2016 IEEE 17th International Conference on Information Reuse and Integration

Statistical measurement and analysis on how the Late-Life Function & Disability Instrument enhances the frailty assessment compared to the national standards used on transcatheter aortic valve patients

Yunkai Liu, Ph.D.
Associate Professor
Gannon University
LIU006@gannon.edu

A. Kate MacPhedran, PT, Ph.D., CCS
Assistant Professor
Gannon University
MACPHEDR001@gannon.edu

Yu Luo, M.S.
Graduate Assistant
Gannon University
LUO001@knights.gannon.edu

Abstract

Transcatheter aortic valve replacement (TAVR) has become a more utilized procedure to perform on patients deemed too frail to handle the demands of a standard open heart approach. How to determine frailty in cardiac patients, particularly TAVR candidates, has been difficult to objectively quantify. The purpose of this research was to statistically measure which outcome tool most accurately depicted frailty in patients who underwent TAVR. Our study was performed based on the comparison between two approaches: the Kansas City Cardiomyopathy Questionnaire (KCCQ), which is the current national assessment standard conducted, and the Fried scale, which tests five frailty domains: gait speed, grip strength, low physical activity, exhaustion, and weight loss. Each domain of the Fried scale was explored and compared alongside the KCCQ with those frail/not frail to the TAVR patients with complications and deaths. Low physical activity was the strongest single-frailty domain predictor. Three statistical models – Logistic Regression, Support Vector Machines (SVM), and Artificial Neural Network (ANN), were used to build classification systems to predict complication conditions. Comparing static numbers, such as Sensitivity, Specificity and Area under Curve (AUC), it is believed that composed models based on the five domains of the Fried scale were able to demonstrate more accurate results than the traditional KCCQ approach. Both SVM and ANN showed significant performance, but further research is necessary to confirm specificity for the Fried scale with the TAVR population.

Keywords: Late-Life Function and Disability Instrument; Kansas City Cardiomyopathy Questionnaire; Fried's frailty phenotype; Logistic Regression; Artificial Neural Networks Model; Support Vector Machines

1. Introduction

The transcatheter aortic valve replacement (TAVR), was introduced in 2002 as an alternative, less invasive approach than traditional open sternotomy aortic valve replacement surgery, and intended specifically for individuals considered too "frail" to withstand traditional approach surgery [1,2]. Research and advancing technology has since led to the emergence of the transapical and transaortic catheter approach in replacing the valve to further ameliorate the risk of morbidity and mortality in this high-risk patient population [1]. A potential candidate for TAVR is assessed through a multi-disciplinary approach which is based predominantly on extensive cardiac testing, surgical feasibility and the various approaches [1, 3]. Surgeons also rely, in part, on a preoperative frailty assessment to help determine appropriateness for TAVR versus open approach and aid them in consideration of short and long-term postoperative survival [3, 4].

Kansas City Cardiomyopathy Questionnaire (KCCQ-12)

The following questions refer to your **heart failure** and how it may affect your life. Please read and complete the following questions. There are no right or wrong answers. Please mark the answer that best applies to you.

1. **Heart failure** affects different people in different ways. Some feel shortness of breath while others feel fatigue. Please indicate how much you are limited by **heart failure** (shortness of breath or fatigue) in your ability to do the following activities over the past 2 weeks.

Activity	**Extremely** Limited	**Quite a bit** Limited	**Moderately** Limited	**Slightly** Limited	**Not at all** Limited	Limited for other reasons or did not do the activity
a. Showering/bathing	O	O	O	O	O	O
b. Walking 1 block on level ground	O	O	O	O	O	O
c. Hurrying or jogging (as if to catch a bus)	O	O	O	O	O	O
	1	2	3	4	5	6

2. Over the past 2 weeks, how many times did you have **swelling** in your feet, ankles or legs when you woke up in the morning?

Every morning	3 or more times per week but not every day	1-2 times per week	Less than once a week	Never over the past 2 weeks
O	O	O	O	O
1	2	3	4	5

3. Over the past 2 weeks, on average, how many times has **fatigue** limited your ability to do what you wanted?

All of the time	Several times per day	At least once a day	3 or more times per week but not every day	1-2 times per week	Less than once a week	Never over the past 2 weeks
O	O	O	O	O	O	O
1	2	3	4	5	6	7

4. Over the past 2 weeks, on average, how many times has **shortness of breath** limited your ability to do what you wanted?

