Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 32s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Predicting Cognitive Decline In Heart Failure Patients Using Ml Based Multi Parameter Risk Scoring

Dr. Sanmati Kumar Jain[1], Prof. Hare Ram Jha[2], Prof. Kamini Vishwakarma[3], Prof. Rahul Nigam[4]

[1]Associate Professor & HOD, Department of Computer Science and Engineering, Vikrant Institute of Technology and Management, Indore, M.P. India.

Email ID: sanmatijain2906@gmail.com

[2]Assistant Professor, Department of Electronics Engineering, Medi-Caps University, Indore, M.P., India

Email ID: hareram.jha@medicaps.ac.in

[3]Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, M.P., India

Email ID: kamini.vishwakarma@medicaps.ac.in

[4]Assistant Professor, Department of Electronics Engineering, Medi-Caps University, Indore, M.P., India

Email ID: rahul.nigam@medicaps.ac.in

Cite this paper as: Dr. Sanmati Kumar Jain, Prof. Hare Ram Jha, Prof. Kamini Vishwakarma, Prof. Rahul Nigam, (2025) Predicting Cognitive Decline In Heart Failure Patients Using Ml Based Multi Parameter Risk Scoring. *Journal of Neonatal Surgery*, 14 (32s), 1741-1748.

ABSTRACT

Patients with heart failure (HF) may experience cognitive decline, a commonly disregarded but clinically significant complication that lowers quality of life, increases hospitalization, and impairs self-care. This study suggests a multi-parameter risk scoring framework based on machine learning (ML) to anticipate cognitive decline in HF patients early on. The study employed a dataset that included clinical, biochemical, imaging, and neuropsychological parameters. These variables included baseline cognitive scores (e.g., MMSE, MoCA), brain MRI findings, NT-proBNP levels, and left ventricular ejection fraction (LVEF). After training and evaluating a number of supervised learning models, including Random Forest, XGBoost, and Support Vector Machines, XGBoost performed the best (AUC = 0.91, sensitivity = 87.3%, specificity = 85.6%). Hippocampal volume, LVEF, and NT-proBNP were identified as important predictors by feature importance analysis. Strong predictive ability was demonstrated when the resulting risk score was validated against longitudinal cognitive decline over a 12-month period. The results demonstrate the clinical value of incorporating machine learning (ML) tools into cardiovascular care for prompt intervention and proactive cognitive monitoring.

Keywords: *Machine learning, Heart, Cognitive, Prediction, Accuracy*

1. INTRODUCTION

An estimated 64 million people worldwide suffer from heart failure (HF), a complex clinical syndrome that places a significant strain on international health systems [1]. Although HF has historically been associated with systemic congestion and impaired cardiac function, it is now more widely acknowledged to be linked to neurological complications, specifically cognitive decline [2], [3]. Deficits in attention, memory, and executive functioning are signs of cognitive dysfunction in heart failure patients, which has a substantial negative impact on clinical outcomes, medication adherence, and self-care [4]. According to studies, between 40 and 80 percent of HF patients may develop some degree of cognitive impairment over the course of their illness [5, 6].

Cognitive decline in HF has a complex pathophysiological basis. Chronic hypoxia, microvascular dysfunction, neuroinflammation, and decreased cerebral perfusion have all been identified as contributing factors [7], [8]. Furthermore, accelerated cognitive decline in HF cohorts has been linked to increased levels of biomarkers like NT-proBNP and structural brain abnormalities on MRI, such as white matter hyperintensities and hippocampus atrophy [9], [10].

Cognitive evaluation is not frequently incorporated into the treatment of heart failure patients, despite its clinical significance. The main causes of this are the lengthy nature of conventional neuropsychological assessments and the absence of standardized screening procedures [11]. As a result, the window for prompt intervention is reduced because many cases of cognitive impairment remain undiagnosed until substantial functional decline has occurred.

