Prediction of Poor Outcome after Transcatheter Aortic Valve Replacement

Suzanne V. Arnold, MD, MHA[a,b], **Jonathan Afilalo, MD, MSc**[c], **John A. Spertus, MD, MPH**[a,b], **Yuanyuan Tang, PhD**[a], **Suzanne J. Baron, MD, MSc**[a,b], **Philip G. Jones, MS**[a], **Michael J. Reardon, MD**[d], **Steven J. Yakubov, MD**[e], **David H. Adams, MD**[f], and **David J. Cohen, MD, MSc**[a,b] **on behalf of the US CoreValve Investigators**

[a]Saint Luke's Mid America Heart Institute, Kansas City, Missouri

[b]Department of Medicine, Saint Luke's Mid America Heart Institute, University of Missouri-Kansas City, Kansas City, Missouri

[c]Department of Medicine, Jewish General Hospital, McGill University, Montreal, Quebec, Canada

[d]Department of Cardiothoracic Surgery, Houston-Methodist-Debakey Heart and Vascular Center, Houston, Texas

[e]Riverside Methodist Hospital, Columbus, Ohio

[f]Department of Cardiothoracic Surgery, Mount Sinai Medical Center, New York, New York

Abstract

BACKGROUND—A series of models have been developed to identify patients at high risk for poor outcomes after transcatheter aortic valve replacement (TAVR) to help guide treatment choices, offer patients realistic expectations of long-term outcomes, and support decision making.

OBJECTIVES—We examined the performance of the previously developed TAVR Poor Outcome risk models in an external dataset and explored the incremental contribution of geriatric domains to model performance.

METHODS—Poor outcome after TAVR was defined as death, poor quality of life (QOL), or decline in QOL, as assessed using the Kansas City Cardiomyopathy Questionnaire. We tested 4 TAVR Poor Outcome risk models: 6-month and 1-year full and clinical (reduced) models. We examined each model's discrimination and calibration in the CoreValve trial dataset, then tested the incremental contribution of frailty and disability markers to the model's discrimination using the incremental discrimination index (IDI).

Corresponding Author: Suzanne V. Arnold, MD, MHA, Saint Luke's Mid America Heart Institute, 4401 Wornall Road, Kansas City, Missouri 64111, Telephone: 816-932-8606, Fax: 816-932-5613, suz.v.arnold@gmail.com.

RESULTS—Among 2,830 patients who underwent TAVR in the CoreValve US Pivotal Extreme and High Risk trials and associated continued access registries, 31.2% experienced a poor outcome at 6 months following TAVR (death: 17.6%; very poor QOL: 11.6%; QOL decline: 2.0%) and 50.8% experienced a poor outcome at 1 year (death: 30.2%; poor QOL: 19.6%; QOL: decline 1.0%). The models demonstrated similar discrimination as in the Placement of Aortic Transcatheter Valves Trial cohorts (c-indexes: 0.637 to 0.665) and excellent calibration. Adding frailty as a syndrome increased the c-indexes by 0.000 to 0.004 (IDI p < 0.01 for all except the 1-year clinical model), with the most important individual components being disability and unintentional weight loss.

CONCLUSIONS—Although discrimination of the TAVR Poor Outcome risk models was generally moderate, calibration was excellent among patients with different risk profiles and treated with a different TAVR device. These findings demonstrated the value of these models for individualizing outcome predictions in high-risk patients undergoing TAVR.

Keywords

Over the past decade, transcatheter aortic valve replacement (TAVR) has emerged as a less invasive option for aortic valve replacement and is currently approved in the United States for patients at high or extreme risk of morbidity or mortality with conventional surgical aortic valve replacement (SAVR). Several randomized trials have demonstrated that TAVR substantially reduces mortality and improves quality of life (QOL) as compared with medical therapy among patients considered to be inoperable (1), outcomes that are similar or superior to those of SAVR in high-risk patients (2,3). Nonetheless, some patients do not improve functionally or live long following TAVR, with approximately 1 in 4 patients treated with TAVR dying within 1 year (4). Moreover, despite substantial improvement in QOL for many patients, a sizeable minority continues to experience significant heart failure symptoms, with associated functional limitation and poor QOL (5–9).

Consequently, there is considerable interest in developing methods to identify patients at high risk of poor outcomes post-TAVR (10–13). A series of TAVR poor outcome risk models have been developed using data from the PARTNER (Placement of Aortic Transcatheter Valve) trials. Based on pre-procedural characteristics, these models were designed to identify such individuals (7) to help guide treatment choices and offer patients realistic expectations of outcomes based on their personal characteristics. Unlike previous studies that focused solely on mortality (14,15), these models integrated both mortality and health status into a single construct (16). This outcome is particularly relevant in this elderly population, as one of the principal benefits of treatment is to improve patient health status. Although these models had moderate discrimination and good calibration with the observed data, external validation could not be performed, nor were other potentially important predictors examined, such as frailty and disability, which have demonstrated prognostic importance (17).

To address these knowledge gaps, we used data from the CoreValve US Pivotal Extreme and High Risk trials and their associated continued access registries to examine the performance

of the prior models in a distinct and broader patient population. Additionally, we sought to determine the incremental contribution of integrating geriatric domains, such as gait speed and disability, in the risk models. Our ultimate goal was to develop models that could be deployed at the bedside.

Methods

The specific details of the CoreValve trials have been published previously (18,19). Briefly, patients had severe, symptomatic aortic stenosis and were estimated to be at high (estimated 30-day risk of mortality of 15%) or extreme (estimated 30-day risk of mortality or irreversible morbidity of 50%) surgical risk. Patients in the high-risk trial were randomized to TAVR versus SAVR; after the randomized trial was enrolled, subsequent patients underwent TAVR as part of a continued access registry. All patients in the extreme-risk cohort underwent TAVR. Our study included only patients who underwent an attempted TAVR procedure, performed using the self-expanding system (CoreValve, Medtronic, Inc., Minneapolis, Minnesota). The study was approved by the institutional review board at each site, and all patients provided written informed consent prior to participation.

Health Status, Functional Status, and Memory

Disease-specific health status was assessed at baseline, 1 month, 6 months, and 1 year after enrollment using the Kansas City Cardiomyopathy Questionnaire (KCCQ) (20). The overall KCCQ summary score (KCCQ-OS) was the primary health status outcome for this study. KCCQ-OS values ranged from 0 to 100; higher scores indicated fewer symptoms and better QOL. The KCCQ-OS score generally correlated with NYHA functional class as follows: class I: KCCQ-OS 75 to 100; class II: 60 to 74; class III: 45 to 59; and class IV: 0 to 44 (21,22). Changes in the KCCQ-OS of 5, 10, and 20 points correspond with small, moderate, or large clinical improvements, respectively (22).

