy. These capabilities are particularly valuable for tasks such as predicting the likelihood of disease progression, identifying high-risk patients, and forecasting readmission rates.

This study investigates the use of ML for healthcare predictive analytics, with a focus on its role in predicting disease outcomes and patient readmissions. These tasks are crucial for improving patient management, reducing healthcare costs and ensuring the efficient allocation of resources. Various ML techniques, including supervised, unsupervised, and reinforcement learning, are explored for their ability to handle the complexities of healthcare data. Additionally, advancements in natural language processing (NLP) and deep learning are highlighted as key enablers for deriving insights from unstructured data.

However, the adoption of ML in healthcare is not without challenges. Issues such as data privacy, model interpretability and biases in training data require careful consideration. This study also emphasizes the importance of developing explainable AI (XAI) solutions to foster trust and acceptance among healthcare professionals.

By examining the capabilities and limitations of ML in healthcare predictive analytics, this work aims to provide a comprehensive understanding of its potential to reshape the healthcare landscape. It highlights opportunities for innovation, identifies barriers to implementation, and outlines strategies for integrating ML-driven analytics into clinical practice. As the healthcare sector continues to evolve, leveraging ML for predictive analytics will play a pivotal role in enabling data-driven, patient-centered care.

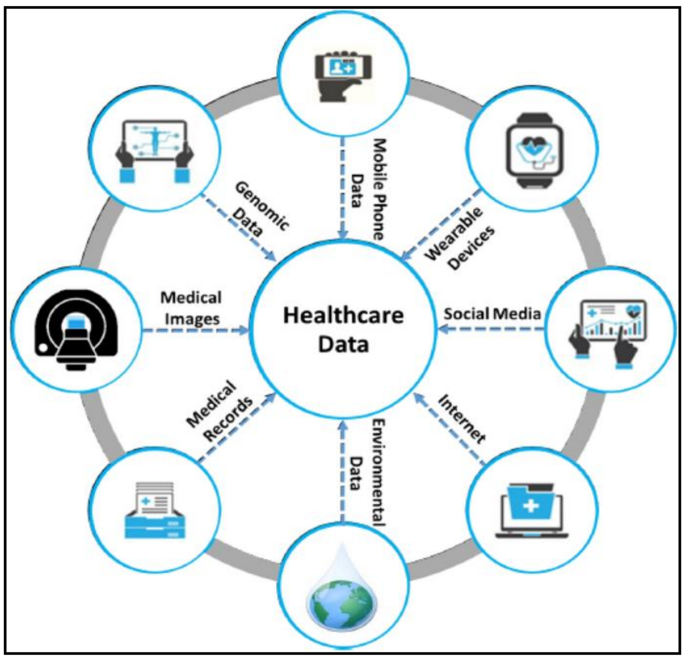

Figure 1: Illustration of heterogeneous sources contributing to healthcare data [2]

Healthcare services generate vast amounts of data daily, encompassing electronic health records, genomics, environmental factors, social media insights, and other sources. The complexity and volume of this data make traditional analysis methods insufficient. Machine learning (ML) and deep learning (DL) provide powerful tools to analyze and interpret this information effectively, transforming raw data into actionable insights. By leveraging these advanced techniques, healthcare professionals can enhance decision-making and patient care [2].

Incorporating diverse data sources, such as medical records, genomic data, and environmental data, further enriches the analytical process, enabling more accurate and personalized predictions. These approaches support four critical applications in healthcare: prognosis, diagnosis, therapy, and clinical workflow optimization. Each of these domains benefits significantly from the predictive and interpretive power of ML and DL, as illustrated in Figure 1, which provides an overview of these data sources. These advancements mark a transformative step in modern healthcare systems.

Patient readmission rates are a critical metric in evaluating the quality of healthcare services and the effectiveness of clinical interventions. Unplanned readmissions within a short period after discharge are often indicative of suboptimal care, inadequate post-discharge planning, or insufficient patient support. These events not only compromise patient outcomes but also impose significant financial burdens on healthcare systems. Consequently, reducing readmission rates has become a priority for healthcare providers, policymakers, and payers alike [3].

The growing adoption of electronic health records (EHRs) offers an opportunity to address this challenge through data-driven approaches. EHRs contain rich, multifaceted data, including patient demographics, diagnoses, treatments, medications, lab results, and clinical notes. Leveraging this information, advanced machine learning (ML) techniques can be employed to predict patient readmission risks, allowing healthcare providers to intervene proactively.

This study explores the use of ML to analyze EHR data and develop predictive models for identifying patients at high risk of readmission. By utilizing structured data, such as medical histories and treatment plans, and unstructured data, such as physician notes, ML models can uncover patterns that traditional statistical methods often miss. Predictive models can also incorporate time-series data to track patient health trends, enhancing their ability to forecast readmissions accurately.