All of the time	Several times per day	At least once a day	3 or more times per week but not every day	1-2 times per week	Less than once a week	Never over the past 2 weeks
O	O	O	O	O	O	O
1	2	3	4	5	6	7

5. Over the past 2 weeks, on average, how many times have you been forced to sleep sitting up in a chair or with at least 3 pillows to prop you up because of **shortness of breath**?

Every night	3 or more times per week but not every day	1-2 times per week	Less than once a week	Never over the past 2 weeks
O	O	O	O	O
1	2	3	4	5

FIGURE 1: Sample Questions of KCCQ

Frailty is multifactorial in nature and to date there has not been a universal definition agreed upon for this

978-1-5090-3207-5/16 $31.00 © 2016 IEEE
DOI 10.1109/IRI.2016.64

concept or a standardized method utilized to accurately identify who is frail [3, 5-7]. For this study, frailty is defined as a decline in physiological systems which are evident in one's physical functions and activities of daily living. The Kansas City Cardiomyopathy Questionnaire (KCCQ) is one of the most widely used health-related quality-of-life measures for patients with congestive heart failure [8, 9]. The KCCQ has been used in hundreds of clinical trials which have involved thousands of patients [8]. There is an inverse relationship between the KCCQ score and New York Heart Association (NYHA) classification, in that patients with low KCCQ scores indicate more advanced congestive heart failure symptoms and decreased quality of life [10, 11]. This valid and reliable tool is used to track progression of patients' conditions, when heart muscle has weakened due to prior heart attacks, heart valve problems, infections and the like, however, its symptoms section is specifically written for heart failure patients [5, 6]. Some sample questions of KCCQ are shown in Figure 1.

With the KCCQs fairly strong correlation to the New York Heart Association Classifications [10, 11], this questionnaire has more recently been used to estimate functional status [9, 12, 13]. The 5-meter walk test and KCCQ are the standard "frailty" measures required by the national Transcatheter Valve Therapy (TVT) Registry for patients with severe aortic stenosis undergoing TAVR [6, 9, 14]. These tools are used to help estimate TAVR surgical risk and eligibility for the necessity of TAVR surgery, as well as help predict surgical outcomes and track postoperative progress [14, 15-18]. Every TAVR candidate must have at least severe aortic stenosis, however, not every patient with severe aortic stenosis has cardiomyopathy and congestive heart failure [18, 19]. Thus, the KCCQ may fail to accurately capture frailty in this population.

Fried et al. proposed a different approach to assessing frailty, referred to clinically as the Fried scale, which is based on a frailty phenotype [7, 20]. The Fried scale's frailty phenotype is comprised of five domains: low physical activity, slow walk speed, unintended weight loss, exhaustion, and weak grip strength [7, 20]. Fried's frailty phenotype has been proven to be a valid frailty measurement and recognized as one the most commonly utilized measure of frailty [6, 16, 21].

The measurements for Fried's frailty scale are obtained through a combination of physiological tests and self-reported questionnaires [8]. To date, the frailty domain "low physical activity" has had wide variability in how it has been measured (E.g., Minnesota Leisure Time Activities Questionnaire; Katz Index of activities of daily living) [2, 3, 7, 20]. The Late-Life Function and Disability Instrument (LLFDI) is a well-established, valid and reliable outcome tool which calculates functional limitation and has been tested on 60 to 90 year olds and older adults with cardiovascular disease and status-post cardiac surgery, making the LLFDI an ideal tool to use for "low physical activity" measurement [22-26].

A brief description of each of the five frailty domains tested is listed as follows.

- Walk Speed – time to walk 15 feet at a comfortable pace. Abbreviated as "Gait" in this paper.
- Exhaustion – based on responses to questions about energy level and effort. Abbreviated as "Exha" in this paper.
- Weight Loss – unintentional weight loss (≥ 10# in the last 12 months); answered as yes/no. Abbreviated as "WeLoss" in this paper.
- Grip Strength – measured using a hand-held dynamometer squeezed with the dominant hand. Abbreviated as "Grip" in this paper.
- Low Physical Activity – based on routine physical activities and daily tasks using the function component of the LLFDI [22]. Abbreviated as "LPA" in this paper.

The purpose of this study was to determine the predictive power of two different measurement tools in depicting frailty in patients with severe aortic stenosis who undergo TAVR. Every domain of the Fried scale is explored in detail and compared against the KCCQ with respect to patient complications and deaths. Three data mining models – Logistic Regression, Support Vector Machines (SVM), and Neural Network, were utilized to build classification systems in order to predict complication conditions. Comparing static numbers such as Sensitivity, Specificity and Area under Curve (AUC), it is believed that composed models based on the five domains of Fried's frailty scale are able to demonstrate more accurate results than the traditional KCCQ approach.