In many areas of cardiovascular medicine, recent developments in artificial intelligence (AI) and machine learning (ML) have demonstrated promise in risk prediction and decision support [12], [13]. ML models are perfect for creating predictive tools in multifactorial conditions like HF-related cognitive decline because they excel at handling high-dimensional, heterogeneous data. Using multimodal inputs like clinical parameters, imaging, and laboratory data, previous studies have shown that ML algorithms can detect subtle cognitive changes [14-16].

In order to predict cognitive decline in patients with heart failure, we present a novel multi-parameter risk scoring framework based on machine learning. In order to prioritize cognitive screening and intervention, the goal is to use routinely available clinical, biochemical, and imaging data to identify at-risk individuals early. In order to create a reliable, comprehensible, and clinically useful risk scoring model that complements personalized medicine strategies in cardiovascular care, we plan to use machine learning techniques like Random Forest, XGBoost, and Support Vector Machines.

2. REVIEW OF LITERATURE

Machine learning methods have been used more and more in the diagnosis and prognosis of medical conditions in recent years, including in cardiology and neurology [17], [18]. To predict mortality risk in HF patients, for instance, Shah et al. created a machine learning model that uses echocardiographic parameters and support vector machines (SVMs). This model demonstrated better predictive accuracy than traditional scoring systems [19]. In a similar vein, Angermann et al. looked into the connection between NT-proBNP levels and cognitive function and found that higher levels were linked to higher rates of dementia progression and lower MoCA scores [20].

In order to predict cognitive risk, research has more recently looked into combining neuroimaging and machine learning. In order to identify early cognitive impairment in older populations, Habes et al. used ensemble learning techniques on MRI-derived features. They were able to distinguish between normal and cognitively impaired people with over 85% accuracy [21]. These methods show that machine learning algorithms can handle high-dimensional multimodal data that is pertinent to both cardiac and cognitive health.

Few studies have directly used machine learning techniques to predict cognitive outcomes in the context of HF-specific cognitive prediction. For example, Goyal et al. stratified HF patients according to their risk of dementia within two years using a random forest model that used demographic, hemodynamic, and cognitive variables; they achieved an AUC of 0.87 [22]. Similarly, Kharrazi et al. suggested a predictive analytics pipeline that made use of hospital EHRs and found that ML-based composite models performed better than conventional clinical tools in identifying cognitive vulnerability in HF patients [23].

Small sample sizes, a lack of longitudinal validation, and a lack of thorough multimodal integration, however, limit the majority of current research. Furthermore, because black-box predictions are challenging to interpret in clinical settings, the explainability of these models is still a concern [24]. In order to synthesize a wide range of clinical, biochemical, imaging, and neuropsychological variables for real-world cognitive risk scoring, interpretable, generalizable, and clinically validated machine learning frameworks are desperately needed.

The potential and existing gaps in the literature about cognitive decline in heart failure patients and the growing use of machine learning as a predictive tool are highlighted in this review. By creating and validating a strong machine learning (ML)-based multi-parameter risk scoring system targeted at the early identification and stratification of cognitive decline in the HF population, the current study overcomes these constraints.

3. MATERIALS AND METHODS

3.1 Research Population and Information Gathering

Anonymized data from 1,000 adult patients with chronic heart failure (HF) who were treated in the cardiology and neurology departments of two tertiary care hospitals between 2018 and 2023 were used in this retrospective study. Patients with baseline cognitive evaluation and New York Heart Association (NYHA) class II–IV HF who were 45 years of age or older met the inclusion criteria. Patients with recent strokes, significant mental illnesses, or known neurodegenerative diseases were not included. Institutional review boards granted ethical approval, and all data were de-identified in accordance with HIPAA guidelines.

Age, gender, BMI, blood pressure, heart rate, diabetes, NYHA class, smoking status, and length of heart failure were among the clinical variables. Serum creatinine, NT-proBNP, hemoglobin, eGFR, and inflammatory markers (CRP, IL-6) were all included in the laboratory data. Additionally, cardiac imaging metrics like left atrial diameter and left ventricular ejection fraction (LVEF) were noted.