While the potential benefits of ML-driven predictions are substantial, challenges such as data privacy, algorithm interpretability, and biases in training data must be addressed to ensure equitable and effective solutions. Furthermore, the integration of these models into existing healthcare workflows is essential for their successful adoption.

2. LITERATURE SURVEY

The application of machine learning and predictive analytics in healthcare has led to significant advancements in predicting disease outcomes and patient readmission rates. The following studies provide insights into methodologies and findings in this domain.

Shameer et al. (2017) [1] developed predictive models for hospital readmissions using the Mount Sinai heart failure cohort. Their study employed electronic medical record (EMR)-wide machine learning techniques, highlighting the value of integrating diverse EMR data to improve prediction accuracy. By analyzing comprehensive datasets, they demonstrated how combining clinical, demographic, and treatment information enhances the performance of readmission prediction models. The study emphasizes the importance of utilizing extensive EMR data to better predict heart failure outcomes and improve healthcare decision-making, ultimately aiding in the prevention of avoidable readmissions.

Weiskopf et al. (2017) [2] addressed the quality of electronic medical record (EMR) data reuse, proposing a data quality assessment guideline to ensure reliable predictive analytics. The study emphasized the importance of robust data preprocessing and validation for achieving high accuracy in predictive models. By outlining key steps for data quality assessment, the framework provides a foundational approach for integrating machine learning with clinical data. This guideline supports the development of effective predictive models and helps ensure the reliability and usability of EMR data in healthcare applications, facilitating more accurate clinical decision-making.

Smith et al. (2018) [3] conducted a systematic review of risk prediction models for acute myocardial infarction readmissions. The review examined model performance metrics and identified significant variability in outcomes, attributed to differences in datasets and methodologies. Their findings emphasized the importance of adopting standardized approaches to improve model reliability and generalizability. Additionally, they highlighted the necessity of incorporating diverse patient cohorts to ensure broader applicability of the models. The study underscored the need for consistency in methods and data sources to enhance the predictive power and accuracy of readmission risk models in clinical settings.

Mahajan and Ghani (2019) [4] investigated the application of ensemble machine learning methods for predicting heart failure readmission risk. Their study showcased the advantages of ensemble techniques over traditional models, as they combined multiple algorithms to enhance prediction accuracy. By leveraging the strengths of different models, this approach improved the robustness of predictions, particularly in complex healthcare datasets. The study highlighted how ensemble learning can effectively manage the intricate and varied nature of healthcare data, demonstrating its potential to improve clinical decision-making and readmission risk predictions for heart failure patients.

Badawy et al. (2023) [5] presented a comprehensive survey on healthcare predictive analytics, highlighting the roles of machine learning and deep learning techniques. The survey explored various methods and their applications in predicting disease outcomes, including readmission rates. It emphasized the increasing adoption of deep learning models due to their ability to process large-scale, unstructured data, such as electronic medical records (EMRs). Additionally, the study noted the flexibility of deep learning in adapting to diverse healthcare scenarios, positioning it as a key approach for improving predictive accuracy and decision-making in healthcare settings.

Golas et al. (2018) [6] conducted a retrospective analysis using electronic health record (EHR) data to develop a machine learning model for predicting 30-day readmission risk in heart failure patients. By integrating demographic, clinical, and medication data, the study demonstrated how incorporating feature-rich datasets can enhance the reliability of predictions. The research highlighted the potential of machine learning to deliver timely, actionable insights for healthcare providers, thereby enabling more informed decision-making and improving patient outcomes. This approach underscores the value of comprehensive data in predictive analytics for heart failure management.

Rojas et al. (2018) [7] investigated the prediction of intensive care unit (ICU) readmissions using machine learning models applied to electronic health record (EHR) data. Their study compared various machine learning algorithms and found that techniques like random forests and gradient boosting outperformed traditional statistical methods. The research emphasized the importance of real-time data usage and highlighted the critical role of feature selection in improving the accuracy of ICU readmission predictions. These findings underscore the potential of machine learning to enhance decision-making in critical care settings and optimize patient management.

Ashfaq et al. (2019) [8] examined the use of deep learning for predicting readmissions using electronic health records (EHRs). Their study highlighted the effectiveness of deep learning models, especially recurrent neural networks (RNNs), in capturing temporal patterns and complex relationships within longitudinal patient data. The research emphasized the ability of deep learning techniques to handle large-scale and unstructured datasets, offering significant advantages over traditional methods. By leveraging these models, the study demonstrated the potential for improved prediction accuracy in readmission risks, particularly in managing complex healthcare data.

Mahmoudi et al. (2020) [9] conducted a systematic review on the use of electronic health records (EHRs) for developing and validating hospital readmission prediction models. The review identified challenges such as data quality issues, population heterogeneity, and the need for external validation. It emphasized the importance of addressing these obstacles to improve the reliability and generalizability of machine learning-based prediction models. The study provided valuable insights into the current state of research and offered recommendations for future work aimed at enhancing the robustness and applicability of readmission prediction models in diverse healthcare settings.