Abbreviations and Acronyms
Gait = walk speed
Exha = exhaustion
WeLoss = weight loss (unintentional)
Grip = grip strength
LPA = low physical activity
LLFDI = Late-Life Function & Disability Instrument
Resp = refers to complications or death status
SVM = Support Vector Machines
ROC = Receiver Operating Characteristic (curve)
AUC = Area Under Curve, implies area under ROC curve

2. Data Collection

A retrospective cohort design was used to assess frailty data on 70 high-risk patients who were referred to Saint Vincent Hospital's multi-disciplinary TAVR team for possible TAVR procedure, between April, 2013 and October, 2014. Patients who were under 65 years, did not communicate fluently in English, or ended up not undergoing TAVR surgery for any reason (25 of the 70 referrals), were excluded from the study. Informed consent was originally obtained on all 70 patients to participate in the TAVR work-up process which included undergoing a preoperative frailty assessment. Institutional Review Board approval was obtained from both Saint Vincent Hospital and Gannon University, as this was a collaborative venture. All individual patient information was de-identified before any analysis conducted.

Frailty assessment measurements were based on the frailty phenotype operationalized by Fried et al [7, 20] which is made up of slow gait speed, weak grip strength, shrinkage (weight loss), exhaustion, and low physical activity. Patients with impairments in at least three of the five domains were considered frail. Each patient had grip strength measured (average of 3 trials in kg) using a Jaymar dynamometer and patients whose average scores were in the lowest 20% of community older adults, as based on gender and body mass index (on average: <30 kg for men, <19 kg for women), met criteria as having weak grip strength. Five meter walk speed was calculated (average of 3 trials) with patients cued to walk at a comfortable pace with or without a device. Gait speeds in the lowest 20% of community older adults as based on sex and height cut-offs (either > 6 or 7 seconds) met criteria for slow gait. Patients also answered 32 standardized questions on routine physical activities and daily tasks using the LLFDI, which asks "How much difficulty do you have…?" on a 1-5 Likert scale which was then converted to 0-100 point weighted scale, with a lower scale indicating increased limitation in physical mobility and activities of daily living, with a cut-off score at 52.5. Exhaustion was assessed according to self-report, using two standardized questions obtained from the Center for Epidemiologic Studies Depression Scale: ''How often in the last week did you feel (1) everything was an effort; (2) you could not get going?'' Answering either question as "3 or more days of the week" was considered "low endurance/exhausted," and met criteria. Unintentional weight loss of 10 or more pounds within the last year was self-reported and anyone noting such was considered having met the criterion for shrinkage/impaired nutrition.

Patients were also assessed according to the KCCQ which is a 12 question self-report on their physical function, congestive heart failure symptoms, quality of life, and social limitations [9]. The KCCQ is on a 1-5 or 1-7 Likert scale (depending on the question) which is converted to 0-100 summary score [9].

Baseline demographic information as well as all-cause mortality and complications which occurred up to 30 days postoperatively were assessed retrospectively via chart reviews. Postoperative complications included major bleeding, minor vascular, major vascular, stroke, new arrhythmia (requiring pacemaker), and acute renal injury, which are the reportable complications per TVT Registry criteria [27]. Verification of chart information (E.g., corrective vascular procedures) was confirmed by TAVR team members and all-cause mortality at 30 days was assessed per follow-up visit notes or phone calls (if visits were missed) by TAVR team members.

TABLE 1: Patient Data Demographics vs. TVT Registry Data

Patient Data (% or mean)	TVT Registry Data (% or mean)
n = 45 (from 04/13–10/14)	**n = 12,182** (from11/11–06/13)
Age 81.5 (60 – 95 range)	Age 84 (79 – 88 range)
Female 22 (48.9%)	Female n = 6,316 (51.9%)
Caucasian 45 (100%)	Caucasian n = 11, 615 (95.3%)
Chronic CHF 36 (80%)	30 day mortality = 7%
TIA or Stroke 9 (20%)	
Cancer 8 (17.8%)	
Chronic renal failure 18 (40%)	**PARTNER Trial Data* (n =348)**
Stroke at 30 days 2 (4.4%)	30 day stroke incidence = *4.1%
Arrhythmia 3 (6.7%)	Pacemaker needed = *6.4%
Bleeding 2 (4.4%)	Bleeding = *15.7%
Minor Vascular 4 (8.9%)	Minor vascular = n/a
Major Vascular 1 (2.2%)	Major and minor vascular = *11.3%
NYHA Class III/IV 43 (95.6%)	NYHA Class III/IV = *94.3%
30-Day Mortality 3 (6.7%)	30 day mortality = **5%
30-Day Complications 12 (26.7%)	30-Day Complications = n/a