Generic health status was assessed with the Medical Outcomes Study Short-Form 12 (SF-12) Health Survey (23). Functional status was assessed using a 6-minute walk test (6MWT), with patients permitted to use assist devices (e.g., walkers, canes), if needed (24). If a patient could not perform the test, the value for the 6MWT distance was set to 0 m. Memory was tested using the Mini-Mental Status Exam (MMSE) and scores were categorized as normal, mild dementia, or severe dementia (MMSE 28–30, 19–27, 0–18, respectively) (25).

Frailty and Disability

A syndrome of decreased physiological reserve, frailty is generally assessed through a range of geriatric domains including slowness, weakness, unintentional weight loss, exhaustion, and inactivity (26). Frailty, as a syndrome, was defined as 3 or more deficits in these 5 geriatric domains. Details about the definitions and measurements of each domain are in the Online Appendix. Disability was defined as 1 dependencies in activities of daily living (ADLs), which included: bathing, dressing, toileting, transferring, continence, and feeding (27). ADLs were considered as a continuous variable, with deficiencies ranging from 0 (fully independent in ADLs) to 6 (fully dependent).

Poor TAVR Outcomes and Risk Models

We modeled 2 levels of poor outcome after TAVR, as previously described. The first represented the worst outcome and was defined as any of the following: 1) death within the first 6 months after TAVR; 2) very poor QOL (KCCQ-OS score <45) at 6 months post-TAVR; or 3) moderate worsening in QOL (decrease of 10 points in the KCCQ-OS score from baseline to 6 months) (16,28). The second level of poor outcome represented a slightly better outcome and was defined as any of the following: 1) death within 1 year post-TAVR; 2) poor QOL (KCCQ-OS score <60) at 1 year after TAVR; or 3) moderate worsening in QOL (decrease of 10 points in the KCCQ-OS score from baseline to 1 year). These combined definitions integrated the 2 potential benefits of TAVR, reduced mortality and improved QOL, but also recognized that patients with good QOL at baseline may not improve symptomatically post-TAVR but could still derive a mortality benefit. It is expected that these outcomes would be used in combination to help patients understand their likelihood of meaningful recovery after TAVR.

The TAVR Poor Outcome risk models were constructed using data from patients who underwent TAVR in the PARTNER 1 trials and associated continued access registries, which included both inoperable and high-risk patients (7). Using a split-sample design, we developed and internally validated 2 multivariable logistic models to identify a parsimonious set of covariates to identify patients at high-risk for poor outcome. Each model considered 25 candidate variables and used Harrell's backward selection, with a 5% loss of information threshold for covariate selection (29). We also created a 1-year model (with a more conservative definition of poor outcome) using a similar approach. Finally, for ease of implementation in a clinical setting, we created 2 reduced ("clinical") models that included the 12-item KCCQ (30), excluded the 6MWT and Short Form-12 Mental summary score as potential predictors, categorized the MMSE (normal, mild dementia, and moderate/severe dementia) (25), and used a 10% loss of information as the cut-point for variable selection (to create more parsimonious models). In the original derivation and internal validation cohorts, each of the 4 models demonstrated moderate discrimination (c-index = 0.64 to 0.66) and good calibration.

Statistical Analysis

Baseline characteristics were compared between patients with and without a poor outcome at 6 months after TAVR using 2-sample Student t tests for continuous variables and chi-square tests for categorical variables. To examine the performance of the 4 models in the cohort of patients, we used the intercept and coefficient estimates from the prior prediction models to calculate the predicted risks of patients in the CoreValve trials. Discrimination was examined with c-indexes. We then plotted observed versus predicted risks by decile of predicted risk, and the regression line was compared against the line of equality (intercept = 0 and slope = 1). We then examined the incremental contribution of frailty and individual geriatric domains (including all 5 individual frailty domains and disability) to the performance of the models, assessed by independent association with the outcome and comparison of c-statistics using integrated discrimination improvement (IDI) statistics (details in the Online Appendix). All analyses were performed using SAS version 9.3 (SAS

Institute, Cary, North Carolina), and a p value of < 0.05 was used to indicate statistical significance.

Results

As part of the CoreValve US Pivotal Extreme Risk and High Risk trials and continued access registries, 3,290 patients underwent an attempted TAVR procedure and had 6 months of follow-up. After excluding 460 (14.0%) patients with missing KCCQ data (no patients were missing survival data), our final analytic cohort included 2,830 patients (Figure 1). Patients with missing outcomes data were generally similar to those in the analytic cohort, although patients with missing data were more likely to be female (52.6% vs. 45.4%; p = 0.004) and slightly more likely to have atrial fibrillation/flutter (49.6% vs. 43.8%; p = 0.021) and had lower MMSE scores (26.1 vs. 26.5; p = 0.004) (Online Table 1).

In the analytic cohort, the mean age was 83.3 years, 45.4% were female, mean aortic valve gradient was 47 mm Hg, and mean Society for Thoracic Surgeons mortality risk score was 9.0%. At 6 months after attempted TAVR, 882 patients (31.2%) had a poor outcome due to death in 498 (17.6%), very poor QOL in 328 (11.6%), and decline in QOL in 56 (2.0%) patients. The baseline characteristics of patients with an acceptable versus poor outcome at 6 months after attempted TAVR are shown in Table 1. Among the 2,325 patients with 1-year outcomes data available, 1,181 patients (50.8%) had a poor outcome: death in 703 (30.2%), poor QOL in 455 (19.6%), and decline in QOL in 23 (1.0%).

External Validation

When we examined the performance of the previously developed 6-month full model among patients in the trials, discrimination was moderate, with a c-index of 0.646, similar to the c-index of 0.661 in the PARTNER derivation cohort. The observed versus predicted risks of a poor outcome post-TAVR within risk deciles are shown in Figure 2. The median predicted risk of a poor outcome at 6 months after TAVR was 30.6%, with 9.4% of patients having a 50% predicted risk of a poor 6-month outcome. Among these high-risk patients, 40.8% were dead and an additional 16.4% had very poor QOL or QOL decline by 6 months after TAVR (Online Figure 1). The model demonstrated excellent calibration with the observed outcomes with an intercept of 0.03, a slope of 0.87, and an R^2 of 96%. Similar performance was observed for the 6-month clinical model (c-index = 0.637), the 1-year full model (c-index = 0.653), and the 1-year clinical model (c-index = 0.665); the calibration of all models was excellent (Figure 2). Based on the 1-year clinical model, the median predicted risk of a poor outcome at 1 year post-TAVR was 49.6%, with 8.4% of patients having a 70% predicted risk of a poor 1-year outcome. Among these very high-risk patients, 60.3% were dead and an additional 16.9% had poor QOL or QOL decline by 1 year after TAVR (Online Figure 1).