Chi et al. (2021) [10] developed a machine learning model to predict mortality and readmission in patients experiencing in-hospital cardiac arrest. By incorporating a broad set of clinical and physiological features, the model effectively stratified patients according to their risk levels. The study demonstrated the potential of advanced machine learning algorithms in critical care, emphasizing their role in identifying high-risk patients and enabling timely, targeted interventions. This approach not only aims to improve patient survival rates but also reduces the likelihood of readmissions, highlighting the value of predictive analytics in enhancing outcomes in critical care settings.

Zhao et al. (2021) [11] investigated the variability in readmission rates following total hip and knee arthroplasty across different datasets. Their findings revealed inconsistencies due to the use of disparate data sources, emphasizing the necessity of standardized methodologies to achieve accurate readmission predictions. The study highlighted the importance of data harmonization to improve the reliability of machine learning applications in healthcare. By addressing these challenges, the research advocates for unified approaches that enhance the consistency and accuracy of predictive models, ultimately supporting better clinical decision-making and patient outcomes.

Matheny et al. (2021) [12] developed electronic health record (EHR)-based prediction models to assess the 30-day readmission risk for patients hospitalized due to acute myocardial infarction. By integrating advanced machine learning techniques with patient demographics, clinical data, and comorbidities, the study demonstrated the effectiveness of these models in providing robust predictions. Their work highlights the potential of machine learning in cardiology, particularly for improving patient outcomes and optimizing resource allocation. This approach showcases how ML can support clinical decision-making and enhance the management of patients with complex cardiovascular conditions.

Okere et al. (2021) [13] applied machine learning to identify risk factors for 30-day readmission and 180-day in-hospital mortality in ischemic heart disease patients. The study ranked these factors by their relative importance, offering valuable insights for targeted healthcare interventions. By leveraging feature importance analysis in ML models, the research emphasized the potential to prioritize risk factors, enabling clinicians to focus on the most critical elements influencing patient outcomes. This approach enhances the predictive power of models, ultimately supporting more effective and personalized healthcare strategies for managing ischemic heart disease patients.

Najafi-Vosough et al. (2021) [14] compared multiple machine learning methods for predicting hospital readmissions among heart failure patients in Iran. Their results showed that ensemble techniques, like random forests and gradient boosting, outperformed traditional methods, highlighting the superior predictive accuracy of ML models. The study also underscored the ability of these models to adapt across diverse healthcare systems, demonstrating their potential to address regional healthcare challenges. This research emphasizes the value of machine learning in improving readmission predictions and enhancing healthcare management, especially in regions with unique healthcare dynamics and resource constraints.

Lv et al. (2021) [15] applied machine learning-driven models to predict prognostic outcomes for heart failure patients using EHR data. Their retrospective study incorporated a broad set of patient-specific features, highlighting how ML can enhance the accuracy of readmission and mortality predictions. The research demonstrated that machine learning could identify critical patterns in patient data, leading to more accurate predictions. Furthermore, it emphasized the importance of retrospective analysis in clinical decision-making, helping healthcare providers make informed decisions and potentially improving patient outcomes by identifying at-risk individuals early in the care process.

Van Grootven et al. (2021) [16] conducted a systematic review and meta-analysis of prediction models for hospital readmissions in heart disease patients. They found significant variability in model performance, which could be attributed to differences in datasets, feature selection, and methodologies. The study underscored the need for standardized evaluation frameworks to ensure more reliable comparisons and broader applicability in clinical practice. By highlighting these challenges, the research emphasized the importance of developing uniform approaches to model development and assessment, which could improve the consistency and effectiveness of readmission prediction models in healthcare settings.

Mohamed et al. (2022) [17] applied machine learning algorithms to predict readmissions in patients with chronic obstructive pulmonary disease (COPD). Their study demonstrated that ensemble methods, like random forests and gradient boosting, significantly improved predictive accuracy compared to traditional models. By integrating a wide array of data, including demographic, clinical, and behavioral factors, the research highlighted the importance of a comprehensive approach to risk assessment. This work emphasized how combining diverse data sources with advanced ML techniques can enhance prediction models, leading to better patient management and targeted interventions in COPD care.

Sharma et al. (2022) [18] explored using administrative data to predict 30-day readmissions in heart failure patients. By applying various machine learning models, their study achieved high accuracy, with ensemble methods outperforming traditional approaches. The research highlighted the potential of administrative datasets as a cost-effective alternative to electronic health records (EHRs) for predictive modeling. This work demonstrates that even with less detailed data, machine learning can still produce reliable predictions, offering a practical solution for healthcare systems with limited access to comprehensive EHR data, ultimately improving patient management and resource allocation.

