Artificial Intelligence for Modeling Complex Treatment Decisions in Aortic Valve Intervention

1st Jie Jun Wong
Cardiology
National Heart Centre Singapore
Singapore, Singapore
jiejun.wong@mohh.com.sg

2nd Glades Tan
Cardiology
National Heart Centre Singapore
Singapore, Singapore
gladeshmm@gmail.com

3rd Xinliu Zhong
Biomedical Engineering
National University of Singapore
Singapore, Singapore
xinliuzhong@u.nus.edu

4th Kay Woon Ho
Cardiology
National Heart Centre Singapore
Singapore, Singapore
ho.kay.woon@singhealth.com.sg

5th Vincent Wei Jun Sim
Cardiology
National Heart Centre Singapore
Singapore, Singapore
weijunsim@ymail.com

6th Si Yong Yeo
Digital Health
Lee Kong Chian School of Medicine
Singapore, Singapore
siyong.yeo@ntu.edu.sg

7th Angela S. Koh
Cardiology
National Heart Centre Singapore
Singapore, Singapore
angela.koh.s.m@singhealth.com.sg

2024 IEEE Conference on Artificial Intelligence (CAI) | 979-8-3503-5409-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/CAI59869.2024.00049

***Abstract*— When making treatment decisions for invasive cardiovascular procedures in older persons, physicians often face a myriad of complex scenarios, such as frailty, cognitive impairment, and multimorbidity. Accounting for these characteristics in real-world practice is challenging in aortic valve replacement for aortic stenosis (AS) as they impact individualized decisions in achieving meaningful postprocedural outcomes without excessive risk. Based on these characteristics, 864 unique scenarios were created that formed the original dataset, which was further split into 70% training and 30% testing datasets. More controversial clinical scenarios were further tuned based on responses from ten cardiologists and processed using multilayered neural network sequential features analysis and deep learning methods. Contrary to guidelines, symptoms and left ventricular function ranked low in physician importance. In contrast, aging-related functional features, including cognition, ambulation, and frailty scores, ranked high with good overall model accuracy (Shapley 0.811, TabNet 0.938). Feature optimization using the top three features showed good model accuracy (Shapley 0.811, TabNet 0.881). Here, AS illustrates a use case scenario in artificial intelligence (AI) that could be applied to complex clinical decision-making and has excellent potential for handling diverse clinical problems and augmenting physician treatment decision-making.**

Keywords*— *artificial intelligence; aortic stenosis; clinical decision-making; deep learning; neural networks

I. Introduction

Degenerative aortic valve stenosis (AS) is a common aging-related progressive heart valve disease stemming from leaflet calcification resulting in abnormal valve opening and impending cardiac outflow [1]. AS is the most common form of valvular heart disease worldwide, and the global prevalence is anticipated to increase further with global aging [1,2]. The mortality of untreated severe AS is alarmingly high at 50-70% in five years, and treatment options are confined to valve replacement (AVR) either surgically (SAVR) or percutaneously transcatheter (TAVR) for higher surgical risk patients [2,3]. However, SAVR is associated with significant morbidity; although TAVR is less invasive, the procedural risks are still substantial [1]. Notably, a considerable proportion of patients do not improve functionally after AVR (estimated up to 45% post-TAVR) due to pre-existing frailty and multimorbidity [4]. Therefore, selecting the right patient to achieve meaningful procedural outcomes without undue risk is a paramount task for physicians.

Current major cardiology guidelines recommend AVR in severe AS if patients have (a) symptoms, such as angina, or (b) impaired cardiac contractility (LVEF <50%) [1]. While less severe forms of AS (i.e., moderate) and asymptomatic patients are at higher mortality risks, there is insufficient evidence to offer AVR routinely due to procedural risks [1]. Other vital clinical factors must be considered in older persons, such as cognition, frailty, and other major organ diseases. These factors are essential in individualized patient selection but are often too complex to be represented in clinical trials or excluded entirely. Treatment decision-making relies heavily on physician gestalt and experience, accounting for innumerable variables not adequately captured by clinical guidance. This short paper examines the effectiveness of artificial intelligence (AI) via multilayered neural networks and deep learning (DL) in complex medical decision-making, incorporating multiple soft clinical variables that are challenging to capture using traditional statistical techniques.