Incremental Contribution of Geriatric Domains to the Models

Frailty was identified in 59.8% of patients, with the most common deficits being slowness, inactivity, and exhaustion. Disability was reported in 16.6% of patients; the most common dependencies were bathing, transferring, and dressing. Except for weakness, all frailty and

disability markers were significantly more common in patients who had a poor outcome (Table 1). For all models except the 1-year clinical model, frailty was associated with an increase in the odds of a poor outcome of 30% to 40% when added to the existing models (Tables 2 and 3) and a significant, albeit small, improvement in discrimination. The increases in the c-indexes were 0.0043 (IDI p = 0.002), 0.0044 (IDI p = 0.001), 0.0036 (IDI p = 0.005), and 0 (IDI p = 0.218) for the 6-month full, 1-year full, 6-month clinical, and 1-year clinical models, respectively (Online Table 2).

When we examined individual components of frailty and disability, unintentional weight loss was associated with ~50% increased odds of poor outcome and each ADL dependency was associated with ~20% increased odds, with added incremental discriminatory ability to all 4 models (Tables 2 and 3). In the 2 full models, self-reported exhaustion was associated with ~35% increased odds of poor outcome, with added incremental discriminatory ability (Table 2), but it did not contribute significantly to the 2 clinical models that included the KCCQ (Table 3).

Discussion

Although TAVR is highly effective at relieving the hemodynamic obstruction of AS and can lead to excellent outcomes, some patients either die shortly after the procedure or do not achieve a functional benefit from the intervention. Using data from the PARTNER trials and continued access registry, we previously developed the TAVR Poor Outcome risk models that could help identify patients at high risk for poor outcomes after TAVR, based on baseline characteristics (7). In this study, these models performed similarly well in an external dataset of patients from the CoreValve trials and continued access registries, supporting their applicability for use in clinical practice. Furthermore, we were able to test the incremental contribution of both frailty and disability markers to the models, variables that were not collected routinely in the earlier PARTNER trials. As hypothesized, these variables added incrementally to the performance of the models, although the improvement in discrimination was small. Dependence in ADLs and unintentional weight loss were the most important predictors of a poor outcome of TAVR after accounting for the functional and clinical variables already included in the model.

Prediction models tend to perform better on the data from which they were constructed than on new data (31,32), making strategies such as bootstrapping for internal validation important to limit the optimism bias inherent in model creation (33). Even with such strategies, however, external validation remains a critical final step if such models are to demonstrate true clinical value (32). Nonetheless, in practice, prediction models are often accepted without external validation (34). In some cases, absence of external validation may relate to lack of an appropriate dataset for validation, insufficient cooperation among researchers with relevant datasets, or an aversion to "replication research." In the case of TAVR, however, we took advantage of a second major clinical trial program with both randomized and continued access components to validate and extend the results of the previous work. By testing the performance of the prior risk models in a completely separate dataset, we believe that the current study provides critical evidence to support the validity of these models for predicting poor outcome post-TAVR, thereby increasing our confidence in

their generalizability and ability to provide useful information to physicians evaluating patients for TAVR.

Frailty and disability have been associated with poor outcomes after TAVR, both for mortality (17,35,36) and the combination of death or poor QOL (6,37). As to their incremental value in risk prediction, frailty and disability improved model discrimination when added to cardiac surgery risk scores among patients undergoing coronary artery bypass and/or valve surgery (38). However, their role in predicting long-term survival and functional outcomes after TAVR has not been explored. While the optimal approach to measuring frailty and disability for preoperative risk assessment remains uncertain (39), practice guidelines recommend that physicians assess frailty when evaluating older adults with valvular heart disease (40). Interestingly, in this study, we found that the optimal clinical model did not require assessment of gait speed or grip strength; these factors, which may be cumbersome to collect, did not add incremental discrimination to existing models that already included a disease-specific assessment of functional status and QOL by means of the KCCQ (currently collected for all U.S. patients undergoing TAVR as part of the Transcatheter Valve Therapy registry).

Clinical Implications

The main goal of developing a prediction model is to inform clinical practice and, with TAVR, to counsel patients and their families as to the expected outcomes of care. Unfortunately, the number of risk models created and published far exceed those used in clinical practice. The reasons for this implementation gap are varied, but include use of variables that are difficult to collect in routine clinic practice, modeling outcomes that are not clinically relevant or actionable, and lack of external validation or concerns about model performance (41). While the first 3 criteria are not relevant to our models, there may be concern about the discriminative ability (as measured by the c-index) of the models, which was only moderate. Importantly, the models were exceedingly well calibrated, which may be even more important than discrimination when prospectively applying these models in clinical care (42). A model with moderate discrimination but excellent calibration means that the model is limited in its ability to reliably identify individuals who are certain not to benefit from TAVR (i.e., medical futility) but has a good ability to inform patients about their probability of a poor outcome (i.e., estimated risk). While some prediction models are highly discriminative – for instance, predicting short-term mortality after angioplasty (43) -- c-indexes of 0.62 to 0.66, as obtained for our models, are similar to those of other cardiovascular prediction models that have improved the application of clinical trials to practice (44,45).

Quality of life outcomes, while exceedingly important to patients, can be challenging to predict (46). Nonetheless, we believe that including QOL in our combined endpoint is a key strength of our study, better aligning our models with the goals of many patients considering TAVR. Notwithstanding these challenges and limitations, we believe our models can be useful in clinical practice. For example, we identified a modest-sized subgroup of patients with a very high risk of a poor outcome at 1 year post-TAVR, and among these patients, 58% were dead and an additional 18% had persistently poor QOL (Central Illustration). Thus,

while we cannot state with certainty which patients will or will not have a poor outcome, this assessment of risk could be quite valuable to physicians when deciding whether TAVR is appropriate for particular patients as well as for patients deciding whether to undergo TAVR. Given the relative ease of collection of the factors included in the models and the importance of risk stratification prior to TAVR in guiding treatment decisions and managing recovery expectations, we believe that these findings support using these models, along with the heart team's clinical judgment, in the preoperative evaluation of patients with AS.