TIA = Transient Ischemic Attack; CHF = Congestive Heart Failure
NYHA = New York Heart Association (Classifications);
TVT = Transcatheter Valve Therapy (Registry)
*PARTNER Trial = National study using TVT Registry to estimate risks
** = mortality rate was based on high-risk TAVR patients

3. Results - Baseline Demographics

Descriptive analysis from chart reviews were conducted on those identified as frail and on the entire sample, to compare to the surgical population as a whole. Baseline demographics (see Table 1) revealed that the study population was comprised entirely of Caucasian individuals, ranging in age from 60 to 95 (mean was 81.5) and females made up 48.9%. Base on the data of those that underwent TAVR (n = 45), at 30 days postoperatively, a total of 2 patients (4.4%) had suffered a stroke, 2 patients had bleeding issues, and 5 patients (11.1%) had vascular issues, either minor or major, with the latter requiring surgical repair. Thirty-day complications overall were noted in 12 patients (26.7%) and the mortality rate at 30-days was 6.7% (3 patients).

4. Each Domain of the Fried Scale

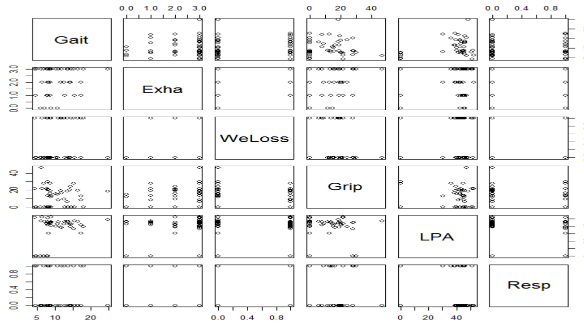

FIGURE 2: Data Distribution of Five Domains and Resp

Terms in Figure 2 and related static values are listed as follows.

- Gait is the score of gait domain. The range was from 0 to 30 (seconds). The higher score means slower gait speed with frail cut-off at/above 6 seconds. In our dataset, the mean was 10.52 seconds and the median was 8.77 seconds.
- Exha is the score of exhaustion domain. The range is from 0 to 3 (Likert scale), with 0 is the best condition, and 3 is the worst condition. In our dataset, the mean was 2.16 and the median was 3.00.
- WeLoss is the score of weight loss domain. 0 means negative for unexpected weight loss, and 1 means positive. In our dataset, the mean was 0.42 and the median was 0.00.
- Grip is the score of grip domain. The range was from 0 to 50 (kg). The higher score means lower grip strength with gross cut-off <30kg (for men) and <19kg (for women) as frail. In our dataset, the mean was 12.95 and the median was 13.67.
- LPA is the score of low physical activity domain. The range was from 0 to 100 (converted LLFDI raw scale). In our dataset, the mean was 37.89 and the median was 43.44. The mean fell in the "severe limitation" category and median fell in
  the "moderate to severe limitation" category.

TABLE 3: Predictive Power of KCCQ

KCCQ Score 45-100 (Not Frail) vs. KCCQ Score 0-44 (Frail)	
True positive (TP) = 12 **False positive (FP) = 27**	**Positive predictive value** = TP / (TP + FP) = 12 / (12 + 27) = 30.77%
True negative (TN) = 4 **False negative (FN) = 2**	**Negative predictive value** = TN / (FN + TN) = 4 / (2 + 4) =67.77%
Sensitivity **= TP / (TP + FN)** =12 / (12 + 2) = 85.7%	**Specificity** = TN / (FP + TN) = 4 / (27 + 4) = 12.9%

- Resp is a Boolean value to demonstrate complication or death status. 0 means patients without complication or death, and 1 means otherwise.

Figure 2 demonstrated some relationship of complications or deaths with each domain. For example, when Resp equals 1, most LPA data are in the range from 40 to 50; when Resp equals 0, most Gait data are in the range from 5 to 15. In addition, the relationship among five domains can be noticed. For instance, when Exha equals 3, most LPA data are in the range from 40 to 50, which means most patients with exhaustion have low physical activities.

Contingency tables were constructed to calculate negative and positive predictor values for frailty and the sensitivity and specificity of each measurement tool. The cohort was dichotomized into 2 groups "with complications or death within 30 days" versus "no complications or death within 30 days."