II. Methods

A. Dataset Development

Prior studies identified eight variables associated with poorer outcomes in AS (age, symptoms, AS severity, LVEF, moderate-advanced kidney disease, cognition, frailty score, ambulation, and other surgery indications). Each variable had 2-3 permutations, producing 864 unique hypothetical clinical scenarios that formed the original dataset; 70% (n=604) were allocated as training and 30% (n=260) as testing datasets. Age was defined as $\geqslant 80$ years, 70-80 years, or <70 years; symptoms were being symptomatic or asymptomatic; AS severity was moderate or severe; LVEF was reduced (<50%) or preserved ($\geqslant 50\%$); moderate-advanced kidney disease (CKD) was eGFR <45ml/min/1.73m^2 not on dialysis; cognition was moderate dementia, mild dementia, or no impairment; frailty score was frail or non-frail; ambulation was able to walk, wheelchair-bound, or bedbound. Three decision outcome options were possible: SAVR, TAVR, or conservative management. The original dataset was completed by one cardiologist, producing counts of 700 conservative, 108 TAVR, and 56 SAVR decision outcomes. From the original dataset, controversial scenarios were identified and crafted into a survey focusing on (a) guideline indications with relative contraindications and (b) non-guideline indications with clinical grounds for AVR. The survey was disseminated in November 2023 to ten independent senior board-certified cardiologists. Responses were analyzed individually.

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00049

B. Model Development and Shapley analysis

The datasets were first processed using a combination of supervised greedy layer-wise multilayer neural networks and Shapley methods [5-7]. Transfer learning was utilized to account for the large dataset, various scenarios, and survey responses. Using the Sequential model Application Programming Interface (API), a Multilayer Perceptron (MLP) neural network with eight inputs, three outputs, and two hidden layers with ten nodes each (Figure 1).

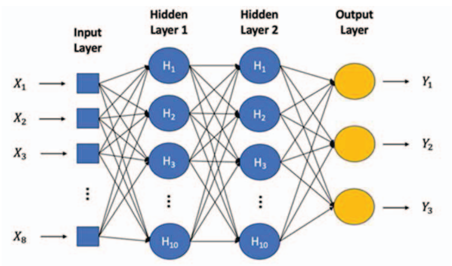

Figure 1. Multi-layer perceptron (MLP) neural network construct.

Shapley Additive Explanations (SHAP) analysis was applied using KernelSHAP to obtain feature attributions by measuring the average marginal contribution of each feature towards the outcome in each scenario and assigning a local importance value to each feature [8]. Sensitivity analysis was then performed using TabNet Classifier analysis [9].

III. Results

A. MLP model performance

The Sequential API MLP neural network model attained an overall test accuracy of 0.811. KernelSHAP analysis without feature optimization found that AS severity had the greatest influence on decision-making, with a mean SHAP value of approximately 0.122 (Figure 2). Among the top-ranked features, frailty scores, cognition, and CKD attained

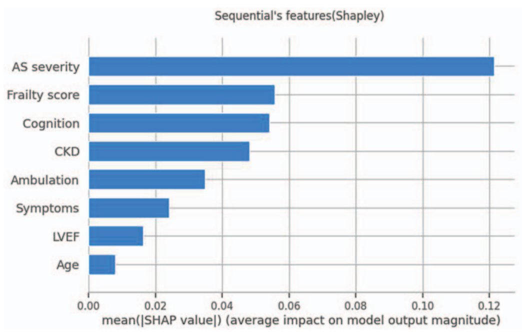

Figure 2. MLP feature importance using KernelSHAP analysis.

higher mean SHAP values of about 0.055, 0.053, and 0.048, respectively. Ambulation status was ranked middle, while the presence of symptoms, LVEF, and age was ranked the lowest. Repeat analysis with feature optimization using the top three features (AS severity, frailty score, and cognition) produced a similar test accuracy of 0.811.