Le Lay et al. (2022) [19] investigated predicting hospital readmissions in multimorbid patients using machine learning models. The study incorporated a wide range of patient data, including comorbidities, medications, and socioeconomic factors, achieving strong predictive performance. Their research highlighted the capability of machine learning to handle the complexities of multimorbidity, offering valuable insights for healthcare planning. By integrating diverse factors that influence patient outcomes, the study demonstrated that ML models could enhance decision-making and resource allocation, ultimately improving care for patients with multiple chronic conditions.

Mohamed et al. (2022) [20] revisited the prediction of COPD readmissions, using advanced machine learning techniques to assess feature importance and model interpretability. Their study highlighted the effectiveness of ML in identifying high-risk patients, showcasing its potential in improving predictive accuracy. The research also emphasized the necessity of explainable models to ensure their clinical adoption, making it easier for healthcare professionals to trust and apply the predictions in decision-making. By focusing on both model performance and transparency, the study contributed to enhancing the practical use of ML in healthcare settings for COPD readmission prediction.

Wang and Zhu (2022) [21] focused on predicting 30-day hospital readmissions by developing a disease-specific model using nationwide hospital admission data. Their study aimed to enhance the accuracy of readmission predictions by incorporating both general hospital admission data and disease-specific variables. The researchers demonstrated that integrating these contextual variables significantly improved the model's predictive performance, making it more robust and tailored to specific patient populations. The inclusion of disease-specific data allowed the model to account for variations in readmission risk based on the underlying health condition, ultimately leading to more accurate predictions. This research emphasizes the importance of contextualizing predictive models to better reflect the complexities of individual diseases and patient demographics. The findings also suggest that the combination of general hospital data and disease-specific factors can offer more reliable insights, providing healthcare providers with improved tools to manage readmission risks and optimize patient care strategies.

Teo et al. (2023) [22] presented an extensive review of the latest trends in hospital readmission prediction approaches. The study examined a wide range of machine learning (ML) techniques, comparing classical statistical methods with advanced deep learning models, and discussed their applicability across diverse healthcare settings. The authors emphasized the growing integration of electronic health records (EHRs) with real-time data, which enhances the accuracy and timeliness of readmission predictions. Additionally, the study highlighted the increasing importance of model interpretability and transparency in ML applications. As healthcare professionals are more likely to trust and adopt predictive models when they can understand how the model generates its results, the research underscored the need for developing explainable models. By focusing on these aspects, the study pointed out the potential for improving clinical decision-making, patient outcomes, and overall trust in machine learning-driven healthcare technologies.

Sabouri et al. (2023) [23] utilized machine learning to predict both readmission and mortality in heart failure patients. Their work applied a range of algorithms, including decision trees, support vector machines, and neural networks, and found that ensemble methods provided superior performance. This research emphasized the dual role of predictive models in both mortality and readmission prediction, suggesting that these models could be valuable tools for improving patient management strategies.

Ruppert et al. (2023) [24] performed a systematic review of predictive models for readmission to intensive care units (ICUs). Their analysis of various machine learning (ML) approaches emphasized the significance of feature selection in model performance, identifying key factors such as age, comorbidities, and clinical markers. The study highlighted the importance of these features in improving the accuracy of readmission predictions. By focusing on these critical factors, the research underscored the potential of ML to enhance decision-making in ICU settings, where precise predictions are essential for efficient resource allocation and patient care optimization.

Talwar et al. (2023) [25] performed a meta-analysis comparing advanced ML algorithms with logistic regression for predicting hospital readmissions. Their findings suggested that models such as gradient boosting and random forests outperformed logistic regression in terms of both accuracy and generalizability. This meta-analysis underscored the need for selecting the right ML algorithm based on the nature of the data and the clinical problem.

Ghasemieh et al. (2023) [26] developed a novel machine learning model using a stacking ensemble learner to predict emergency readmissions in heart disease patients. By combining the strengths of multiple models, they achieved high prediction accuracy. The study emphasized the effectiveness of ensemble techniques in handling the complexity and variability inherent in predicting readmissions for a diverse patient population. Their approach demonstrated how integrating different models can enhance the robustness of predictions, making it a valuable tool for clinical decision-making and improving patient care outcomes in heart disease management.

Nair et al. (2024) [27] introduced a machine learning-based approach for predicting readmissions in heart failure patients, focusing on evaluating its real-world impact through a quasi-experimental study design. Their research protocol aimed to assess the effectiveness of integrating electronic health record (EHR)-based machine learning models into clinical practice. This approach not only involved the development of predictive models but also considered the broader implications of adopting such models in healthcare settings. By assessing how the model performs when deployed in real clinical environments, the study aimed to address concerns about model accuracy, usability, and potential barriers to adoption. The authors emphasized the significance of understanding the model's real-world impact on patient outcomes and clinical workflows. This work underscores the need for rigorous evaluation to ensure that machine learning models are not only accurate but also effective and sustainable when integrated into day-to-day clinical decision-making processes.