TABLE 2: Predictors with Each Domain of Fried Scale

Complic-ations or deaths in 0-30 days	**Gait**	**Exha**	**WeLoss**	**Grip**	**LPA**
Positive Predictors					
N= 14	12	12	6	11	11
Percentage	85.71%	85.71%	42.86%	78.57%	78.57%
Negative Predictors					
N= 31	28	20	13	21	28
Percentage	90.00%	64.00%	41.00%	67.00%	90.00%

After exploring each domain of the Fried scale (see Table 2), it is obvious that most of the 5 domains have strong correlations as both positive and negative predictors. Only "Weight Loss" is neutral in our experiment. "Gait" and "Low Physical Activity" both have very strong correlation affecting to "complications or deaths in 0-30 days." Two domains have similar accuracy with the KCCQ approaches (see Table 3). The results from Table 2 and Table 3 show that domains of the Fried scale have strong accuracy to predict frailty, even when we assume that each domain is independent.

Receiver operating characteristic (ROC) curves were computed to determine the predictive accuracy in identifying frailty, as those who truly had complications or death within 30 days. The ROC curves between five domains of the Fried scale and KCCQ are demonstrated in Figure 3.

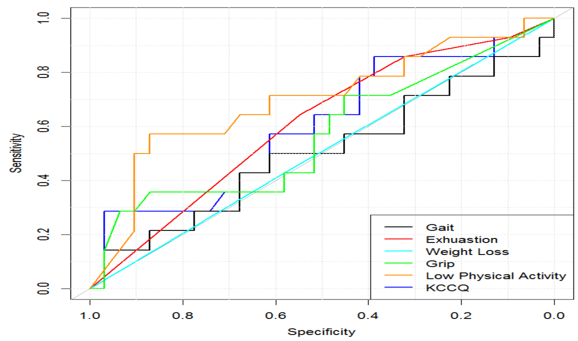

FIGURE 3: ROC Curves between Domains of Fried scale and KCCQ

Area under the ROC curve (AUC) was computed for each of Fried's frailty domains, the composite frailty phenotype (including the LLFDI scale), and the KCCQ and dichotomous cut points for predicting frailty were determined based on optimal crossing for sensitivity and specificity.

TABLE 4: AUC Values between Domains of Fried scale and KCCQ

Gait	Exha	WeLoss	Grip	LPA	KCCQ
0.5023	0.6094	0.5046	0.5703	0.6947	0.6025

Based on results of Figure 3 and Table 4, it is easy to conclude that "Low Physical Activity" is the most important and accurate domain of the Fried scale. Furthermore, KCCQ does not have strong accuracy and high AUC numbers, even though it is a popular approach.

4. Composed Domains of Fried Scale

However, those phenotypes are naturally derived from medical observations. There is no guarantee for them to be independent. Thus, it is rationale to compose domains for a more accurate predictor.

One simple and frequently-used approach is to compose domains based on the numbers of success. In this section, we study two different criteria from the approach and the results are demonstrated in Table 5.

- Criteria I: if any patient meets at least 3 domain criteria, then he/she will be considered as "Frail". Else, the patient is "Not Frail".

- Criteria II: patient has to meet at least 4 domain criteria to be considered as "Frail". Else, the patient is "Not Frail".

TABLE 5: Predictive Power of Composed Domains of Fried Scale

Met 0-2 Domain Criteria (Not Frail) vs Met 3-5 Domain Criteria (Frail)		
Total Population 45	Condition Positive 14	Condition Negative 31
Test Positive 40	True Positive 13	False Positive 27
Test Negative 5	False Negative 1	True Negative 4
Accuracy 38%	Sensitivity 92.86%	Specificity 12.9%
Met 0-1 Domain Criteria (Not Frail) vs Met 4-5 Domain Criteria (Frail)		
Total Population 45	Condition Positive 14	Condition Negative 31
Test Positive 28	True Positive 9	False Positive 19
Test Negative 17	False Negative 5	True Negative 12
Accuracy 46.67%	Sensitivity 64.29%	Specificity 38.71%

The comparison results in Table 5 show the limitation of such approach. Both criteria do not show high accuracy. Criteria I shows high sensitivity, which is common in medical study. However, the false positive number is too high. Criteria II's static numbers are neutral in most perspectives. It does not show the power of prediction of the Fried scale.