B. Sensitivity analysis via TabNet Classifier

TabNet Classifier sensitivity analysis identified cognition as the most critical feature in determining valve replacement (feature importance 0.363), while ambulation and AS severity ranked 2nd and 3rd (0.249 and 0.152, respectively). Age and CKD ranked in the middle, whereas symptoms, LVEF, and frailty score ranked the lowest.

IV. Discussion

This paper describes an efficient pipeline method of utilizing machine learning and deep learning modeling techniques to unravel complex treatment decisions by cardiologists when deciding AVR. Cognition, ambulation, and frailty scores ranked highly important to physicians deciding on aortic valve intervention. Moderate-advanced CKD, representing multimorbidity, was also crucial, likely due to implications on periprocedural dialysis and bleeding risks [10]. The presence of symptoms and LVEF were ranked low importance using Shapley and TabNet models, which were contrary to guideline recommendations. In another retrospective longitudinal cohort study using LASSO and decision tree techniques on patients undergoing AVR, frailty score was the most predictive factor for TAVR or SAVR selection; other high-value predictors were age, surgical risk score, and CKD [11]. Frailty-related features require greater recognition in treatment decisions of invasive cardiovascular procedures and will inform future implementation research and treatment guidance.

References

[1] M. Rana, "Aortic Valve Stenosis: Diagnostic Approaches and Recommendations of the 2021 ESC/EACTS Guidelines for the Management of Valvular Heart Disease -A Review of the Literature," (in eng), *Cardiol Cardiovasc Med,* vol. 6, no. 3, pp. 315-324, Jun 2022, doi: 10.26502/fccm.92920267.

[2] A. Hartley et al., "Trends in Mortality From Aortic Stenosis in Europe: 2000-2017," (in eng), Front Cardiovasc Med, vol. 8, p. 748137, 2021, doi: 10.3389/fcvm.2021.748137.

[3] G. S. Gheorghe, A. S. Hodorogea, A. C. D. Gheorghe, I. T. Nanea, and A. Ciobanu, "Medical management of symptomatic severe aortic stenosis in patients non-eligible for transcatheter aortic valve implantation," (in eng), J Geriatr Cardiol, vol. 17, no. 11, pp. 704-709, Nov 28 2020, doi: 10.11909/j.issn.1671-5411.2020.11.002.

[4] M. Amaki et al., "Cost-effective analysis of Transcatheter aortic valve replacement in patients with severe symptomatic aortic stenosis: A prospective multicenter study," (in eng), J Cardiol, Aug 3 2023, doi: 10.1016/j.jjcc.2023.07.018.

[5] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," Advances in neural information processing systems, vol. 19, 2006.

[6] M. Sundararajan and A. Najmi, "The many Shapley values for model explanation," in *International conference on machine learning*, 2020: PMLR, pp. 9269-9278.

[7] S. Haykin, Neural networks: a comprehensive foundation. Prentice Hall PTR, 1998.

[8] I. Covert and S.-I. Lee, "Improving kernelshap: Practical shapley value estimation using linear regression," in International Conference on Artificial Intelligence and Statistics, 2021: PMLR, pp. 3457-3465.

[9] S. Ö. Arik and T. Pfister, "Tabnet: Attentive interpretable tabular learning," in Proceedings of the AAAI conference on artificial intelligence, 2021, vol. 35, no. 8, pp. 6679-6687.

[10] J. A. Naser et al., "Impact on Kidney Function and Medium-Term Outcomes of Transcatheter Aortic Valve Replacement in Patients With Chronic Kidney Disease," (in eng), Am J Cardiol, vol. 210, pp. 163-171, Jan 1 2024, doi: 10.1016/j.amjcard.2023.10.014

[11] R. Chokesuwattanaskul et al., "Machine Learning-Based Predictive Model of Aortic Valve Replacement Modality Selection in Severe Aortic Stenosis Patients," (in eng), Med Sci (Basel), vol. 12, no. 1, Dec 29 2023, doi: 10.3390/medsci120100