Original Paper

Enhancing Frailty Assessments for Transcatheter Aortic Valve Replacement Patients Using Structured and Unstructured Data: Real-World Evidence Study

Mamoun T Mardini[1], MS, PhD; Chen Bai[1], MS; Anthony A Bavry[2], MPH, MD; Ahmed Zaghloul[2], MD; R David Anderson[3], MD; Catherine E Crenshaw Price[4,5], PhD; Mohammad A Z Al-Ani[3], MD

[1]Health Outcomes and Biomedical Informatics, College of Medicine, University of Florida, Gainesville, FL, United States

[2]Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, United States

[3]Division of Cardiovascular Medicine, College of Medicine, University of Florida, Gainesville, FL, United States

[4]Department of Clinical and Health Psychology, College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

[5]Perioperative Cognitive Anesthesia Network for Alzheimer's Disease and Related Dementias, University of Florida, Gainesville, FL, United States

Corresponding Author:

Mamoun T Mardini, MS, PhD
Health Outcomes and Biomedical Informatics
College of Medicine, University of Florida
1889 Museum Rd, 7th Floor, Suite 7000
Gainesville, FL, 32608
United States
Phone: 1 352 627 9467
Email: mamounmardini@gmail.com

Abstract

Background: Transcatheter aortic valve replacement (TAVR) is a commonly used treatment for severe aortic stenosis. As degenerative aortic stenosis is primarily a disease afflicting older adults, a frailty assessment is essential to patient selection and optimal periprocedural outcomes.

Objective: This study aimed to enhance frailty assessments of TAVR candidates by integrating real-world structured and unstructured data.

Methods: This study analyzed data from 14,000 patients between January 2018 and December 2019 to assess frailty in TAVR patients at the University of Florida. Frailty was identified using the Fried criteria, which includes weight loss, exhaustion, walking speed, grip strength, and physical activity. Latent Dirichlet allocation for topic modeling and Extreme Gradient Boosting for frailty prediction were applied to unstructured clinical notes and structured electronic health record (EHR) data. We also used least absolute shrinkage and selection operator regression for feature selection. Model performance was rigorously evaluated using nested cross-validation, ensuring the generalizability of the findings.

Results: Model performance was significantly improved by combining unstructured clinical notes with structured EHR data, achieving an area under the receiver operating characteristic curve of 0.82 (SD 0.07), which surpassed the EHR-only model's area under the receiver operating characteristic curve of 0.64 (SD 0.08). The Shapley Additive Explanations analysis found that congestive heart failure management, back problems, and atrial fibrillation were the top frailty predictors. Additionally, the latent Dirichlet allocation topic modeling identified 7 key topics, highlighting the role of specific medical treatments in predicting frailty.

Conclusions: Integrating unstructured clinical notes and structured EHR data led to a notable enhancement in predicting frailty. This method shows great potential for standardizing frailty assessments using real-world data and improving patient selection for TAVR.

Introduction

Degenerative severe aortic stenosis is estimated to affect 3.4% of the older population [1]. It results in reduced cardiac output, which reduces activity tolerance and often restricts the ability to perform activities of daily living [2,3]. This stage is associated with high short-term mortality [4]. The most common treatment that helps many patients resume an active lifestyle is transcatheter aortic valve replacement (TAVR), a less invasive procedure than open-heart surgery that replaces the diseased or damaged aortic valve [5]. Unfortunately, about 25% to 35% of patients either die or achieve no functional, morbidity, or mortality benefit from TAVR [6,7]. To optimize patient selection and postsurgical outcomes, accurate tools to acquire personalized, pertinent health data for TAVR candidates can provide valuable insights for patients and their clinicians to make informed decisions in presurgical clinics [6,8].

The American College of Cardiology's guidelines [9], along with various clinical studies [8,10,11], have identified frailty as a risk factor for post-TAVR mortality and morbidity. In a systematic review, Sepehri et al [12] demonstrated the relationship between frailty and adverse postsurgical outcomes in the cardiac population. In addition, research in health economics has shown that frail patients undergoing cardiac procedures incur higher hospitalization costs than nonfrail patients [13]. Given the clinical and economic implications of the TAVR procedure, there is a need to standardize the frailty assessment. This can aid in formulating strategies to enhance post-TAVR outcomes in frail patients. Such strategies may include partnering with geriatricians, opting for monitored anesthesia care, and employing 2 surgeons for complicated procedures, among several others [14].

Existing risk scores, such as the Society of Thoracic Surgeons score and the European System for Cardiac Operative Risk Evaluation, are commonly used for patients referred for cardiac surgery. These scores are useful, but they are stratified into lower-risk patient groups referred for cardiac surgery and do not meet the specific needs of TAVR patients [15-17]. The current methods of assessing frailty in patients undergoing TAVR have significant limitations. A wide variety of tests and instruments are used across studies, many of which have not been robustly evaluated for this specific patient population [15,18]. For example, some studies use subjective methods such as the eyeball test, which relies on visual assessments to gauge frailty but lacks standardization and objective criteria [15]. Other instruments like the Katz Index of Independence in Activities of Daily Living [19] and the Kansas City Cardiomyopathy Questionnaire [20] are often employed in TAVR clinics to assess functional dependency in older individuals, focusing on daily living activities like feeding, bathing, and dressing. While these assessments are valuable in evaluating different aspects of frailty, they vary significantly in their focus and methodology,

leading to inconsistencies in frailty classification among TAVR patients. Additionally, they are not specific to the TAVR population. The Fried frailty phenotype [21] and the Rockwood Frailty Index [22] are well-established and validated frailty assessments. However, these assessments require a detailed physical examination, patient interviews [21], and the completion of surveys [22], which may be time consuming and resource intensive. For example, the Fried frailty phenotype assessment takes 15-20 minutes [23], and the Rockwood Frailty Index requires answering a 70-item questionnaire [22], which can burden health care providers and patients, hindering the adoption of frailty assessments and raising concerns about their feasibility, particularly in high-volume and resource-limited settings.