Huberts et al. (2024) [28] emphasized the importance of explainable predictive analytics for cardiovascular patient readmission and mortality. Their research highlighted the necessity of interpretability in machine learning (ML) models, ensuring that clinicians can understand and trust the predictions generated by these algorithms. By focusing on transparency, the study aimed to enhance the practical application of ML models in healthcare settings. This work contributes to the growing body of research advocating for transparent and actionable insights, ultimately supporting clinicians in making informed decisions based on reliable predictive models.

Park et al. (2024) [29] developed a predictive model for 90-day readmissions following total joint arthroplasty by integrating electronic health records (EHRs) and patient-reported outcome measures (PROMs). Their study showed that incorporating PROMs significantly improved the accuracy of readmission predictions in post-surgical patients. By combining clinical data with subjective patient-reported metrics, the model was able to provide more precise risk assessments. This research highlights the value of including diverse data sources, such as PROMs, to enhance the predictive power of machine learning models in surgical patient care.

Zeinalnezhad and Shishehchi (2024) [30] combined data mining algorithms with meta-heuristic techniques to predict readmission risk in diabetic patients. Their innovative approach sought to enhance the performance of predictive models by tackling the complexity and variability of diabetes-related data. By integrating advanced techniques, the study aimed to improve the accuracy of readmission predictions, contributing to the development of more personalized medicine. This research highlights the growing potential of predictive analytics in managing chronic conditions like diabetes, ensuring better-targeted interventions and improving patient outcomes.

Zarghani (2024) [31] conducted a comparative analysis between Long Short-Term Memory (LSTM) neural networks and traditional machine learning models, such as decision trees and logistic regression, for predicting diabetes patient readmission. The study highlighted the superiority of LSTM networks in capturing temporal dependencies in patient data, which traditional models often fail to account for. LSTM's ability to process sequences of patient health records over time was shown to improve prediction performance significantly compared to conventional models. This work contributed to the growing body of literature advocating for the use of deep learning techniques in patient readmission prediction.

Zhang et al. (2024) [32] presented an updated review of predictive modeling for hospital readmissions in heart disease patients from 2012 to 2023. Their comprehensive review emphasized the evolution of machine learning models, noting a shift from simpler statistical techniques to more complex models such as random forests, support vector machines (SVMs), and deep learning algorithms. They also discussed the integration of multimodal data sources, such as clinical data, imaging, and genetic information, which have further enhanced the predictive capabilities of these models. Their review underscores the significant improvements in model accuracy over the years, owing to advances in both data availability and algorithm sophistication.

Da Silva et al. (2024) [33] explored the application of machine learning for predicting hospital readmissions in the pediatric population. This study addressed the unique challenges of pediatric data, which often involve fewer historical records and more variability in treatment protocols compared to adult populations. The researchers applied ensemble methods and deep learning models to pediatric readmission prediction, with results indicating that specialized models tailored to pediatric care could yield improved prediction accuracy compared to general models used in adult populations.

Adeniran et al. (2024) [34] discussed the broader impact of data-driven decision-making in healthcare, focusing on the role of predictive modeling in improving patient outcomes. Their study explored various machine learning techniques, such as decision trees and neural networks, and emphasized the importance of integrating EHR data with real-time patient monitoring systems. They argued that predictive models could significantly enhance clinical decision-making by providing timely insights into patient risks, ultimately reducing readmissions and improving care quality.

Gaso et al. (2024) [35] examined the use of machine learning and deep learning techniques for predicting readmission cases in diabetes patients. Their study compared various models, including neural networks and gradient boosting machines, demonstrating that deep learning models, when trained on large EHR datasets, offered superior accuracy. This research is particularly relevant as diabetes is a prevalent condition with high readmission rates, and better predictive models could lead to more effective management of diabetic patients and prevent avoidable hospitalizations.

Tsai et al. (2024) [36] utilized machine learning to analyze monthly blood test data for forecasting 30-day hospital readmissions among maintenance hemodialysis patients. By leveraging time-series data from routine blood tests, the study demonstrated the potential of using continuous clinical monitoring to predict readmissions. The research highlighted how integrating regularly collected data, such as lab results, into predictive models can improve both the accuracy and timeliness of readmission predictions. This approach has significant implications for healthcare systems, as it enables more proactive interventions based on real-time data, potentially reducing readmission rates. By focusing on a vulnerable patient population, maintenance hemodialysis patients, the study also underlined the importance of personalized care and the utility of predictive analytics in improving outcomes for patients with chronic conditions. This innovative use of ongoing clinical data further strengthens the case for integrating real-time health data into healthcare decision-making processes.

Hu et al. (2025) [37] developed and validated a machine learning model specifically designed to predict one-year readmissions in patients with Heart Failure with preserved Ejection Fraction (HFpEF). Their model incorporated a wide range of clinical features, including comorbidities, lab results, and treatment history, and was able to predict readmission risk with high accuracy. The study contributed valuable insights into the specific needs of HFpEF patients, whose clinical outcomes can be challenging to predict due to the heterogeneous nature of the disease.