Study Limitations

This study's results should be considered in light of several important limitations. Although the pre-existing models validated well in an external dataset collected at a different point in time and using a different TAVR device, the overall risk of poor outcome in the CoreValve trial cohort was similar to the derivation cohort. As such, the performance of these models in lower-risk patients remains unknown. Since model discrimination partly depends on the range of risk encompassed by the study population, it is likely that the c-statistic will increase significantly once lower-risk patients begin to be incorporated in the validation set. Until then, these models are likely to be most useful among high-risk and inoperable patients for whom a decision about pursuing TAVR is most relevant. Also, when interpreting our findings, it is important to recognize that this study was not designed to specifically examine the association of frailty and disability with poor outcomes but instead sought to test whether these factors improved the performance of the previous published prediction models. Indeed, the TAVR Poor Outcome risk models include many factors that are correlated with frailty, such as functional status (i.e., 6MWT or KCCQ), nutrition (BMI), exhaustion (KCCQ), and cognition (MMSE). Although beyond the scope of our study, it is likely that exclusion of these correlated factors from the analysis would have resulted in a stronger association between markers of frailty and disability and poor outcomes after TAVR.

Conclusions

Using data from a large, multicenter cohort of patients with severe symptomatic AS, we externally validated a series of previously published models for predicting poor outcome (combining both mortality and poor QOL) after TAVR. By demonstrating moderate discrimination and excellent model calibration, these analyses supported the validity of the previously derived models and demonstrated their potential applicability to clinical practice. Based on easily identified variables and limited survey data, the clinical models should be readily implementable in routine care to help guide challenging treatment decisions and manage patient (and physician) expectations. We also found that disability and selected frailty markers added incremental discriminatory ability to the models, suggesting they, too, should be considered as part of a comprehensive assessment prior to TAVR. Future research into the usefulness of these models in clinical care is needed to further support their value in the care of patients with severe aortic stenosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

ABBREVIATIONS AND ACRONYMS

6MWT   6-minute walk test

ADLs   activities of daily living

BMI   body mass index

IDI   integrated discrimination improvement

KCCQ-OS   Kansas City Cardiomyopathy Questionnaire-overall summary score

MMSE   Mini-Mental Status Exam

TAVR   transcatheter aortic valve replacement

References

1. Kapadia SR, Leon MB, Makkar RR, et al. 5-year outcomes of transcatheter aortic valve replacement compared with standard treatment for patients with inoperable aortic stenosis (PARTNER 1): a randomised controlled trial. Lancet. 2015; 385:2485–2491. [PubMed: 25788231]

2. Mack MJ, Leon MB, Smith CR, et al. 5-year outcomes of transcatheter aortic valve replacement or surgical aortic valve replacement for high surgical risk patients with aortic stenosis (PARTNER 1): a randomised controlled trial. Lancet. 2015; 385:2477–2484. [PubMed: 25788234]

3. Reardon MJ, Adams DH, Kleiman NS, et al. 2-Year Outcomes in Patients Undergoing Surgical or Self-Expanding Transcatheter Aortic Valve Replacement. J Am Coll Cardiol. 2015; 66:113–121. [PubMed: 26055947]

4. Holmes DR Jr, Brennan JM, Rumsfeld JS, et al. Clinical outcomes at 1 year following transcatheter aortic valve replacement. JAMA. 2015; 313:1019–1028. [PubMed: 25756438]

5. Arnold SV, Reynolds MR, Wang K, et al. Health Status After Transcatheter or Surgical Aortic Valve Replacement in Patients With Severe Aortic Stenosis at Increased Surgical Risk: Results From the CoreValve US Pivotal Trial. JACC Cardiovasc Interv. 2015; 8:1207–1217. [PubMed: 26292584]

6. Osnabrugge RL, Arnold SV, Reynolds MR, et al. Health status after transcatheter aortic valve replacement in patients at extreme surgical risk: results from the CoreValve U.S. trial. JACC Cardiovasc Interv. 2015; 8:315–323. [PubMed: 25700755]

7. Arnold SV, Reynolds MR, Lei Y, et al. Predictors of poor outcomes after transcatheter aortic valve replacement: results from the PARTNER (Placement of Aortic Transcatheter Valve) trial. Circulation. 2014; 129:2682–2690. [PubMed: 24958751]

8. Reynolds MR, Magnuson EA, Lei Y, et al. Health-related quality of life after transcatheter aortic valve replacement in inoperable patients with severe aortic stenosis. Circulation. 2011; 124:1964–1972. [PubMed: 21969017]

9. Reynolds MR, Magnuson EA, Wang K, 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; 60:548–558. [PubMed: 22818074]

10. Vahanian A, Alfieri O, Al-Attar N, et al. Transcatheter valve implantation for patients with aortic stenosis: a position statement from the European Association of Cardio-Thoracic Surgery (EACTS) and the European Society of Cardiology (ESC), in collaboration with the European Association of Percutaneous Cardiovascular Interventions (EAPCI). Eur Heart J. 2008; 29:1463–1470. [PubMed: 18474941]

11. Holmes DR Jr, Mack MJ, Kaul S, et al. 2012 ACCF/AATS/SCAI/STS expert consensus document on transcatheter aortic valve replacement. J Am Coll Cardiol. 2012; 59:1200–1254. [PubMed: 22300974]

12. Administration USFaD. , editor. U.S. Food and Drug Administration. [Accessed May 5, 2012] FDA Executive Summary: Edwards SAPIEN™ Transcatheter Heart Valve. 2011 Jul 20. Available at: www.fda.gov/downloads/…/UCM262930.pdf.

13. Centers for Medicare & Medicaid Services. [Accessed May 5, 2012] Decision Memo for Transcatheter Aortic Valve Replacement (TAVR) (CAG-00430N). Available at: https://www.cms.gov/medicare-coverage-database/details/nca-proposed-decision-memo.aspx?NCAId=257

14. Gilard M, Eltchaninoff H, Iung B, et al. Registry of transcatheter aortic-valve implantation in high-risk patients. N Engl J Med. 2012; 366:1705–1715. [PubMed: 22551129]

15. Durand E, Borz B, Godin M, et al. Performance analysis of EuroSCORE II compared to the original logistic EuroSCORE and STS scores for predicting 30-day mortality after transcatheter aortic valve replacement. Am J Cardiol. 2013; 111:891–897. [PubMed: 23337835]

16. Arnold SV, Spertus JA, Lei Y, et al. How to define a poor outcome after transcatheter aortic valve replacement: conceptual framework and empirical observations from the placement of aortic transcatheter valve (PARTNER) trial. Circ Cardiovasc Qual Outcomes. 2013; 6:591–597. [PubMed: 24021691]