The goal of this study is to create a novel method that utilizes both structured and unstructured real-world data, along with machine learning (ML) techniques, to construct a frailty index that is specific to patients with aortic stenosis who are scheduled to undergo a TAVR procedure.

Methods

Ethical Considerations

This retrospective study utilized perioperative data collected from January 2018 to December 2019, as part of a federal study (R01 AG055337, CP) approved by the institutional review board of the University of Florida, with a waiver for Health Insurance Portability and Accountability Act compliance and an honest broker for medical record retrieval. The research adhered to the Declaration of Helsinki and followed all the institutional protocols. Participants provided written informed consent at the time of data collection.

Study Population

The study extracted data on 14,000 patients from a deidentified historical database. A specific subset, the TAVR cohort, was identified using the Current Procedural Terminology code 33361, which narrowed the focus to 131 patients. All these patients had aortic stenosis as one of their symptoms.

Frailty Phenotype and Target Variables

A person is classified as frail if they meet 3 or more of the following conditions: (1) unintentional weight loss of 10 pounds or more over the past 6 months; (2) a sense of persistent exhaustion, characterized by a moderate feeling that normal activities were strenuous over the past week or a moderate inability to initiate activities; (3) a reduced pace in walking, as assessed by a nurse; (4) diminished grip strength, determined by a gender-specific T-score below −2.5, based on the highest grip strength recorded across 3 attempts using a Jamar hydraulic hand dynamometer (Model J00105, Lafayette Instrument Europe); and (5) a self-reported decrease in physical activity, gauged by the Duke Activity Status Index [24]. The scoring range

is from 0 to 5; individuals meeting fewer than 3 of these criteria are considered nonfrail. This frailty assessment is documented in electronic health records (EHRs) to assist in planning perioperative care. In the preoperative anesthesia clinic, both the nursing staff and anesthesia professionals undergo training to perform frailty assessments, guided by published criteria [21,25], alongside documenting educational background and conducting cognitive evaluations.

Topic Modeling

Latent Dirichlet allocation (LDA) is a statistical model used in the field of natural language processing to uncover the hidden thematic structure in extensive collections of documents. We used LDA to generate 150 topics on clinical notes collected 60 days before surgery. A total of 543,520 clinical notes from 18,513 patients were collected. We removed preoperative evaluation notes that documented the results of the Fried frailty phenotype score to avoid information contamination. During data preprocessing, we found that some clinical notes contained similar information about the patients. To overcome this issue, we calculated the similarity between clinical notes that share the same NOTE_KEY using cosine distance and removed the notes with a similarity score greater than 80%. Special characters were removed, and all uppercase letters were converted to lowercase. Three types of words from the clinical notes were removed: (1) high-frequency words that were not closely related to medical symptoms, (2) words with a frequency less than 30, and (3) common stop words. We further performed stemming to reduce the words to their root form to simplify and standardize the vocabulary. After the data preprocessing, a total of 170,731 notes from 18,012 patients were included. Topic modeling was based on 50,000 randomly selected clinical notes. Topics of the remaining clinical notes were predicted and aggregated to the individual level to represent if the patients had a mention of a certain topic.

Structured EHR Data

Preoperative attributes extracted from structured EHR data encompassed patient details available up to 6 months prior to their frailty evaluation. These included sociodemographic data (like age, gender, race, and years of education), medical history and severity (such as cancer history and the American Society of Anesthesiologists [ASA] physical status classification [26]), along with the latest biochemical markers typically recorded in preoperative settings (eg, hemoglobin and hematocrit levels).

Sociodemographic attributes captured information such as the patient's age at frailty evaluation, gender, race, marital and employment status, and educational attainment. We divided age into 3 categories: under 70 years, between 70 and 79 years, and over 80 years. Education level was quantified by the total years of formal education [27], with each year of progression counted and repetitions not contributing to the total. For instance, completion of high school was marked as 12 years, a bachelor's degree as 16 years, and so on. Medical history and acuity features included various medical conditions, the ASA physical status score, and the type of clinical encounter.

Medical conditions were binary indicators based on whether a patient had certain conditions, with the *International Classification of Diseases and Related Health Problems*, *Tenth Revision*, diagnosis codes condensed into Clinical Classifications Software categories for a more clinically relevant summarization. Conditions appearing in at least 10% of a broader cohort were selected as features for the ML models. The ASA score, ranging from 1 to 5, assessed patient health, with scores 1-2 indicating low-risk or mild systemic disease, and scores 3 and above signaling serious systemic illness.

Biochemical features included hematocrit, hemoglobin, and platelet count, with hemoglobin and hematocrit levels recommended for evaluation in older patients (65 years and above) undergoing significant surgery or those with a history of severe anemia undergoing minor surgery.

Integration of Topic Features and Structured EHR Features

We manually reviewed topics generated from topic modeling and removed those that were not related to medical findings, resulting in 121 topic features in total. TAVR patients without clinical notes were removed (n=34), resulting in 97/131 patients in the training and evaluating ML models. Table S1 in Multimedia Appendix 1 compares the demographics between patients who were removed due to the absence of clinical notes and those who were used for modeling. After combining features generated from topic modeling and those from structured EHR data, we performed least absolute shrinkage and selection operator (LASSO) regression to reduce the feature space. Different parameters for regularization strength were examined, and the one that achieved the highest area under the receiver operating characteristic curve (AUC) for predicting frailty was selected for feature selection. We further applied Extreme Gradient Boosting (XGBoost) on features selected from LASSO to predict frailty. XGBoost is an ensemble learning method that uses decision tree frameworks. It creates models to improve the efficacy of preceding trees by diminishing errors through gradient descent. The resulting frailty index from XGBoost is a continuous score between 0 and 1, representing the probability of a patient being predicted as frail; high scores indicate a greater likelihood of frailty in TAVR patients.