Li et al. (2025) [38] focused on predicting ICU readmissions for patients with intracerebral hemorrhage using the MIMIC III and IV databases. Their study emphasized the value of large, high-quality clinical databases in training predictive models. The researchers demonstrated that ICU-specific models could enhance patient care by identifying high-risk individuals early, enabling timely interventions. This approach underscores the potential of utilizing comprehensive clinical datasets to improve prediction accuracy and ultimately reduce readmission rates, thereby optimizing ICU resource allocation and improving patient outcomes. Their work contributes to the growing use of data-driven strategies in critical care settings.

Buddhiraju et al. (2025) [39] compared the prediction accuracy for 30-day readmission following primary total knee arthroplasty using the ACS-NSQIP risk calculator and a novel artificial neural network model. Their findings showed that the neural network model outperformed the traditional risk calculator in predicting readmissions. This research highlights the potential of advanced machine learning models to enhance postoperative care, offering more accurate predictions for orthopedic patients. By improving prediction accuracy, these models could help reduce readmission rates and optimize resource allocation, ultimately contributing to better patient outcomes in orthopedic surgery.

Purbasari et al. (2024) [40] applied machine learning in conjunction with hyperparameter optimization using Bayesian techniques to predict readmission risk following total hip arthroplasty. The study demonstrated that optimizing hyperparameters significantly improved the performance of predictive models, leading to more accurate predictions of readmission risk. By leveraging Bayesian methods for hyperparameter tuning, the researchers were able to enhance the model's ability to capture complex patterns in patient data. This research emphasizes the critical role of hyperparameter optimization in machine learning, particularly in healthcare, where accurate predictions are essential for improving patient outcomes. The study highlights how refining model parameters can lead to better predictive accuracy, ultimately supporting healthcare providers in making informed decisions, reducing readmission rates, and ensuring efficient resource allocation.

3. CONCLUSION

In conclusion, recent studies highlight the ongoing advancement of machine learning techniques in predicting hospital readmissions. A key trend is the development of disease-specific models and the increasing use of deep learning approaches, which enable more accurate and context-sensitive predictions. Moreover, the integration of real-time health data, such as electronic health records and patient-reported outcomes, has enhanced the timeliness and effectiveness of these models. As machine learning models continue to evolve, they hold significant potential to improve patient outcomes, optimize resource allocation, and reduce readmissions across diverse patient populations. However, challenges remain in integrating heterogeneous data sources, ensuring model transparency, and assessing their real-world impact. Future research should focus on refining predictive models, enhancing interpretability, and conducting rigorous evaluations in clinical settings to better understand their practical implications and further support healthcare decision-making. These advancements could ultimately lead to more personalized and efficient care for patients.

REFERENCES

[1] Shameer, Khader, Kipp W. Johnson, Alexandre Yahi, Riccardo Miotto, L. I. Li, Doran Ricks, Jebakumar Jebakaran et al. "Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: a case-study using Mount Sinai heart failure cohort." In Pacific symposium on biocomputing 2017, pp. 276-287. 2017.

[2] Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. EGEMS (Wash DC) 2017;5:14.

[3] L.N. Smith, A.N. Makam, D. Darden, H. Mayo, S.R. Das, E.A. Halm, O.K. Nguyen, Acute myocardial infarction readmission risk prediction models: a systematic review of model performance, Circ.: Cardiovasc. Qual. Outcomes 11 (1) (2018) e003885.

[4] S. Mahajan, R. Ghani, Predicting risk of readmission for heart failure using ensemble machine learning methods, Stud. Health Technol. Inform. 264 (2019) 243–247.

[5] Mohammed Badawy, Nagy Ramadan and Hesham Ahmed Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: a survey", Journal of Electrical Systems and Inf Technol (2023).

[6] Golas, Sara Bersche, Takuma Shibahara, Stephen Agboola, Hiroko Otaki, Jumpei Sato, Tatsuya Nakae, Toru Hisamitsu et al. "A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data." BMC medical informatics and decision making 18 (2018): 1-17.

[7] Rojas, Juan C., Kyle A. Carey, Dana P. Edelson, Laura R. Venable, Michael D. Howell, and Matthew M. Churpek. "Predicting intensive care unit readmission with machine learning using electronic health record data." Annals of the American Thoracic Society 15, no. 7 (2018): 846-853.

[8] Ashfaq, Awais, Anita Sant'Anna, Markus Lingman, and Slawomir Nowaczyk. "Readmission prediction using deep learning on electronic health records." Journal of biomedical informatics 97 (2019): 103256.

[9] Mahmoudi, Elham, Neil Kamdar, Noa Kim, Gabriella Gonzales, Karandeep Singh, and Akbar K. Waljee. "Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review." bmj 369 (2020).

[10] Chi, Chien-Yu, Shuang Ao, Adrian Winkler, Kuan-Chun Fu, Jie Xu, Yi-Lwun Ho, Chien-Hua Huang, and Rohollah Soltani. "Predicting the mortality and readmission of in-hospital cardiac arrest patients with electronic health records: a machine learning approach." Journal of medical Internet research 23, no. 9 (2021): e27798.