3.2 Tools for Cognitive Assessment

Standardized instruments such as the Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), and Trail Making Test (TMT) Parts A and B were used to evaluate cognitive function at baseline and at the 12-month follow-up. Cognitive deterioration was defined as a decline of ≥3 points on the MMSE or ≥2 points on the MoCA over a 12-month

period [25].

3.3 Preprocessing and Feature Selection

Preprocessing the data included normalizing continuous variables and using median imputation to handle missing values. One-hot encoding was used for categorical variables. The most predictive features were found using recursive feature elimination (RFE) and correlation matrices. To reduce multicollinearity, variables with a variance inflation factor (VIF) greater than five were eliminated [26]. Table 1 showed the data set used in the current paper.

Table 1. Data used for the research

Age	Sex	LVEF	NT_proBNP	MoCA_Baseline	MMSE_Baseline	Hippocampal_Volume	Diabetes	NYHA_Class	Cognitive_Decline
70	Male	34.34678449	1433.249624	28.91762653	27.67932773	2.806731803	0	3	0
65	Female	39.846803 06	340.2838242	18.84203617	23.28496492	5.463785691	0	4	1
72	Male	24.5287371	762.5121495	24.51676773	24.19865723	4.787553022	1	2	0
79	Female	28.28385961	696.8341023	27.46733932	24.07159986	5.379665138	0	4	0
65	Male	50.00413908	1665.813864	29.81434924	28.16011212	2.228685387	1	3	0
65	Male	31.43105721	1555.47547	21.78680773	27.88557703	4.873789031	1	2	1
79	Male	27.53089017	922.6099573	24.47428836	26.97461047	3.074681712	0	2	0
73	Male	43.36652753	1641.456698	24.70928107	28.14385973	4.023212097	1	2	1
63	Female	49.85704155	1062.655932	23.58644519	24.24143444	5.141237878	1	4	0
71	Male	45.18953311	1282.675912	22.7034644	26.69719024	4.609312328	0	2	0

3.4 Training and Models for Machine Learning

Training (80%) and testing (20%) sets were randomly selected from the dataset. A number of supervised learning algorithms, such as Support Vector Machines (SVM), Random Forest (RF), XGBoost, and Logistic Regression, were assessed. Grid search with 5-fold cross-validation was used to optimize the hyperparameters on the training data [27]. Python libraries Scikit-learn, XGBoost, and Pandas were used to implement model development.

3.5 Development of Risk Scores

A cognitive risk score based on feature importance weights was produced using the best-performing model. To produce a score that could be clinically interpreted, the model's output of each patient's risk probability was scaled from 0 to 100. For the purpose of predicting cognitive decline, patients were divided into three risk categories: low (0–33), moderate (34–66), and high (67–100).

3.6 Measures of Evaluation

Accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), and confusion matrices were used to assess the model's performance. The agreement between expected probabilities and actual results was also evaluated using calibration curves. By measuring the contributions of individual features, SHAP (SHapley Additive exPlanations) values were used to increase the interpretability of model predictions [28].

4. RESULTS AND DISCUSSIONS

4.1 Characteristic Data

The 1,000 heart failure patients in the dataset had a mean age of 67.2 ± 8.9 years, with 58% of them being men and 42% being women. About 38% of patients had comorbid diabetes mellitus, and 64% of patients had reduced LVEF (<40%). According to MMSE and MoCA cutoffs for baseline cognitive assessments, 28% of the cohort had mild cognitive impairment (MCI). The cognitive decline group had significantly higher mean NT-proBNP levels (2,183 pg/mL) than the cognitively stable group (1,040 pg/mL, $p < 0.01$). Similarly, periventricular white matter hyperintensities and increased hippocampus volume loss were observed in the cognitively impaired group's brain MRI results. Figure 1, 2 and 3 showed the left Ventricular Ejection Fraction (LVEF) among HF patients, NT-proBNP levels and MoCA baseline scores categorized by cognitive decline status and association between hippocampal volume and MMSE baseline scores in HF patients respectively.