To assess the additive value of topic features derived from clinical notes, we also evaluated the performance of XGBoost on predicting frailty using only features from structured EHR data. We used the k-nearest neighbor imputation technique to impute missing values in numerical features with the mean value from the k-nearest neighbors (k=5) found in the training dataset. Missing values in categorical features were imputed as a new label (unknown). Numerical features were scaled to 0-1, and categorical features were encoded before training the ML model. Figure 1 illustrates the overview of the analytic pipeline.

Figure 1. Overview of the analytic pipeline. EHR: electronic health record; LASSO: least absolute shrinkage and selection operator; LDA: latent Dirichlet allocation; TAVR: transcatheter aortic valve replacement; XGBoost: Extreme Gradient Boosting.
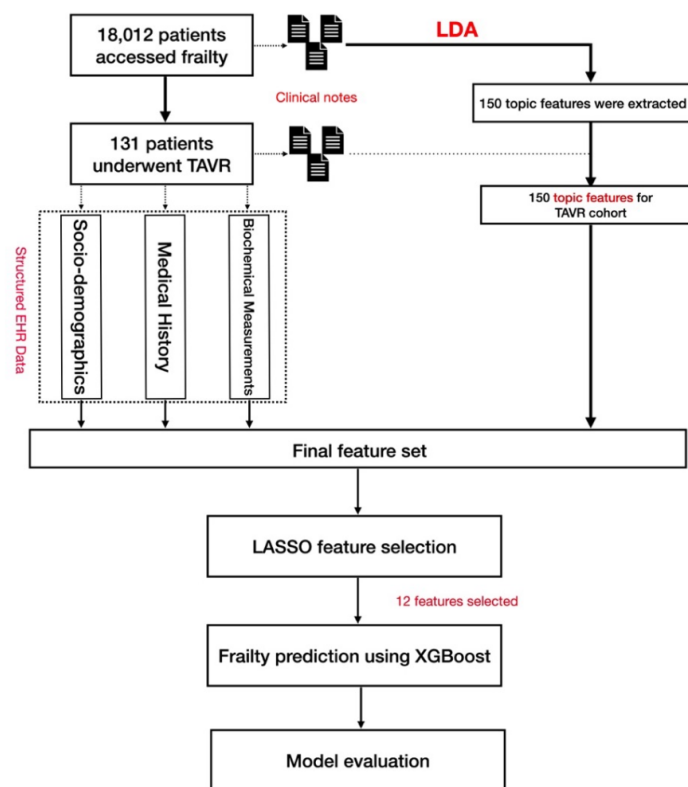
Model Performance and Evaluation

We employed a 5×5 nested cross-validation technique to evaluate our ML models. This technique consisted of 5 outer and 5 inner folds. During each outer fold, we set aside one-fifth of the patient records as an independent testing set, while the remaining four-fifths formed the training set. This outer training subset was then split into 5 inner folds for further validation. Each inner fold served as a stand-alone validation set, while the other 4 acted as the training set for the inner loop. The inner loop was responsible for training the models and fine-tuning the hyperparameters, taking a methodical approach in searching for the model's ideal hyperparameter settings. Meanwhile, the outer loop estimated errors and evaluated generalization capabilities. To optimize hyperparameters, we utilized a grid search strategy to systematically explore various combinations of predefined hyperparameters to train our models. We calculated and reported the mean and SD for metrics like AUC, accuracy, sensitivity, and specificity across the 5 outer folds. This thorough approach boosts confidence in the models' generalizability and scalability.

Model Interpretation and Feature Ranking

We used the Shapley Additive Explanations (SHAP) method to explain how our trained ML models work. SHAP is a widely used model-agnostic explanatory approach that helps in understanding the outputs of ML models. We created a SHAP summary plot to demonstrate the importance of features and their impact on the outcome. This impact is shown through a sign and magnitude: the SHAP value's sign indicates the direction of the feature's impact on the outcome (eg, a positive SHAP value indicates that the feature in question increases the likelihood of frailty), while its magnitude reflects the feature's predictive influence.

Statistical Analysis

Categorical features between the frail and nonfrail cohorts were compared using a χ^2 test for features with expected cell counts greater than 5 and a Fisher exact test for features with expected cell count less than or equal to 5 (ie, race). Age differences between the frail and nonfrail cohorts were compared using a Wilcoxson rank sum test. The significance level was set at $P<.05$.

Results

Table 1 compares patient characteristics between frail and nonfrail patients who underwent TAVR. The most significant difference observed was that frail patients were older and had lower hematocrit and hemoglobin levels compared to nonfrail patients. However, there were no significant differences in other characteristics such as gender, race, marital status, education, ASA score, or employment status.

Table 1. Comparison of patient characteristics between frail and nonfrail TAVR[a] patients.

Features	Frail (n=46)	Nonfrail (n=51)	*P* value
Age (years), mean (SD)	80.22 (7.45)	77.31 (6.05)	.04
Hematocrit (%), mean (SD)	36.16 (5.38)	39.93 (4.99)	<.001
Hemoglobin (g/DL), mean (SD)	12.14 (1.89)	13.4 (1.78)	.002
Platelet count (platelets/mL), mean (SD)	199.95 (86.34)	202.12 (51.91)	.88
Gender, n (%)			.69
Female (n=37)	19 (51.4)	18 (48.6)	
Male (n=60)	27 (45)	33 (55)	
Race, n (%)			.54
White (n=93)	43 (46.2)	50 (53.8)	
Others (n=4)	3 (75)	1 (25)	
Marital status, n (%)			.32
Married (n=65)	28 (43.1)	37 (56.9)	
Others (n=32)	18 (56.3)	14 (43.7)	
Years of education, n (%)			.41
Greater than 12 (n=58)	25 (43.1)	33 (56.9)	
Less than or equal to 12 (n=49)	21 (53.8)	28 (46.2)	
ASA[b] physical status score, n (%)			.47
Less than or equal to 3 (n=21)	8 (38.1)	13 (61.9)	
Greater than 3 (n=76)	38 (50)	38 (50)	
Employment status, n (%)			.79
Retired (n=76)	35 (46.1)	41 (53.9)	
Not retired (n=21)	11 (52.4)	10 (47.6)	

[a]TAVR: transcatheter aortic valve replacement.
[b]ASA: American Society of Anesthesiologists.