17. Green P, Woglom AE, Genereux P, et al. The impact of frailty status on survival after transcatheter aortic valve replacement in older adults with severe aortic stenosis: a single-center experience. JACC Cardiovasc Interv. 2012; 5:974–981. [PubMed: 22995885]

18. Adams DH, Popma JJ, Reardon MJ, et al. Transcatheter aortic-valve replacement with a self-expanding prosthesis. N Engl J Med. 2014; 370:1790–1798. [PubMed: 24678937]

19. Popma JJ, Adams DH, Reardon MJ, et al. Transcatheter aortic valve replacement using a self-expanding bioprosthesis in patients with severe aortic stenosis at extreme risk for surgery. J Am Coll Cardiol. 2014; 63:1972–1981. [PubMed: 24657695]

20. Green CP, Porter CB, Bresnahan DR, Spertus JA. 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. [PubMed: 10758967]

21. Arnold SV, Spertus JA, Lei Y, et al. Use of the Kansas City Cardiomyopathy Questionnaire for Monitoring Health Status in Patients With Aortic Stenosis. Circ Heart Fail. 2013; 6:61–67. [PubMed: 23230306]

22. Spertus J, Peterson E, Conard MW, et al. Monitoring clinical changes in patients with heart failure: a comparison of methods. Am Heart J. 2005; 150:707–715. [PubMed: 16209970]

23. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996; 34:220–233. [PubMed: 8628042]

24. Afilalo J. Frailty in Patients with Cardiovascular Disease: Why, When, and How to Measure. Curr Cardiovasc Risk Rep. 2011; 5:467–472. [PubMed: 21949560]

25. Folstein MF, Robins LN, Helzer JE. The Mini-Mental State Examination. Arch Gen Psychiatry. 1983; 40:812. [PubMed: 6860082]

26. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001; 56:M146–M156. [PubMed: 11253156]

27. Shelkey M, Wallace M. Katz Index of Independence in Activities of Daily Living. Journal of gerontological nursing. 1999; 25:8–9.

28. Allen LA, Gheorghiade M, Reid KJ, et al. Identifying patients hospitalized with heart failure at risk for unfavorable future quality of life. Circ Cardiovasc Qual Outcomes. 2011; 4:389–398. [PubMed: 21693723]

29. Harrell, FE. Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag; 2001.

30. Jones PG, Gosch KL, Li Y, et al. The KCCQ-12: A Short Version of the Kansas City Cardiomyopathy Questionnaire. Circ Cardiovasc Qual Outcomes. 2013; 6:A248.

31. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol. 2003; 56:826–832. [PubMed: 14505766]

32. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol. 2016; 69:245–247. [PubMed: 25981519]

33. Steyerberg EW, Harrell FE Jr, Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 2001; 54:774–781. [PubMed: 11470385]

34. Siontis GC, Tzoulaki I, Castaldi PJ, Ioannidis JP. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. J Clin Epidemiol. 2015; 68:25–34. [PubMed: 25441703]

35. Rodes-Cabau J, Webb JG, Cheung A, et al. Long-term outcomes after transcatheter aortic valve implantation: insights on prognostic factors and valve durability from the canadian multicenter experience. J Am Coll Cardiol. 2012; 60:1864–1875. [PubMed: 23062535]

36. Stortecky S, Schoenenberger AW, Moser A, et al. Evaluation of multidimensional geriatric assessment as a predictor of mortality and cardiovascular events after transcatheter aortic valve implantation. JACC Cardiovasc Interv. 2012; 5:489–496. [PubMed: 22625186]

37. Green P, Arnold SV, Cohen DJ, et al. Relation of frailty to outcomes after transcatheter aortic valve replacement (from the PARTNER trial). Am J Cardiol. 2015; 116:264–269. [PubMed: 25963221]

38. Afilalo J, Mottillo S, Eisenberg MJ, et al. Addition of frailty and disability to cardiac surgery risk scores identifies elderly patients at high risk of mortality or major morbidity. Circ Cardiovasc Qual Outcomes. 2012; 5:222–228. [PubMed: 22396586]

39. Afilalo J, Alexander KP, Mack MJ, et al. Frailty assessment in the cardiovascular care of older adults. J Am Coll Cardiol. 2014; 63:747–762. [PubMed: 24291279]

40. Nishimura RA, Otto CM, Bonow RO, et al. 2014 AHA/ACC Guideline for the Management of Patients With Valvular Heart Disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014; 63:e57–e185. [PubMed: 24603191]

41. Arnold SV, Kosiborod M. Predicting the Benefit of Statins in Patients With Diabetes Mellitus: A Case of Perfect Being the Enemy of Good? Circ Cardiovasc Qual Outcomes. 2016; 9:191–193. [PubMed: 27166211]

42. Coppus SF, van der Veen F, Opmeer BC, et al. Evaluating prediction models in reproductive medicine. Hum Reprod. 2009; 24:1774–1778. [PubMed: 19395365]

43. Peterson ED, Dai D, DeLong ER, et al. Contemporary mortality risk prediction for percutaneous coronary intervention: results from 588,398 procedures in the National Cardiovascular Data Registry. J Am Coll Cardiol. 2010; 55:1923–1932. [PubMed: 20430263]

44. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. BMC Med Res Methodol. 2006; 6:18. [PubMed: 16613605]

45. Kent DM, Rothwell PM, Ioannidis JP, et al. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. Trials. 2010; 11:85. [PubMed: 20704705]

46. Arnold SV, Masoudi FA, Rumsfeld JS, et al. Derivation and validation of a risk standardization model for benchmarking hospital performance for health-related quality of life outcomes after acute myocardial infarction. Circulation. 2014; 129:313–320. [PubMed: 24163068]

PERSPECTIVES

COMPETENCY IN PATIENT CARE & PROCEDURAL SKILLS

Patients undergoing TAVR with poor functional status, poor QOL, low mean aortic valve gradient, lung disease, and kidney disease are at particularly high risk post-procedure for death or persistently low QOL. Dependence on others for assistance with ADLs and unintentional weight loss also correlated with poor outcomes after TAVR.