[11] Zhao S, Kendall J, Johnson AJ, Sampson AAG, Kagan R. Disagreement in readmission rates after total hip and knee arthroplasty across data sets. Arthroplasty Today 2021; 9:73e7.

[12] M.E. Matheny, I. Ricket, C.A. Goodrich, R.U. Shah, M.E. Stabler, A.M. Perkins, C. Dorn, J. Denton, B.E. Bray, R. Gouripeddi, et al., Development of electronic health record-based prediction models for 30-day readmission risk among patients hospitalized for acute myocardial infarction, JAMA Netw. Open 4 (2021) e2035782.

[13] A. Okere, V. Sanogo, H. Alqhtani, V. Diaby, Identification of risk factors of 30- day readmission and 180-day in-hospital mortality, and its corresponding relative importance in patients with Ischemic heart disease: a machine learning approach, Expert Rev. Pharmacoecon. Outcomes Res. 21 (2021) 1043–1048.

[14] R. Najafi-Vosough, J. Faradmal, S. Hosseini, A. Moghimbeigi, H. Mahjub, Predicting hospital readmission in heart failure patients in Iran: A comparison of various machine learning methods, Healthc. Inform. Res. 27 (2021) 307–314.

[15] Lv, Haichen, Xiaolei Yang, Bingyi Wang, Shaobo Wang, Xiaoyan Du, Qian Tan, Zhujing Hao, Ying Liu, Jun Yan, and Yunlong Xia. "machine learning–driven models to predict prognostic outcomes in patients hospitalized with heart failure using electronic health records: retrospective study." Journal of medical Internet research 23, no. 4 (2021): e24996.

[16] Van Grootven, Bastiaan, Patricia Jepma, Corinne Rijpkema, Lotte Verweij, Mariska Leeflang, Joost Daams, Mieke Deschodt, Koen

Milisen, Johan Flamaing, and Bianca Buurman. "Prediction models for hospital readmissions in patients with heart disease: a systematic review and meta-analysis." BMJ open 11, no. 8 (2021): e047576.

[17] Mohamed, Israa, Mostafa M. Fouda, and Khalid M. Hosny. "Machine learning algorithms for COPD patients readmission prediction: A data analytics approach." IEEE Access 10 (2022): 15279-15287.

[18] V. Sharma, V. Kulkarni, F. McAlister, D. Eurich, S. Keshwani, S. Simpson, D. Voaklander, S. Samanani, Predicting 30-day readmissions in patients with heart failure using administrative data: A machine learning approach, J. Card. Fail. (2022).

[19] Le Lay, Jules, Edgar Alfonso-Lizarazo, Vincent Augusto, Bienvenu Bongue, Malek Masmoudi, Xiaolan Xie, Baptiste Gramont, and Thomas Célarier. "Prediction of hospital readmission of multimorbid patients using machine learning models." PLoS One 17, no. 12 (2022): e0279433.

[20] Mohamed, Israa, Mostafa M. Fouda, and Khalid M. Hosny. "Machine learning algorithms for COPD patients readmission prediction: A data analytics approach." IEEE Access 10 (2022): 15279-15287.

[21] Wang, Shuwen, and Xingquan Zhu. "Nationwide hospital admission data statistics and disease-specific 30-day readmission prediction." Health Information Science and Systems 10, no. 1 (2022): 25.

[22] Teo, Kareen, Ching Wai Yong, Joon Huang Chuah, Yan Chai Hum, Yee Kai Tee, Kaijian Xia, and Khin Wee Lai. "Current trends in readmission prediction: an overview of approaches." Arabian journal for science and engineering 48, no. 8 (2023): 11117-11134.

[23] Sabouri, Maziar, Ahmad Bitarafan Rajabi, Ghasem Hajianfar, Omid Gharibi, Mobin Mohebi, Atlas Haddadi Avval, Nasim Naderi, and Isaac Shiri. "Machine learning based readmission and mortality prediction in heart failure patients." Scientific Reports 13, no. 1 (2023): 18671.

[24] Ruppert, Matthew M., Tyler J. Loftus, Coulter Small, Han Li, Tezcan Ozrazgat-Baslanti, Jeremy Balch, Reed Holmes et al. "Predictive modeling for readmission to intensive care: a systematic review." Critical Care Explorations 5, no. 1 (2023): e0848.

[25] Talwar, Ashna, Maria A. Lopez-Olivo, Yinan Huang, Lin Ying, and Rajender R. Aparasu. "Performance of advanced machine learning algorithms overlogistic regression in predicting hospital readmissions: A meta-analysis." Exploratory Research in Clinical and Social Pharmacy 11 (2023): 100317.