Table S2 in Multimedia Appendix 1 shows the number of features selected by LASSO alongside corresponding performance metrics across different regularization parameter values. Lower regularization parameter values (ie, inversed regularization strength) correspond to increased penalization of coefficient magnitudes and reduced features retained in the final model. Specifically, an inversed regularization strength of 0.25 imposed the highest penalty, leading to the smallest feature selection. Comparatively, an inversed regularization strength of 0.5 resulted in 12 selected features and achieved the highest AUC (0.67) and accuracy (0.65).

Table S3 in Multimedia Appendix 1 and Figure 2 show the performance of the ML models in identifying frailty. Results showed significant enhancement in the predictive performance when topic features derived from clinical notes were integrated with structured EHR data, compared to when only structured EHR data were used. Specifically, the combined topic + EHR model demonstrated a superior AUC of 0.82 (SD 0.07), indicating a robust ability to distinguish between frail and nonfrail individuals. This was a notable improvement over the EHR-only model, which achieved an AUC of 0.64 (SD 0.08).

Figure 3 shows the rank of features in predicting the frailty phenotype based on their SHAP values. Notably, the treatment and management of congestive heart failure (CHF) with carvedilol, spondylosis, intervertebral disc disorders, and other back problems and the treatment and management of atrial fibrillation (AFib) were found to have the strongest influence. These features exhibited high SHAP values, suggesting they strongly predict the frailty phenotype. Other significant factors included musculoskeletal symptoms, anemia, and age ≥80 years.

Figure 2. Receiver operating characteristic curves of frailty prediction using structured EHR data only and an integration of structured EHR data and topic features. Structured EHR features included sociodemographics, medical history and severity, and biochemical measurements. Topics features included features derived from clinical notes using latent Dirichlet allocation. The final features used were the ones selected from least absolute shrinkage and selection operator logistic regression. AUC: area under the receiver operating characteristic curve; EHR: electronic health record.
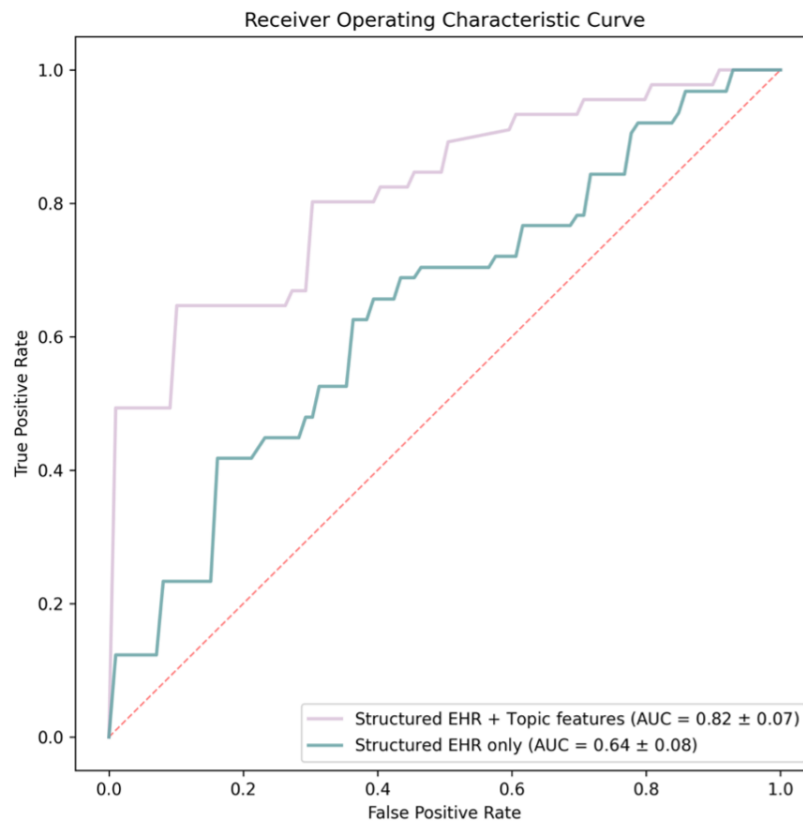
Figure 3. The rank of features in predicting the frailty phenotype. SHAP values greater than 0 (towards the right) contributed to predicting frailty, whereas SHAP values less than 0 contributed to predicting nonfrailty. For diagnosis and topic features, red dots indicate the presence, while blue dots indicate the absence of the condition. Red dots in "Years of education" represent patients with less than 12 years of education. Red dots in "Age >= 80" represent patients aged over 80 years old. SHAP: Shapley Additive Explanations.
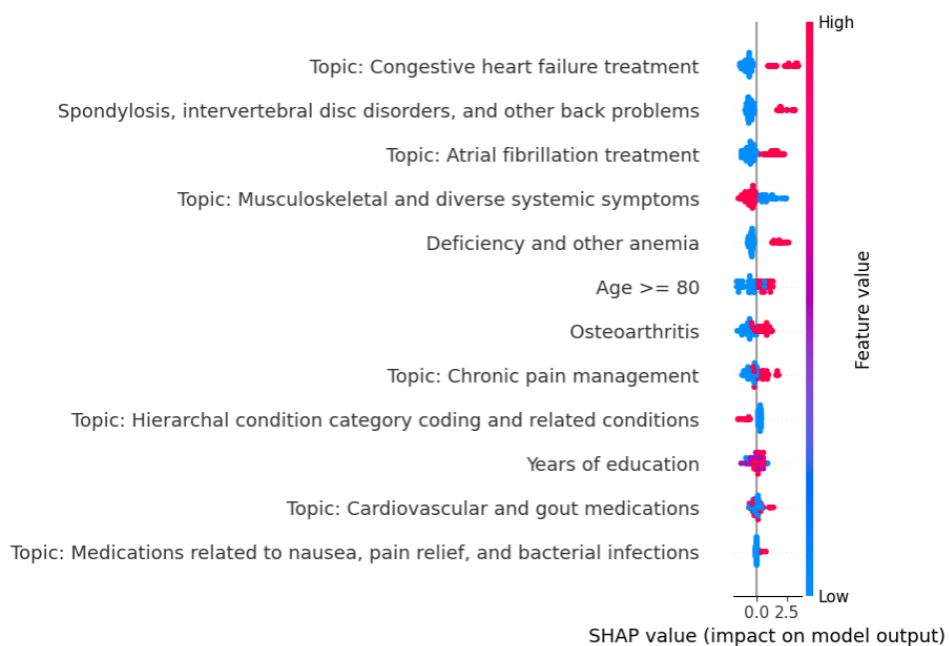
Table 2 presents 7 topics identified through LDA of clinical notes and further selected by LASSO. Of these, 5 topics related to the treatment and management of specific medical comorbidities in the cohort. The LDA algorithm identified underlying topics by clustering co-occurring terms across the clinical corpus. Each topic's top 10 weighted terms indicated which terms appeared most prominently within a given cluster. Higher term weights represented stronger associations between the term and the topic. For instance, the medications "Carvedilol" (term weight 0.097) and "Coreg" (term weight 0.094) had the highest weights within the first topic. Based on the prominence of these cardioprotective drugs, the topics were interpreted as medical treatment and management of CHF.