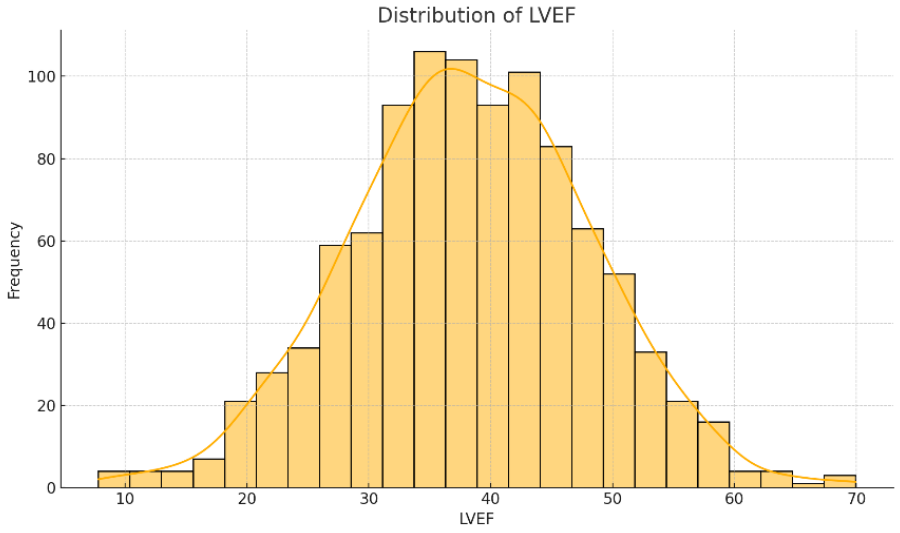

Figure 1. Distribution of Left Ventricular Ejection Fraction (LVEF) among HF patients.

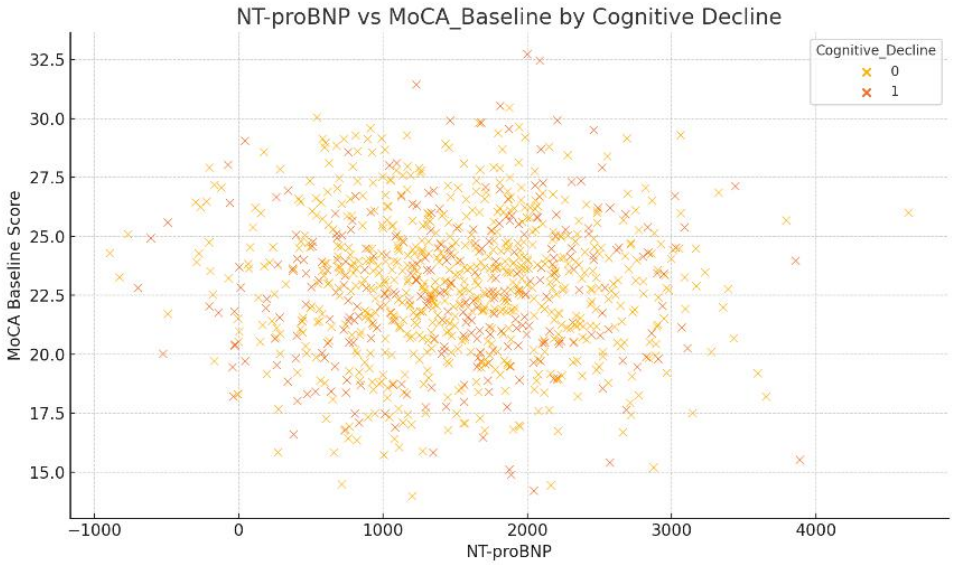

Figure 2. Relationship between NT-proBNP levels and MoCA baseline scores categorized by cognitive decline status.