TRANSLATIONAL OUTLOOK

It is important to investigate the usefulness of these prediction models to guide selection of candidates for TAVR.
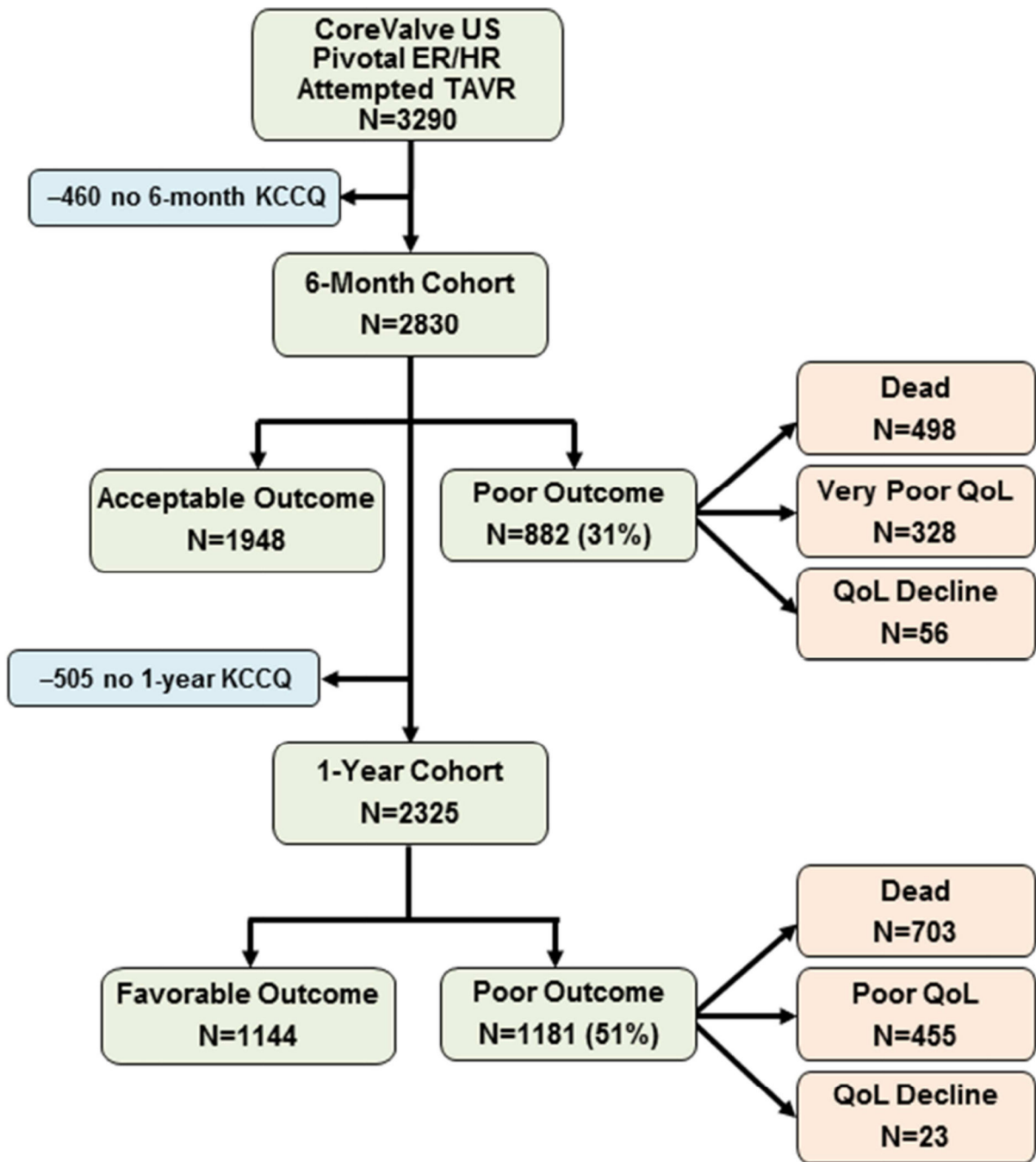
Figure 1. Study Flow
Patients from the CoreValve US Pivotal Extreme Risk and High Risk (ER/HR) trials were enrolled in this validation study. KCCQ = Kansas City Cardiomyopathy Questionnaire; QOL = quality of life; TAVR = transcatheter aortic valve replacement.
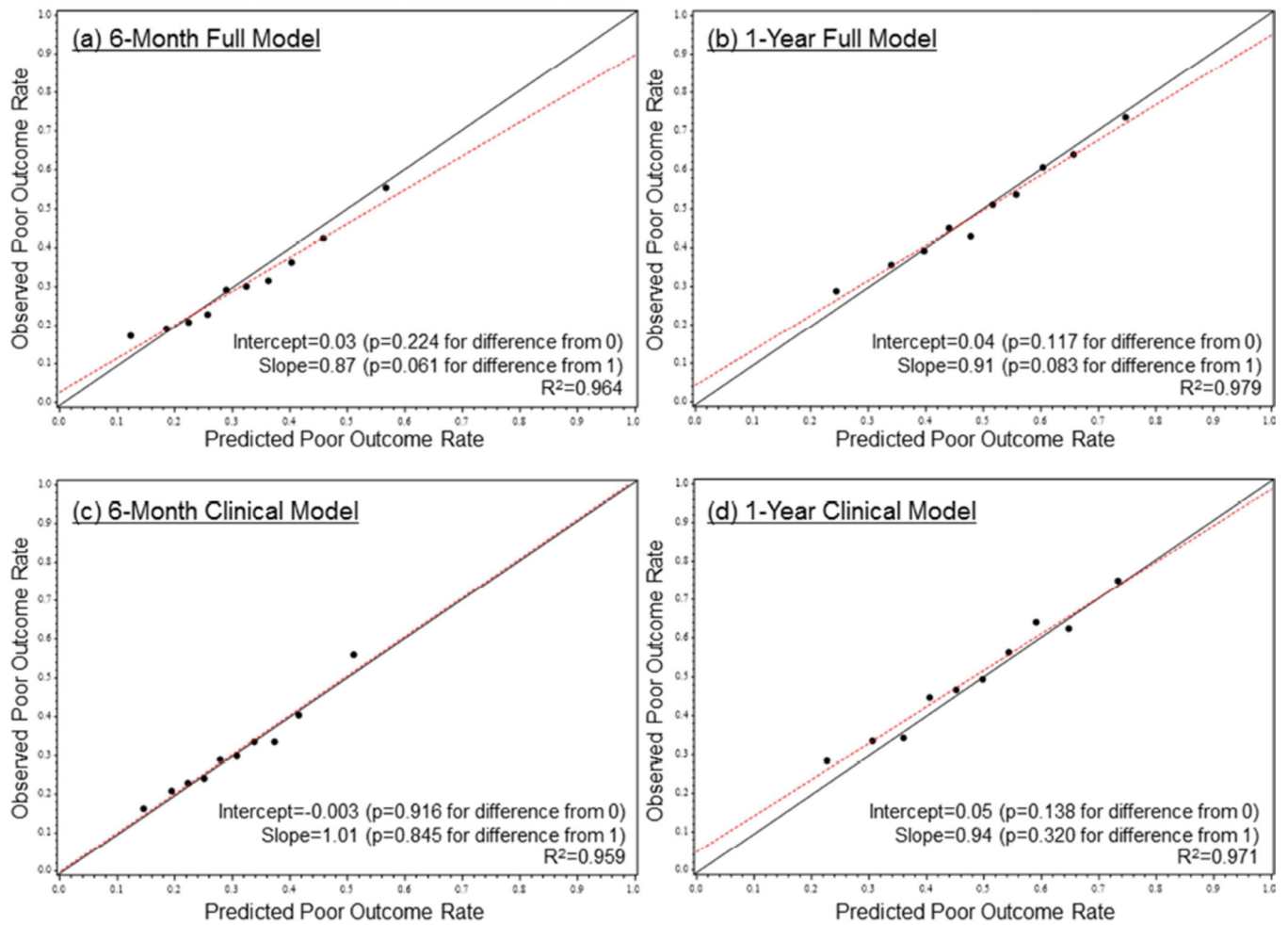
Figure 2. Risk Model Calibration