Table 2. Topics selected by least absolute shrinkage and selection operator regression for predicting the frailty phenotype in transcatheter aortic valve replacement patients, along with the top associated terms (stemmed) and their respective weights. The topics and terms listed are critical factors identified by the least absolute shrinkage and selection operator model that contribute to the prediction of frailty. The weights indicate the strength and direction of each term's association with the frailty phenotype, providing insights into the underlying factors influencing patient outcomes.

Topics	Details
Topic: Congestive heart failure treatment	0.097*"carvedilol"+0.094*"coreg"+ 0.056*"chf"+0.056*"lasix"+ 0.034*"bumex"+0.034*"spironolacton"+ 0.032*"cardiomyopathi"+0.029*"meal"+ 0.029*"bp"+0.027*"aldacton"
Topic: Atrial fibrillation treatment	0.168*"eliqui"+0.071*"fib"+ 0.069*"apixaban"+0.052*"xarelto"+ 0.047*"diltiazem"+0.035*"rate"+ 0.029*"stop"+0.028*"sotalol"+ 0.027*"hold"+0.025*"cardiovers"
Topic: Musculoskeletal and diverse systemic symptoms	0.058*"movement"+0.044*"swell"+ 0.041*"joint"+0.039*"musculoskelet"+ 0.033*"respiratori"+0.029*"extra"+ 0.026*"mucou"+0.026*"jaundic"+ 0.025*"psychiatr"+0.025*"rate"
Topic: Chronic pain management	0.197*"everi"+0.131*"tramadol"+ 0.091*"ultram"+0.080*"acetaminophen"+ 0.051*"nightli"+0.045*"ointment"+ 0.043*"appli"+0.038*"norco"+ 0.034*"bactroban"+0.032*"mupirocin"
Topic: Hierarchal condition category coding and related conditions	0.498*"hcc"+0.412*"code"+ 0.020*"mellitu"+0.014*"problem"+ 0.011*"anemia"+0.010*"list"+ 0.004*"neuropathi"+0.004*"pvd"+ 0.003*"coronari"+0.002*"defici"
Topic: Cardiovascular and gout medications	0.105*"tartrat"+0.102*"lopressor"+ 0.065*"gout"+0.064*"allopurinol"+ 0.057*"atorvastatin"+0.049*"zyloprim"+ 0.047*"lipitor"+0.042*"male"+ 0.042*"nightli"+0.036*"physic"
Topic: Medications related to nausea, pain relief, and bacterial infections	0.210*"everi"+0.073*"zofran"+ 0.060*"ondansetron"+ 0.043*"acetaminophen"+ 0.038*"amoxicillin"+0.034*"releas"+ 0.027*"odt"+0.027*"keflex"+ 0.025*"augmentin"+0.024*"delay"

Discussion

Principal Findings

This study developed a frailty index using ML techniques by combining structured and unstructured real-world data. Incorporating the features obtained from unstructured clinical notes and structured EHR data significantly improved the model's performance in identifying frailty. SHAP revealed that cardiovascular conditions, specifically CHF and AFib, as well as musculoskeletal disorders, were the primary factors that predict frailty among patients. This is consistent with the mainstream of clinical evidence [28-30]. The prominence of age as a predictor supports the well-established association between advancing age and increased vulnerability to frailty and limited quality of life improvement post TAVR [31]. The impact of years of formal education on frailty prediction was relatively lower (ranked 10th in importance), indicating that while socioeconomic factors contribute to health outcomes, clinical characteristics are more determinative in the context of frailty. This clinical ML application can easily be integrated into pre-TAVR settings to guide perioperative management, heart team discussions, and resource utilization. Our model provides actionable and semiquantitative insights and can potentially highlight targets for optimization (eg, anemia). In addition, the data processing and derivation pipeline we practically demonstrate allows ML model retraining to accommodate updates in practice and interinstitutional context differences.