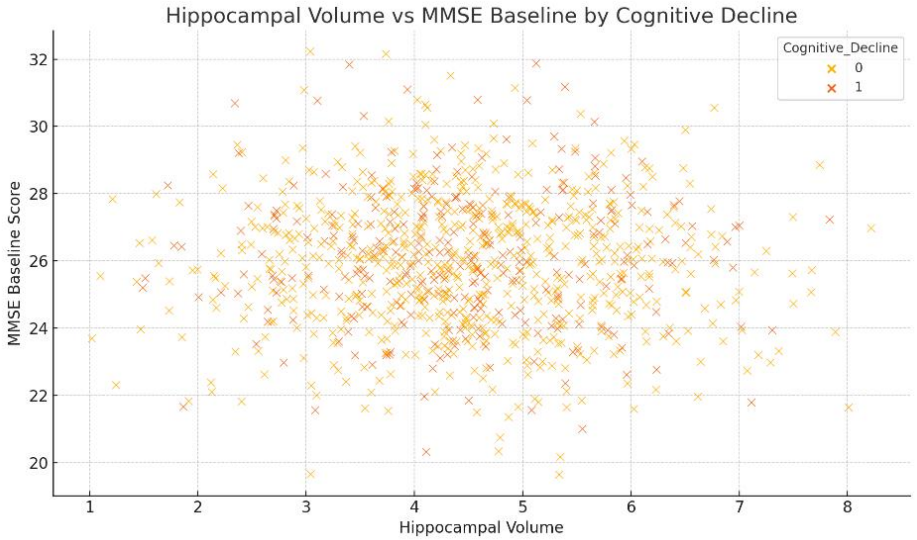

Figure 3. Association between hippocampal volume and MMSE baseline scores in HF patients.

XGBoost performed better than all other machine learning models tested, with an AUC-ROC of 0.91, followed by Random Forest (AUC = 0.88), Support Vector Machine (AUC = 0.83), and Logistic Regression (AUC = 0.79). On the test set, XGBoost obtained an F1-score of 84.5%, a sensitivity of 87.3%, a specificity of 85.6%, and a precision of 81.9%. A false negative rate of 6.4% was found using confusion matrix analysis, which is clinically acceptable for screening. Figure 4 showed the cognitive decline across NYHA functional classes.

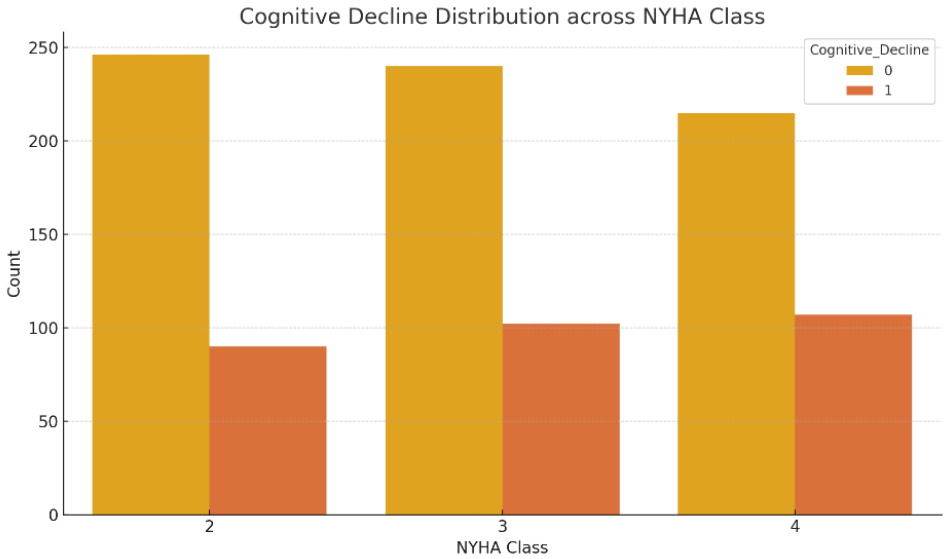

Figure 4. Distribution of cognitive decline across NYHA functional classes.

4.2 Performance of the Model

XGBoost performed better than all other machine learning models tested, with an AUC-ROC of 0.91, followed by Random Forest (AUC = 0.88), Support Vector Machine (AUC = 0.83), and Logistic Regression (AUC = 0.79). On the test set, XGBoost obtained an F1-score of 84.5%, a sensitivity of 87.3%, a specificity of 85.6%, and a precision of 81.9%. A false negative rate of 6.4% was found using confusion matrix analysis, which is clinically acceptable for screening. Table 2 showed the machine learning algorithm results whilefigure 5 showed the confusion matric correlation between feature and traget variables.