The calibration plots for the TAVR Poor Outcome risk models prediction are shown for the **(A)** 6-month full model, **(B)** 1-year full model, **(C)** 6-month clinical model, and **(D)** 1-year clinical model. Red lines = fitted regression lines; black lines = lines of equality (intercept = 0, slope = 1). Abbreviations as in Figure 1.
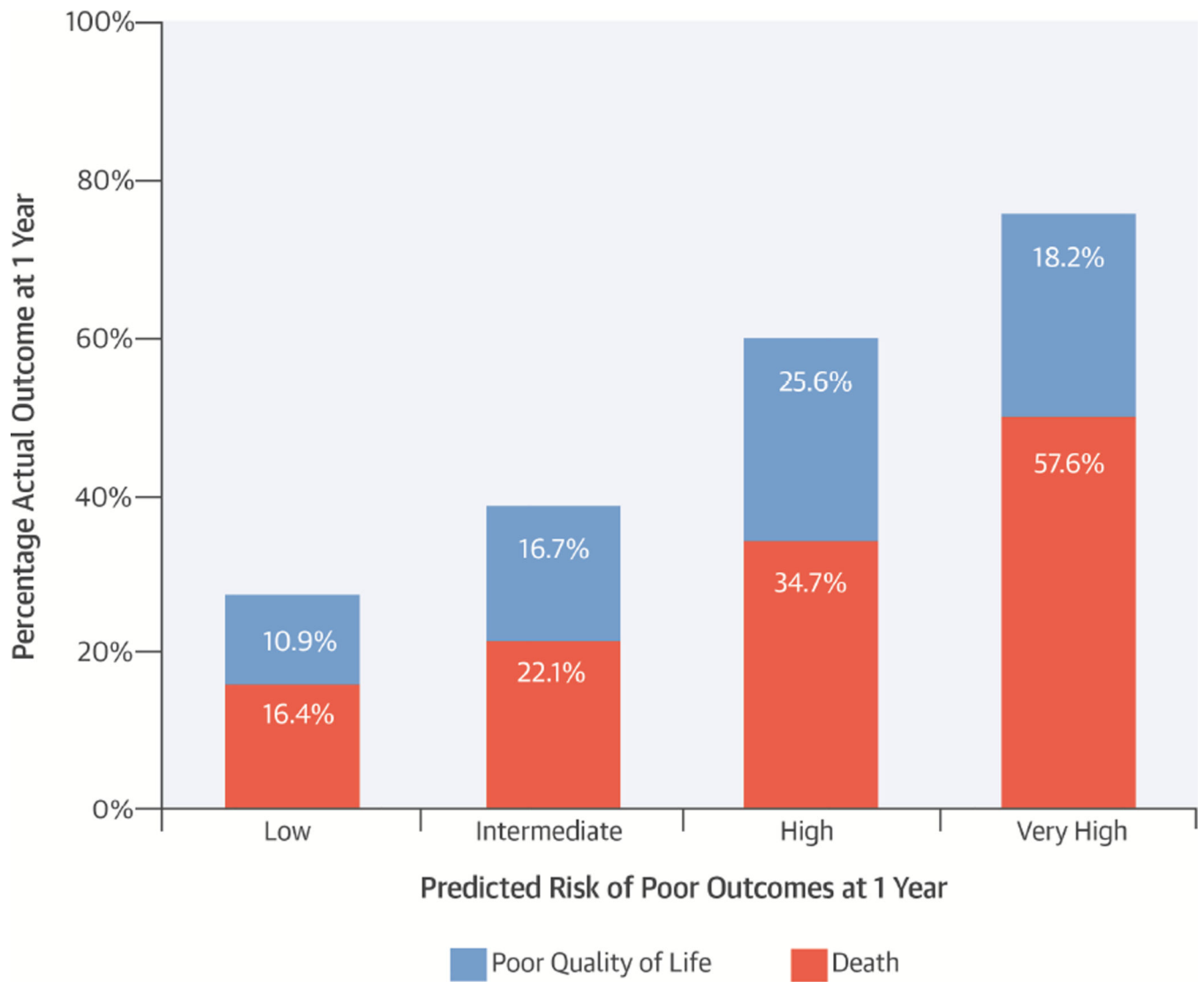
Central illustration. Poor Outcome after Transcatheter Aortic Valve Replacement
In this study to externally validate previously developed models, poor outcome after transcatheter aortic valve replacement (TAVR) was defined as death, poor quality of life (QOL), or decline in QOL. The models demonstrated excellent calibration between predicted levels of risk and 1-year outcomes post-TAVR. Inclusion of QOL better aligned these models with the goals of many patients contemplating TAVR.