Structured real-world data that contain rich historical and current information across various domains (eg, diagnoses, procedures, demographics) have been leveraged to predict frailty. Segal et al [32] constructed ML models using medical claims data to predict the Fried frailty phenotype, achieving an AUC of 0.76 with a gradient boosting machines algorithm. Similarly, Le Pogam et al's [33] exploration using inpatient discharge data obtained an AUC of 0.71, while Bai et al [34] harnessed structured EHR data to predict preoperative frailty, achieving an AUC of 0.74 using XGBoost. Notably, a considerable amount of information contained within EHRs exists as unstructured text. Recent advancements in ML and natural language processing have enabled researchers to harness this valuable resource, which traditionally posed substantial challenges for manual analysis. For instance, Shao et al [35] used topic modeling to extract frailty-related topics from clinical notes and defined frailty using the number of frailty-related topics present. This frailty index was associated with a 2-fold increased risk of the composite outcome combining 1-year hospitalization and mortality. Further, Chen et al [36] and Martin et al [37] developed deep learning models to predict sentences indicating geriatric syndromes or actionable frailty aspects and achieved an F_1-score of 0.61 and a scaled Brier score of 0.52. Despite these advances, a gap remains in the literature concerning the integration of structured and unstructured real-world data for frailty prediction. Moreover, the development of frailty indices tailored to preoperative contexts, especially for TAVR patients, remains unexplored.

Our results demonstrated that clinical notes serve as a complement to structured EHR data in predicting preoperative frailty. Incorporating features extracted from clinical notes significantly improved the model's performance over using only structured EHR data, with the integrated model using only 12 features achieving an AUC of 0.82 compared to 0.64 for the EHR-only model in the test set. This performance is superior to the one-size-fits-all statistical models such as the Society of Thoracic Surgeons score, which has an AUC of 0.64-0.7 in risk stratifying patients [38-40].

Identifying the key topics and associated terms provides actionable insights into understanding and managing frailty-associated factors. For instance, knowing that the patient has heart failure requiring cardiac medications and contributing to frailty may trigger putting the patient on a more intensive pre- and post-TAVR optimization pathway, such as right heart catheterization, to optimize heart failure therapy preprocedures. Similar approaches can be applied to other cardiovascular conditions or musculoskeletal disorders that are partially correctable with focused pre- or postprocedural conditioning. The topic of CHF treatment and management, with high-weight terms like "Carvedilol," can highlight high systemic vascular resistance hemodynamic phenotypes, systolic heart failure, or arrhythmia requiring beta blockade. Similarly, the topic focusing on AFib management, with terms like "Eliquis" and "Apixaban," highlights the need to evaluate and manage arrhythmias, which are risk factors for stroke and also are associated with obesity, hypoventilation, and chronically low physical and cardiopulmonary conditioning. The specificity of terms and their weights in each topic provides a framework for health care professionals to identify and prioritize areas of concern in pre-TAVR patients, ensuring a comprehensive and tailored approach to frailty assessment. Identifying "high-risk candidates" is crucial in determining a patient's readiness for TAVR and optimizing postprocedural outcomes by addressing the multifaceted aspects of frailty and medical complexity in this vulnerable patient population [41,42]. A considerable advantage to using ML models is that they are translatable to new TAVR techniques, adaptable to clinical settings and automated facilitating implementation, and updatable to accommodate shifts in populations and practices.

Compared to the traditional frailty measurements, a key advantage of our developed method is its applicability to newly admitted patients with severe aortic stenosis being considered for TAVR. The model can process a patient's available information, including both structured EHR data and unstructured clinical notes, to generate a frailty index score once the patient is admitted. This assessment provides an objective measure of a patient's frailty status, which has the potential to enhance the clinical shared decision-making between physicians and the patient. By offering an early and timely indication of a patient's frailty status, this approach allows for a more informed discussion about the risks and benefits of TAVR, potential alternative treatments, and the need for additional preoperative interventions and postoperative care plans, ultimately improving the care of TAVR patients.

This study, while insightful, acknowledges specific limitations. It utilizes retrospective EHR data, a method that provides a valuable foundation for analysis despite its inherent limitations, such as the potential for minor reporting inconsistencies, recall variations, and lead time considerations. The study's focus on frailty, although a topic of ongoing discussion due to its varied definitions and intersection with medical complexity, enriches the discourse by offering a nuanced view of its implications. Additionally, while our ML approach effectively integrates structured and unstructured real-world data to predict frailty, it does not extensively address potential confounding factors such as comorbidities, medications, or socioeconomic status. This focus on predictive modeling, rather than causal analysis, may limit the ability to fully account for these variables. Furthermore, this study has not undergone external validation; however, this aspect is mitigated by the rigorous internal review processes, which reduce the risk of overfitting and ensure the model's preliminary reliability. Rather than significantly limiting, these considerations underscore the study's contribution to ongoing research and its potential as a springboard for further investigation and external validation efforts.

Conclusion

This study demonstrates the effectiveness of combining unstructured clinical notes with structured EHR data to assess frailty in TAVR patients. This approach can potentially improve risk stratification for TAVR, adding insight into targets for perioperative optimization and potentially improving TAVR outcomes. Future efforts should integrate these models into clinical settings to improve TAVR planning and patient care, emphasizing continuous medical education and innovation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary material.
[DOCX File (Microsoft Word File), 17 KB-Multimedia Appendix 1]

References

1. Osnabrugge RLJ, Mylotte D, Head SJ, et al. Aortic stenosis in the elderly: disease prevalence and number of candidates for transcatheter aortic valve replacement: a meta-analysis and modeling study. J Am Coll Cardiol. Sep 10, 2013;62(11):1002-1012. [doi: 10.1016/j.jacc.2013.05.015] [Medline: 23727214]

2. Kim DH, Afilalo J, Shi SM, et al. Evaluation of changes in functional status in the year after aortic valve replacement. JAMA Intern Med. Mar 1, 2019;179(3):383-391. [doi: 10.1001/jamainternmed.2018.6738] [Medline: 30715097]