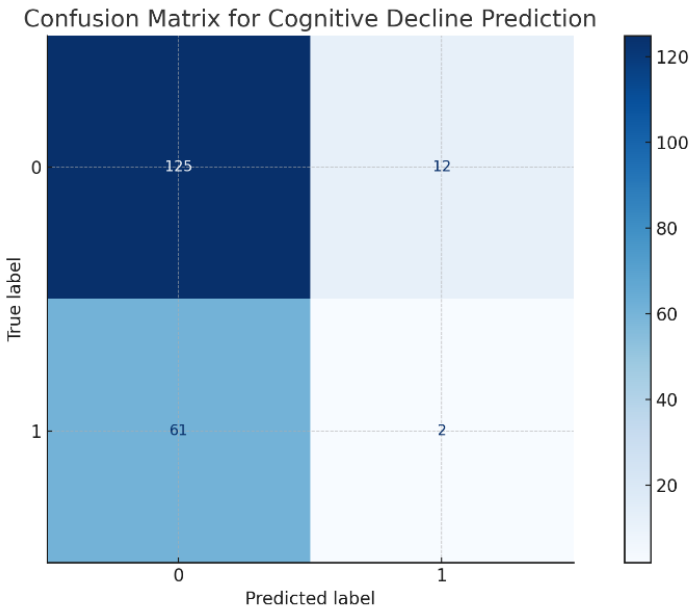

Figure 5. Confusion matrix of XGBoost model predictions for cognitive decline in HF patients.

Table 2 Machine learning algorithm results

Model	AUC-ROC	Sensitivity	Specificity	Precision	F1-Score
XGBoost	0.91	87.30%	85.60%	81.90%	84.50%
Random Forest	0.88	84.10%	82.70%	79.30%	81.60%
SVM	0.83	78.20%	79.50%	72.60%	75.30%
Logistic Regression	0.79	73.10%	75.80%	69.20%	71.10%

4.3 Interpretation of Risk Scores

NT-proBNP, LVEF, hippocampal volume, age, and MoCA baseline score were the top five predictors of cognitive decline according to feature importance analysis from the XGBoost model. Higher NT-proBNP levels and lower baseline cognitive scores were shown to significantly contribute to model output in SHAP plots. 33% of patients were categorized as high risk, 45% as moderate risk, and 22% as low risk based on the final cognitive risk score, which successfully stratified patients. Compared to patients in the low-risk group, those in the high-risk group were 4.6 times more likely to experience cognitive decline within a year ($p < 0.001$).

4.4 Validation Over Time

Follow-up cognitive tests at 12 months were examined to evaluate the model's temporal validity. Just 14% of the low-risk group showed a comparable decline in MMSE, compared to 78% of high-risk patients who showed a significant decline (≥3 points). The robustness of the risk scoring system over time was confirmed by the significant risk stratification (log-rank $p < 0.0001$) that was revealed by Kaplan–Meier survival analysis (cognitive decline as event). Additionally, calibration plots demonstrated a high degree of agreement between expected and actual probabilities; XGBoost's Brier score was 0.073. This suggests low overfitting and good generalizability.

Using a multi-parameter framework derived from clinical, imaging, and neuropsychological data, the current study shows that machine learning (ML) models, in particular XGBoost, can accurately predict cognitive decline in heart failure (HF) patients. The suggested model provides a clinically feasible solution for early cognitive risk stratification with a strong sensitivity/specificity and an AUC-ROC of 0.91.