[26] Ghasemieh, Alireza, Alston Lloyed, Parsa Bahrami, Pooyan Vajar, and Rasha Kashef. "A novel machine learning model with Stacking Ensemble Learner for predicting emergency readmission of heart-disease patients." Decision Analytics Journal 7 (2023): 100242.

[27] Nair, Monika, Lina E. Lundgren, Amira Soliman, Petra Dryselius, Ebba Fogelberg, Marcus Petersson, Omar Hamed, Miltiadis Triantafyllou, and Jens Nygren. "Machine Learning Model for Readmission Prediction of Patients With Heart Failure Based on Electronic Health Records: Protocol for a Quasi-Experimental Study for Impact Assessment." JMIR Research Protocols 13, no. 1 (2024): e52744.

[28] Huberts, Leo CE, Sihan Li, Victoria Blake, Louisa Jorm, Jennifer Yu, Sze-Yuan Ooi, and Blanca Gallego. "Predictive analytics for cardiovascular patient readmission and mortality: An explainable approach." Computers in Biology and Medicine 174 (2024): 108321.

[29] Park, Jaeyoung, Xiang Zhong, Emilie N. Miley, Rachel S. Rutledge, Jaquelyn Kakalecik, Matthew C. Johnson, and Chancellor F. Gray. "Machine Learning-Based Predictive Models for 90-Day Readmission of Total Joint Arthroplasty Using Comprehensive Electronic Health Records and Patient-Reported Outcome Measures." Arthroplasty Today 25 (2024): 101308.

[30] Zeinalnezhad, Masoomeh, and Saman Shishehchi. "An integrated data mining algorithms and meta-heuristic technique to predict the readmission risk of diabetic patients." Healthcare Analytics 5 (2024): 100292.

[31] Zarghani, Abolfazl. "Comparative Analysis of LSTM Neural Networks and Traditional Machine Learning Models for Predicting Diabetes Patient Readmission." arXiv preprint arXiv:2406.19980 (2024).

[32] Zhang, Wei, Weihan Cheng, Koichi Fujiwara, Richard Evans, and Chengyan Zhu. "Predictive Modeling for Hospital Readmissions for Patients with Heart Disease: An updated review from 2012-2023." IEEE Journal of Biomedical and Health Informatics (2024).

[33] Da Silva, Nayara Cristina, Marcelo Keese Albertini, André Ricardo Backes, and Geórgia das Graças Pena. "Machine learning for hospital readmission prediction in pediatric population." Computer Methods and Programs in Biomedicine 244 (2024): 107980.

[34] Adeniran, Ibrahim Adedeji, Christianah Pelumi Efunniyi, Olajide Soji Osundare, and Angela Omozele Abhulimen. "Data-driven decision-making in healthcare: Improving patient outcomes through predictive modeling." Engineering Science & Technology Journal 5, no. 8 (2024).

[35] Gaso, Mekia Shigute, Remudin Reshid Mekuria, Al Khan, Mohammad Imtiyaz Gulbarga, Iskender Tologonov, and Zuhra Sadriddin. "Utilizing Machine and Deep Learning Techniques for Predicting Re-admission Cases in Diabetes Patients." In Proceedings of the International Conference on Computer Systems and Technologies 2024, pp. 76-81. 2024.

[36] Tsai, Cheng-Han, Dong-Her Shih, Jue-Hong Tu, Ting-Wei Wu, Ming-Guei Tsai, and Ming-Hung Shih. "Analyzing Monthly Blood Test Data to Forecast 30-Day Hospital Readmissions among Maintenance Hemodialysis Patients." Journal of Clinical Medicine 13, no. 8 (2024): 2283.

[37] Hu, Yue, Fanghui Ma, Mengjie Hu, Binbing Shi, Defeng Pan, and Jingjing Ren. "Development and validation of a machine learning model to predict the risk of readmission within one year in HFpEF patients: Short title: Prediction of HFpEF readmission." International Journal of Medical Informatics 194 (2025): 105703.

[38] LI, HEXIN, Rachel Monger, Maryam Pishgar, and Elham Pishgar. "ICU Readmission Prediction for Intracerebral Hemorrhage Patients using MIMIC III and MIMIC IV Databases." medRxiv (2025): 2025-01.

[39] Buddhiraju, Anirudh, Michelle Riyo Shimizu, Tony Lin-Wei Chen, Henry Hojoon Seo, Blake M. Bacevich, Pengwei Xiao, and Young-Min Kwon. "Comparing prediction accuracy for 30-day readmission following primary total knee arthroplasty: the ACS-NSQIP risk calculator versus a novel artificial neural network model." Knee Surgery & Related Research 37, no. 1 (2025): 3.

[40] Purbasari, Intan Yuniar, Athanasius Priharyoto Bayuseno, R. Rizal Isnanto, and Tri Indah Winarni. "Readmission Risk Prediction after Total Hip Arthroplasty Using Machine Learning and Hyperparameters Optimized with Bayesian Optimization." Available at SSRN 5071220.