3. Carabello BA, Paulus WJ. Aortic stenosis. Lancet. Mar 14, 2009;373(9667):956-966. [doi: 10.1016/S0140-6736(09)60211-7] [Medline: 19232707]

4. Bevan GH, Zidar DA, Josephson RA, Al-Kindi SG. Mortality due to aortic dtenosis in the United States, 2008-2017. JAMA. Jun 11, 2019;321(22):2236-2238. [doi: 10.1001/jama.2019.6292] [Medline: 31184728]

5. Makkar RR, Thourani VH, Mack MJ, et al. Five-year outcomes of transcatheter or surgical aortic-valve replacement. N Engl J Med. Jan 29, 2020;382(9):799-809. [doi: 10.1056/NEJMoa1910555] [Medline: 31995682]

6. Puri R, Iung B, Cohen DJ, Rodés-Cabau J. TAVI or no TAVI: identifying patients unlikely to benefit from transcatheter aortic valve implantation. Eur Heart J. Jul 21, 2016;37(28):2217-2225. [doi: 10.1093/eurheartj/ehv756] [Medline: 26819226]

7. Arnold SV. Calculating risk for poor outcomes after transcatheter aortic valve replacement. J Clin Outcomes Manag. May 2019;26(3):125-129. [Medline: 31427851]

8. Li Z, Dawson E, Moodie J, et al. Measurement and prognosis of frail patients undergoing transcatheter aortic valve implantation: a systematic review and meta-analysis. BMJ Open. Mar 4, 2021;11(3):e040459. [doi: 10.1136/bmjopen-2020-040459] [Medline: 33664067]

9. Otto CM, Kumbhani DJ, Alexander KP, et al. 2017 ACC Expert Consensus Decision Pathway for transcatheter aortic valve replacement in the management of adults with aortic stenosis: a report of the American College of Cardiology Task Force on Clinical Expert Consensus Documents. J Am Coll Cardiol. Mar 14, 2017;69(10):1313-1346. [doi: 10.1016/j.jacc.2016.12.006] [Medline: 28063810]

10.   Kim DH, Kim CA, Placide S, Lipsitz LA, Marcantonio ER. Preoperative frailty assessment and outcomes at 6 months or later in older adults undergoing cardiac surgical procedures: a systematic review. Ann Intern Med. Nov 1, 2016;165(9):650-660. [doi: 10.7326/M16-0652] [Medline: 27548070]

11.   Afilalo J, Lauck S, Kim DH, et al. Frailty in older adults undergoing aortic valve replacement: the FRAILTY-AVR study. J Am Coll Cardiol. Aug 8, 2017;70(6):689-700. [doi: 10.1016/j.jacc.2017.06.024] [Medline: 28693934]

12.   Sepehri A, Beggs T, Hassan A, et al. The impact of frailty on outcomes after cardiac surgery: a systematic review. J Thorac Cardiovasc Surg. Dec 2014;148(6):3110-3117. [doi: 10.1016/j.jtcvs.2014.07.087] [Medline: 25199821]

13.   Goldfarb M, Bendayan M, Rudski LG, et al. Cost of cardiac surgery in frail compared with nonfrail older adults. Can J Cardiol. Aug 2017;33(8):1020-1026. [doi: 10.1016/j.cjca.2017.03.019] [Medline: 28754387]

14.   van Steenbergen GJ, van Veghel D, Schulz DN, et al. Better survival after transcatheter aortic valve replacement by process improvements. Neth Heart J. Apr 2021;29(4):193-200. [doi: 10.1007/s12471-020-01526-7] [Medline: 33284420]

15.   Hinterbuchner L, Strohmer B, Hammerer M, Prinz E, Hoppe UC, Schernthaner C. Frailty scoring in transcatheter aortic valve replacement patients. Eur J Cardiovasc Nurs. Oct 2016;15(6):384-397. [doi: 10.1177/1474515115596640] [Medline: 26216870]

16.   Shroyer ALW, Coombs LP, Peterson ED, et al. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. Ann Thorac Surg. Jun 2003;75(6):1856-1864. [doi: 10.1016/s0003-4975(03)00179-6] [Medline: 12822628]

17.   Nashef SAM, Sharples LD, Roques F, Lockowandt U. EuroSCORE II and the art and science of risk modelling. Eur J Cardiothorac Surg. Apr 2013;43(4):695-696. [doi: 10.1093/ejcts/ezs468] [Medline: 22922694]

18.   Morley JE, Vellas B, van Kan GA, et al. Frailty consensus: a call to action. J Am Med Dir Assoc. Jun 2013;14(6):392-397. [doi: 10.1016/j.jamda.2013.03.022] [Medline: 23764209]

19.   Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. JAMA. Sep 21, 1963;185:914-919. [doi: 10.1001/jama.1963.03060120024016] [Medline: 14044222]

20.   Spertus JA, Jones PG. Development and validation of a short version of the Kansas City Cardiomyopathy Questionnaire. Circ Cardiovasc Qual Outcomes. Sep 2015;8(5):469-476. [doi: 10.1161/CIRCOUTCOMES.115.001958] [Medline: 26307129]

21.   Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. Mar 2001;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]

22.   Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol Sci Med Sci. Jul 2007;62(7):722-727. [doi: 10.1093/gerona/62.7.722] [Medline: 17634318]

23.   Kim H, Higgins PA, Canaday DH, Burant CJ, Hornick TR. Frailty assessment in the geriatric outpatient clinic. Geriatr Gerontol Int. Jan 2014;14(1):78-83. [doi: 10.1111/ggi.12057] [Medline: 23530477]

24.   Hlatky MA, Boineau RE, Higginbotham MB, et al. A brief self-administered questionnaire to determine functional capacity (the Duke Activity Status Index). Am J Cardiol. Sep 15, 1989;64(10):651-654. [doi: 10.1016/0002-9149(89)90496-7] [Medline: 2782256]