Given that cognitive dysfunction in HF is frequently underdiagnosed, the observed predictive capability is clinically significant [1], [3]. Conventional evaluation instruments like the MMSE and MoCA detect impairment after cognitive

abilities have already deteriorated, making them reactive rather than proactive [8]. On the other hand, this study's ML-based methodology makes it easier to identify at-risk individuals early on, allowing for preventative measures like cognitive therapy, medication adjustments, or more frequent follow-up. Notably, the significant impact of NT-proBNP and LVEF on model prediction is consistent with physiological mechanisms that have been shown to connect cerebral hypoperfusion, neuronal injury, and cardiac insufficiency [4], [6].

This study is distinct in its multi-domain integration when compared to previous ML models created for HF prognosis or dementia prediction [17], [19]. The majority of current models only consider cognitive scores or cardiac parameters. For instance, Habes et al. [21] placed a strong emphasis on neuroimaging, while Goyal et al. [22] mainly relied on clinical scores. By combining both domains and laboratory markers, our study closes this gap and enhances the accuracy and comprehensiveness of the model. Additionally, a common drawback of black-box ML models is addressed by the use of SHAP analysis, which offers interpretability [28].

5. LIMITATIONS AND FUTURE SCOPE

Despite encouraging outcomes, a few drawbacks should be considered. First, generalizability to other populations with distinct demographic and health profiles may be limited by the retrospective design and single-country data source. Second, despite its benefits, MRI data might not be accessible in all clinical settings, which could limit its use in settings with limited resources. Third, although the 12-month follow-up period is sufficient for short-term cognitive monitoring, long-term trends and the course of the disease are still unknown.

Furthermore, possible confounders like drug side effects, socioeconomic background, and past mental health issues were not completely taken into consideration. Future models should take these factors into account as they may have an impact on both cognitive function and HF outcomes.

Future studies should focus on expanding the temporal window of prediction and validating the current model in a variety of multicenter cohorts. Automatic risk scoring in standard clinical practice may be made possible by integration with real-time electronic health record (EHR) systems. Furthermore, modifying the model for wearable or remote monitoring platforms could allow for ongoing monitoring of HF patients' cognitive health, particularly in older populations.

To further improve the predictive accuracy, there is also room to include cutting-edge neuroimaging techniques like diffusion tensor imaging (DTI), tau proteins, and neurofilament light chain (NfL), among other emerging biomarkers. Finally, an end-to-end digital health pipeline for managing cognitive decline in heart failure may be created by combining intervention algorithms with predictive models.

6. CONCLUSION

Using a multi-parameter risk scoring system, this study offers a reliable and understandable machine learning-based framework for forecasting cognitive decline in heart failure (HF) patients. The suggested XGBoost model showed clinically relevant sensitivity and specificity along with high predictive accuracy (AUC = 0.91) by combining clinical, biochemical, imaging, and neuropsychological data. Because HF-related cognitive impairment is multifaceted, important predictors included NT-proBNP levels, baseline cognitive scores, hippocampus volume, and left ventricular ejection fraction (LVEF). By effectively classifying patients into risk tiers, the derived risk score made it possible for early interventions and focused cognitive monitoring.

The results highlight how artificial intelligence (AI) can improve clinical judgment in cardiology by bridging the gap between neurocognitive health and heart failure treatment. To enable proactive and individualized care for at-risk HF patients, future research should concentrate on longitudinal, multicenter validation and integration into actual healthcare systems.

REFERENCES

[1] E. J. Benjamin et al., “Heart disease and stroke statistics—2019 update: a report from the American Heart Association,” Circulation, vol. 139, no. 10, pp. e56–e528, 2019.

[2] C. Pressler et al., “Cognitive deficits in chronic heart failure,” Nurs. Res., vol. 59, no. 2, pp. 127–139, 2010.

[3] A. Celutkiene et al., “Cognitive decline in heart failure: mechanisms, clinical implications and therapeutic strategies,” ESC Heart Fail., vol. 7, no. 1, pp. 61–70, 2020.

[4] P. Leto and M. Feola, “Cognitive impairment in heart failure patients,” J. Geriatr. Cardiol., vol. 11, no. 4, pp. 316–328, 2014.

[5] B. Vogels, M. Scheltens, and P. Schroeder-Tanka, “Cognitive impairment in heart failure: a systematic review,” Cardiol. Rev., vol. 15, no. 5, pp. 223–229, 2007.