5. Classify Domains of Fried Scale using Three Data Mining Models

In this section, three different data mining models are applied to classify the domains of the Fried scale. The purpose is to find a better mathematical model to predict frail. A comparison with the KCCQ approach is also studied. All calculations are performed under R environments. Several R packages, such as General Linear Modeling, SVM and neural network model are applied in the research. And the major measurements to evaluate those three methods are specificity, sensitivity, and AUC.

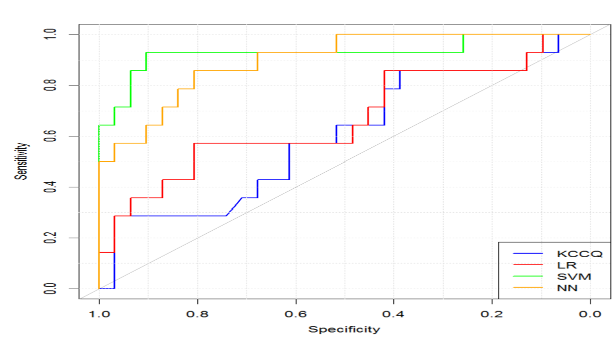

FIGURE 4: ROC Curves of KCCQ Classification System, Logistic Regression Model, SVM Model and Neural Network Model

Logistic regression analysis [30, 31] is a statistical analysis to determine the interdependent relationship between two or more quantitative variables, which is used widely. Multiple linear regression analysis models are used to predict a binary response from a binary predictor, used for predicting the outcome of a categorical dependent variable based on more predictor variables. It can measure the relationship between the dependent variable and independent variables and test how much weights each factor will take. During the calculation, either "Binomial" or "Gaussian" is used to describe the error distribution and link function. Logistic regression model with family 'Gaussian' shows better results (Specificity 81%, Sensitivity 58% and AUC 0.6682) than using family "Binomial" (Specificity 74%, Sensitivity 58% and AUC 0.6659). Thus, the ROC of logistic regression showed in Figure 4 and Table 6 is based on "Gaussian" family.

SVM [32, 33] is a supervised learning model with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. Support Vector learning is based on simple ideas which originated in statistical learning theory [29]. The simplicity comes from the fact that SVM apply a simple linear method to the data but in a high-dimensional feature space non-linearly related to the input space. Moreover, even though people can think of SVM as a linear algorithm in a high-dimensional space, in practice, it does not involve any computations in that high dimensional space. This simplicity combined with state of the art performance on many learning problems (classification, regression, and novelty detection) has contributed to the popularity of the SVM. During our calculation, epsilon regression (eps-Regression) is used as a regression machine. Two different kernels ("Radial Basis" and "Polynomial") are used in training and predicting. SVM model with type "eps-Regression" and kernel "Radial Basis" show better results (Specificity 91%, Sensitivity 92% and AUC 0.9286) than using kernel "Polynomial" (Specificity 87%, Sensitivity 87% and AUC 0.8917). Thus, the ROC of SVM showed in Figure 4 and Table 6 is based on type "eps-Regression" and kernel "Radial Basis".

Artificial neural network model [34, 35] is built on numerous nodes and the connection among these nodes. Each node represents a specific output function called activation function. Artificial neural networks are generally presented as systems of interconnected "neurons" which exchange messages between each other. The connections have numeric weights that can be tuned based on experience, making neural nets adaptive to inputs and capable of learning. During our calculation, two neural network models are applied: One Hidden Layer with One Hidden Neuron and One Hidden Layer with Two Hidden Neuron. Artificial neural network model with one hidden layer and two hidden neurons shows better results (Specificity 86%, Sensitivity 81% and AUC 0.8986) than using one hidden layer and one hidden neurons (Specificity 77%, Sensitivity 79% and AUC 0.8249). Thus, the ROC of neural network showed in Figure 4 and Table 6 is based on artificial neural network model with one hidden layer and two hidden neurons.

TABLE 6: Specificity, Sensitivity, AUC of KCCQ Classification System, Logistic Regression Model, SVM Model and Neural Network Model

	KCCQ Classification System	Logistic Regression Model	SVM Model	Neural Network Model
Specificity	62%	81%	91%	86%
Sensitivity	59%	58%	92%	81%
AUC	0.6025	0.6682	0.9286	0.8986

Some conclusions are drawn based on the comparison results from Figure 4 and Table 6.

- KCCQ classification system has lower specificity, sensitivity and AUC. It is only similar with simple approach, such as logistic regression.
- SVM has the higher performance in specificity, sensitivity and AUC. Neural network model also shows relatively high performance.

The study shows the feasibility to use statistical model to predict frail based on domains of the Fried scale.