25.   Amini S, Crowley S, Hizel L, et al. Feasibility and rationale for incorporating frailty and cognitive screening protocols in a preoperative anesthesia clinic. Anesth Analg. Sep 2019;129(3):830-838. [doi: 10.1213/ANE.0000000000004190] [Medline: 31425227]

26.   Horvath B, Kloesel B, Todd MM, Cole DJ, Prielipp RC. The evolution, current value, and future of the American Society of Anesthesiologists physical status classification system. Anesthesiology. Nov 1, 2021;135(5):904-919. [doi: 10.1097/ALN.0000000000003947] [Medline: 34491303]

27.   Price C, Cole DJ. Adding 'Formal Years of Education' to patients' preoperative screening. ASA Monitor. May 1, 2023;87(5):1-4. [doi: 10.1097/01.ASM.0000935232.57361.7f]

28.   Damluji AA, Rodriguez G, Noel T, et al. Sarcopenia and health-related quality of life in older adults after transcatheter aortic valve replacement. Am Heart J. Jun 2020;224:171-181. [doi: 10.1016/j.ahj.2020.03.021] [Medline: 32416332]

29.   Krishnan A, Suarez-Pierre A, Zhou X, et al. Comparing frailty markers in predicting poor outcomes after transcatheter aortic valve replacement. Innovations (Phila). Feb 2019;14(1):43-54. [doi: 10.1177/1556984519827698] [Medline: 30848712]

30.   Zahid S, Din MTU, Khan MZ, et al. Trends, predictors, and outcomes of 30-day readmission with heart failure after transcatheter aortic valve replacement: insights from the US nationwide readmission database. J Am Heart Assoc. Aug 16, 2022;11(16):e024890. [doi: 10.1161/JAHA.121.024890] [Medline: 35929464]

31.   Attinger-Toller A, Ferrari E, Tueller D, et al. Age-related outcomes after transcatheter aortic valve replacement: insights from the SwissTAVI Registry. JACC Cardiovasc Interv. May 10, 2021;14(9):952-960. [doi: 10.1016/j.jcin.2021.02.031] [Medline: 33865734]

32. Segal JB, Chang HY, Du Y, Walston JD, Carlson MC, Varadhan R. Development of a claims-based frailty indicator anchored to a well-established frailty phenotype. Med Care. Jul 2017;55(7):716-722. [doi: 10.1097/MLR.0000000000000729] [Medline: 28437320]

33. Le Pogam MA, Seematter-Bagnoud L, Niemi T, et al. Development and validation of a knowledge-based score to predict Fried's frailty phenotype across multiple settings using one-year hospital discharge data: the electronic frailty score. E Clin Med. Jan 10, 2022;44:101260. [doi: 10.1016/j.eclinm.2021.101260] [Medline: 35059615]

34. Bai C, Al-Ani M, Amini S, et al. Developing and validating an electronic health record-based frailty index in pre-operative settings using machine learning. J Intell Inf Syst. Oct 14, 2023;62(2):339-354. [doi: 10.1007/s10844-023-00818-9]

35. Shao Y, Mohanty AF, Ahmed A, et al. Identification and use of frailty indicators from text to examine associations with clinical outcomes among patients with heart failure. AMIA Annu Symp Proc. Feb 10, 2017;2016:1110-1118. [Medline: 28269908]

36. Chen T, Dredze M, Weiner JP, Kharrazi H. Identifying vulnerable older adult populations by contextualizing geriatric syndrome information in clinical notes of electronic health records. J Am Med Inform Assoc. Aug 1, 2019;26(8-9):787-795. [doi: 10.1093/jamia/ocz093] [Medline: 31265063]

37. Martin JA, Crane-Droesch A, Lapite FC, et al. Development and validation of a prediction model for actionable aspects of frailty in the text of clinicians' encounter notes. J Am Med Inform Assoc. Dec 28, 2021;29(1):109-119. [doi: 10.1093/jamia/ocab248] [Medline: 34791302]

38. Reiff C, Gurevich S, Bertog S, Sorajja P, Kelly R, Garcia S. Validation of STS/ACC TVT-TAVR score in veterans undergoing transcatheter aortic valve replacement. J Invasive Cardiol. Dec 2018;30(12):447-451. [Medline: 30218556]

39. D'Ascenzo F, Capodanno D, Tarantini G, et al. Usefulness and validation of the survival post TAVI score for survival after transcatheter aortic valve implantation for aortic stenosis. Am J Cardiol. Dec 15, 2014;114(12):1867-1874. [doi: 10.1016/j.amjcard.2014.09.031] [Medline: 25438915]

40. Al-Azizi K, Shih E, DiMaio JM, et al. Assessment of TVT and STS risk score performances in patients undergoing transcatheter aortic valve replacement. J Soc Cardiovasc Angiogr Interv. Apr 21, 2023;2(3):100600. [doi: 10.1016/j.jscai.2023.100600] [Medline: 39130722]

41. Rogers T, Thourani VH, Waksman R. Transcatheter aortic valve replacement in intermediate- and low-risk patients. J Am Heart Assoc. May 12, 2018;7(10):e007147. [doi: 10.1161/JAHA.117.007147] [Medline: 29754127]

42. Reyes M, Reardon MJ. Transcatheter valve replacement: risk levels and contemporary outcomes. Methodist Debakey Cardiovasc J. 2017;13(3):126-131. [doi: 10.14797/mdcj-13-3-126] [Medline: 29743997]

Abbreviations

AFib: atrial fibrillation
ASA: American Society of Anesthesiologists
AUC: area under the receiver operating characteristic curve
CHF: congestive heart failure
EHR: electronic health record
LASSO: least absolute shrinkage and selection operator
LDA: latent Dirichlet allocation
ML: machine learning
SHAP: Shapley Additive Explanations
TAVR: transcatheter aortic valve replacement
XGBoost: Extreme Gradient Boosting