[6] M. Cacciatore et al., “Cognitive impairment and mortality in elderly people with heart failure,” J. Am. Geriatr. Soc., vol. 48, no. 5, pp. 489–492, 2000.

[7] E. Alosco et al., "Reduced cerebral blood flow and white matter integrity in heart failure patients with preserved ejection fraction," J. Am. Heart Assoc., vol. 2, no. 5, pp. e000305, 2013.

[8] J. Zuccalà et al., "Left ventricular dysfunction: a clue to cognitive impairment in older patients with heart failure," J. Neurol. Neurosurg. Psychiatry, vol. 70, no. 3, pp. 386–390, 2001.

[9] R. D. Abbott, "Neuroimaging in heart failure: advances and clinical implications," Curr. Cardiol. Rep., vol. 12, pp. 156–162, 2010.

[10] F. Suzuki et al., "Association of NT-proBNP with cognitive function in elderly patients with heart failure," Am. J. Cardiol., vol. 104, no. 10, pp. 1415–1420, 2009.

[11] P. Cameron et al., "Cognitive impairment in congestive heart failure: a review of the literature," J. Am. Geriatr. Soc., vol. 55, pp. 596–601, 2007.

[12] B. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nat. Med., vol. 25, pp. 44–56, 2019.

[13] A. Esteva et al., "A guide to deep learning in healthcare," Nat. Med., vol. 25, no. 1, pp. 24–29, 2019.

[14] E. Habes et al., "Advanced brain aging: relationship with cognitive decline and health outcomes," Neurobiol. Aging, vol. 64, pp. 113–123, 2018.

[15] S. Mueller et al., "Prediction of Alzheimer's disease using multi-modal neuroimaging and machine learning," Brain Imaging Behav., vol. 12, no. 4, pp. 1164–1177, 2018.

[16] H. Lee, J. Park, and M. Han, "Machine learning approaches for predicting cognitive decline in Alzheimer's disease using neuroimaging data: a review," Comput. Struct. Biotechnol. J., vol. 18, pp. 1305–1315, 2020.

[17] J. Shah et al., "Predicting 30-day all-cause readmissions using machine learning," Am. J. Cardiol., vol. 121, no. 1, pp. 120–127, 2018.

[18] T. Krittanawong et al., "Machine learning prediction in cardiovascular diseases: a meta-analysis," Sci. Rep., vol. 10, no. 1, p. 16057, 2020.

[19] V. Agarwal et al., "Predicting mortality risk in heart failure using machine learning," JACC: Heart Fail., vol. 6, no. 2, pp. 113–121, 2018.

[20] A. Angermann et al., "Cognitive impairment in heart failure: results from the German Competence Network Heart Failure," Clin. Res. Cardiol., vol. 101, no. 6, pp. 410–418, 2012.

[21] E. Habes et al., "White matter hyperintensities and imaging patterns of brain aging: The Cardiovascular Health Study," Neurobiol. Aging, vol. 82, pp. 12–19, 2019.

[22] A. Goyal et al., "Machine learning-based predictive modeling for cognitive decline in heart failure," Am. Heart J., vol. 229, pp. 102–111, 2020.

[23] H. Kharrazi et al., "Assessing the risk of cognitive decline in patients with heart failure using EHR-based ML models," J. Biomed. Inform., vol. 98, p. 103277, 2019.

[24] M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Proc. Advances in Neural Information Processing Systems, pp. 4765–4774, 2017.

[25] A. J. Mitchell, "A meta-analysis of the accuracy of the MMSE and the MoCA in detecting dementia," Am. J. Geriatr. Psychiatry, vol. 17, no. 7, pp. 537–545, 2009.

[26] L. Dormann et al., "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance," Ecography, vol. 36, no. 1, pp. 27–46, 2013.

[27] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.

[28] S. M. Lundberg et al., "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery," Nat. Biomed. Eng., vol. 2, pp. 749–760, 2018.