6. Conclusions and Discussion

First off, patient characteristics of the study revealed that our demographics were a classic representation of what has been seen in the literature with TAVR patients [10, 16, 27]. Among our sample as well as patients undergoing TAVR in the U.S., both comprised similar demographics in gender make-up, mean age, and ethnicity [10, 16, 27]. Also similar in findings were the incidence of postoperative complications, both in our sample and the national TVT Registry data, examining 30-day incidence of stroke, vascular complications and overall mortality rate [16, 27]. Despite the data being obtained from a single center, the sample appears to well represent the general population of TAVR patients as a whole.

In this research paper, domains of the Fried scale were carefully explored. Based on our result, traditional KCCQ approach only show neutral accuracy. KCCQ and individual domains of the Fried scale have high sensitivity in detecting "truly frail." Among five domains, "Low Physical Activity," as measured by the LLFDI, appears to be the strongest single-frailty phenotype predictor. Exhaustion and grip strength also act as strong predictors of postoperative complications.

Furthermore, three different data mining models are applied to integrate domains of the Fried scale. SVM model shows the most impressive performance based on specificity, sensitivity and AUC. Artificial neural network model also shows higher performance than the KCCQ approach.

However, the study is still preliminary. This sample had numerous comorbidities preoperatively, and in terms of TAVR surgical outcomes, this certainly could have impacted the findings. These patients were part of the first year this TAVR program was in practice, hence, there is a learning curve for any TAVR team, and until the experience levels and skill sets of the doctors and surgeons improve over time, this can also negatively impact postoperative complications, as has been documented in research and now appearing in national guidelines [2, 27, 28].

In the future, researchers may want to conduct repeated measures of the Fried scale postoperatively, beyond the initial preoperative frailty data, to calculate hazard ratios in order to determine survival trends between those that had complications versus no complications. There was difficulty with a small dataset to generalize findings and for such a reason, we suspect the Artificial Neural Network model may provide more convincing results with a bigger dataset. Furthermore, with a small sample, we were unable to dichotomize groups between frail and not frail to explore further nuances.

7. References

[1] A. Alfirevic, AR. Mehta, LG. Svensson, "Trans catheter aortic valve replacement", *Anesthesiology Clin*, 2013, 31: 355-381.

[2] M. Puls, B. Sobisiak, A. Bleckmann, et al. "Impact of frailty on short- and long-term morbidity and mortality after transcatheter aortic valve implantation: risk assessment by Katz Index of activities of daily living", *EuroIntervention*, 2014, 10:609-19.

[3] J. Afilalo, K. Alexander, M. Mack et al., "Frailty assessment in the cardiovascular care of older adults", *J Am Coll Cardiol*, 2014, 63: 747-762.

[4] N. Bagnall, O. Faiz, A. Darzi, et al., "What is the utility of preoperative frailty assessment for risk stratification in cardiac surgery?", *Interact Cardio Vasc Thoarc Surg*, 2013, 14:398-402.

[5] NM. deVries, JB. Staal, van Ravensberg, et al. "Outcome instruments to measure frailty: A systematic review", *Ageing Res Rev.*, 2010, 10: 104-114.

[6] M. Mack. "Frailty and aortic valve disease", *J Cardiovasc Surg*, 2013, 145: S7-10.

[7] M. Makary, D. Segev, P. Pronovost, et al. "Frailty as a predictor of surgical outcomes in older patients", *J Am Coll Surg*, 2010, 210:901-908.

[8] O. Garin, M. Ferrer, A. Pont et al., "Disease-specific health-related quality of life questionnaires for heart failure: a systematic review with meta-analyses", *Qual Life Res*, 2009, 18:71–85.

[9] C. Green, C. Porter, D. Bresnaham, et al. "Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure", *J Am Coll Cardiol*, 2000, 35:1245-1255.

[10] M. Reynolds, E. Magnuson, K. Wang et al. "Health-related quality of life after transcatheter or surgical aortic valve replacement in high-risk patients with severe aortic stenosis: Results from the PARTNER (Placement of AoRTic TraNscathetER Valve) Trial (cohort A)", *J Am Coll Cardiol*, 2012, 60:548–58.

[11] J. Spertus, E. Peterson, MW. Conard, et al. "Monitoring clinical changes in patients with heart failure: a comparison of methods", *Am Heart J.*, 2005, 150:707–15.

[12] AW. Schoenenberger, S. Stortecky, S. Neumann, et al. "Predictors of functional decline in elderly patients undergoing transcatheter aortic valve implantation (TAVI)". *Eur Heart J.*, 2012, Epub ahead of print.

[13] S. Arnold, J. Spertus, Y. Lei, et al., "Use of the Kansas City cardiomyopathy questionnaire for monitoring health status in patients with aortic stenosis", *Cir Heart Fail*, 2013, 6:61-67.

[14] J. Afilalo, M. Eisenberg, JF. Morin, et al. "Gait speed as an incremental predictor of mortality and major morbidity in elderly patients undergoing cardiac surgery", *J Am Coll Cardiol*. 2010, 56(20):1668-76.

[15] P. Green. “Defining and Measuring Frailty: An Important Assessment Feature of High Risk TAVR Patients”, *J Am Coll Cardiol Intv*, 2012a, 5: 974.

[16] P. Green, AE. Woglom, P. Genereux, et al,. “The impact of frailty status on survival after transcatheter aortic valve replacement in older adults with severe aortic stenosis”, *J Am Coll Cardiol Intv*, 2012b, 5: 974-981.

[17] CM. Wilson, SRK. Kostsuca, and JA. Boura. “Utilization of a 5-Meter Walk Test in evaluating self-selected gait speed during preoperative screening of patients scheduled for cardiac surgery”, *Cardiopul Phys Thera Journ*, 2013, 24: 36-43.

[18] JF. Bean, DD. Olveczky, DK. Kiely, et al., “Performance-based versus patient-reported physical function: what are the underlying predictors?”, *Phys Ther.*, 2011, 91:1804-1811.

[19] E. Dineen, and B. Brent. “Aortic valve stenosis: comparison of patient with to those without chronic congestive heart failure”, *Am J Cardiol*, 1986, 57:419-422.

[20] LP. Fried, CM. Tangen, J. Walston, et al. “Frailty in older adults: evidence for a phenotype”, *J Gerontol*, 2000, 56:146-56.

[21] K. Bouillon, M. Kivimaki, M. Hamer, et al. “Measures of frailty in population based studies: an overview”, *BMC Geriatrics*, 2013, 13(64):2-11.

[22] SM. Haley, AM. Jette, WJ. Coster, et al. “Late Life Function and Disability Instrument: II. Development and evaluation of the function component”, *J Gerontol: Med Sci,* 2002, *57A*:M217-M222.

[23] TK. Kinney LaPier and R. Mizner. “Outcome measures in cardiopulmonary physical therapy: focus on the Late Life Function and Disability Instrument (LLFDI)”, *Cardiopulm Phys Ther J*, 2009, 20(2):32-35.

[24] TK. Kinney LaPier. “Utility of the late life function and disability instrument as an outcome measure in patients participating in outpatient cardiac rehabilitation: a preliminary study”, *Physiother Can*, 2012, 64(1):53-62.

[25] SP. Sayers, AM. Jette, SM. Haley, et al. “Validation of the Late-Life Function and Disability Instrument”, *J Am Geriatr Soc,* 2004, *52*:1554-1559.

[26] N. Dubuc, SM. Haley, P. Ni, et al. “Function and disability in late life: Comparision of the Late-Life Function and Disability Instrument to the Short-Form-36 and the London Handicap Scale”*, Disabil Rehabil 2004, 26*:362-370.

[27] DR. Holmes, MJ. Mack, S. Kaul, et al. “ACCF/ AATS/SCAI/STS expert consensus document on transcatheter aortic valve replacement”, *J Am Coll Cardiol*, 2012, 59(13):1200-54.

[28] RLJ. Osnabrugge, D. Mylotte, SJ. Head, 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*, 2013, 62(11): 1002-12.

[29] VN Vapnik, and V Vapnik. *Statistical Learning Theory*, Wiley New York, 1998.

[30] A. J. Dobson. *An Introduction to Generalized Linear Models*, London: Chapman and Hall, 1990.

[31] W. N. Venables, and B. D. Ripley. *Modern Applied Statistics with S*, New York: Springer, 2002.

[32] K. P. Bennett, and C. Campbell. “Support vector machines: Hype or hallelujah?”, *SIGKDD Explorations*, 2000, 2(2).

[33] C. Cortes, and V. Vapnik. “Support-vector network”, *Machine Learning*, 1995, 20, 1-25.

[34] Murata et al. “Network information criterion determining the number of hidden units for an artificial neural network model”. *IEEE Transactions on Neural Networks* 5 (6), 1994, pp 865-871.

[35] A. Anastasiadis et. al. “New globally convergent training scheme based on the resilient propagation algorithm”, *Neurocomputing* 64, 2005, pp 253-270.