Table 1

Baseline Characteristics[*]

	Poor Outcome n = 882	Acceptable Outcome n = 1,948	p Value
Demographic and clinical characteristics			
Age, yrs	83.6 ± 7.6	83.1 ± 7.9	0.135
Male	55.3	54.2	0.579
Coronary artery disease	79.5	79.9	0.782
Cerebrovascular disease	29.5	25.5	0.027
Peripheral vascular disease	47.4	45.9	0.459
Diabetes mellitus	38.4	35.8	0.183
Atrial fibrillation/flutter	49.7	41.2	<0.001
Home oxygen	33.6	18.1	<0.001
Creatinine, mg/dl	1.3 ± 0.6	1.2 ± 1.3	0.468
Hemoglobin, g/dl	11.6 ± 2.2	11.9 ± 2.1	<0.001
Mean arterial pressure, mm Hg	87.3 ± 12.7	88.6 ± 12.5	0.014
Disease severity			
Left ventricular ejection fraction, %	53.7 ± 14.2	54.3 ± 13.4	0.253
Mean aortic valve gradient, mm Hg	45.2 ± 13.2	48.3 ± 13.4	<0.001
STS mortality risk score, %	10.1 ± 5.2	8.5 ± 4.6	<0.001
6MWT distance, m	97.7 ± 108.0	134.6 ± 120.6	<0.001
KCCQ overall summary	38.2 ± 23.3	45.2 ± 22.9	<0.001
Geriatric domain measures			
Frailty syndrome, 3 deficits	67.2	56.5	<0.001
Walk speed, m/s	0.46 ± 0.32	0.57 ± 1.16	0.004
Slowness	82.1	73.0	<0.001
Grip strength	22.3 ± 10.4	23.5 ± 11.8	0.005
Weakness	42.2	38.7	0.074
Body mass index, kg/m^2	27.3 ± 6.5	28.0 ± 6.4	0.015
Unintentional weight loss	14.7	9.3	<0.001
Exhaustion	72.1	62.7	<0.001
Inactivity	80.4	72.2	<0.001
Disability (any)	24.7	13.0	<0.001
Needs assistance with bathing	20.6	9.5	<0.001
Needs assistance with dressing	15.1	6.2	<0.001
Needs assistance with toileting	8.7	3.0	<0.001
Needs assistance with transferring	13.6	7.1	<0.001
Incontinent	5.4	2.7	<0.001
Needs assistance with feeding	2.0	0.7	0.002

Values are mean ± SD or %.

[*] Patients with and without a poor outcome 6 months after transcatheter aortic valve replacement.

6MWT = 6-minute walk test; KCCQ = Kansas City Cardiomyopathy Questionnaire; STS = Society of Thoracic Surgeons.

TABLE 2

Full Models: Association of Frailty and Geriatric Domains with Poor Outcomes

	6-month Full Model		1-year Full Model	
	OR (95% CI)	p Value	OR (95% CI)	p Value
Base model [*]				
6MWT distance, per 10 m	0.97 (0.96–0.98)		0.97 (0.96–0.98)	
Mean aortic valve gradient, per 10 mm Hg	0.82 (0.75–0.89)		0.84 (0.77–0.90)	
Home oxygen	1.77 (1.23–2.54)		1.80 (1.25–2.61)	
Serum creatinine, per 1 mg/dl	1.32 (1.03–1.70)		1.41 (1.11–1.79)	
Mini-Mental Status Exam, per 1 point	0.96 (0.92–1.00)		0.94 (0.90–0.97)	
Atrial fibrillation/flutter	1.29 (1.02–1.63)		1.13 (0.91–1.40)	
Male	1.23 (0.96–1.57)		1.22 (0.97–1.53)	
Body mass index, per 1 kg/m^2	0.98 (0.96–1.00)		1.00 (0.98–1.02)	
Diabetes mellitus	0.82 (0.63–1.06)			
Mean arterial pressure, per 1 mm Hg	1.01 (1.00–1.02)			
	c-index = 0.646 [†]		c-index = 0.653 [†]	
Frailty syndrome	1.33 (1.11–1.59)	0.002	1.42 (1.18–1.69)	<0.001
	c-index = 0.651	IDI p = 0.002	c-index = 0.658	IDI p<0.001
Geriatric components				
Disabilities, per 1 ADL	1.25 (1.16–1.35)	<0.001	1.19 (1.09–1.30)	<0.001
Unintentional weight loss	1.52 (1.17–1.96)	0.001	1.61 (1.21–2.14)	0.001
Exhaustion	1.33 (1.10–1.60)	0.003	1.35 (1.12–1.63)	0.002
	c-index = 0.667	IDI p = 0.002	c-index = 0.666	IDI p = 0.021

[*] Odds ratios (OR) and confidence intervals (CI) for the base models are based on the original model derivations in PARTNER (7).

[†] Represents the discrimination of the existing base models in the CoreValve dataset.

ADL = activity of daily living; IDI = integrated discrimination improvement; other abbreviations as in Table 1.

Table 3

Clinical Models: Association of Frailty and Geriatric Domains with Poor Outcomes

	6-month Clinical Model		1-year Clinical Model	
	OR (95% CI)	p Value	OR (95% CI)	p Value
Base model [*]				
KCCQ-12-os	0.92 (0.87–0.96)		0.81 (9.78–0.85)	
Mean aortic valve gradient, per 10 mm Hg	0.81 (0.76–0.87)		0.86 (0.80–0.92)	
Home oxygen	1.62 (1.22–2.17)		2.54 (1.13–2.10)	
Serum creatinine, per 1 mg/dl	1.24 (1.02–1.51)		1.37 (1.12–1.68)	
Mild dementia/mild cognitive impairment [†]	1.31 (1.08–1.59)		1.30 (1.07–1.59)	
Moderate/severe dementia [†]	1.72 (0.82–3.62)		1.75 (0.78–3.97)	
Atrial fibrillation/flutter	1.40 (1.16–1.70)		1.28 (1.05–1.55)	
Diabetes mellitus	0.77 (0.63–0.94)		0.93 (0.76–1.14)	
	c-index = 0.637 [‡]		c-index = 0.665 [‡]	
Frailty syndrome	1.29 (1.08–1.55)	0.004	1.13 (0.93–1.36)	0.209
	c-index = 0.641	IDI p = 0.005	c-index = 0.665	IDI p = 0.218
Geriatric components				
Disabilities, per 1 ADL	1.29 (1.19–1.39)	<0.001	1.22 (1.12–1.33)	<0.001
Unintentional weight loss	1.51 (1.17–1.95)	0.001	1.53 (1.15–2.04)	0.003
	c-index = 0.656	IDI p = 0.007	c-index = 0.676	IDI p = 0.008

[*] Odds ratios and confidence intervals for the base model are based on the original model derivations in PARTNER (7).

[†] Mini-mental status exam score: 0–18 for moderate/severe dementia, 19–27 for mild dementia/mild cognitive impairment (25).

[‡] Represents the discrimination of the existing base models in the CoreValve dataset.

KCCQ-12-os = Kansas City Cardiomyopathy Questionnaire-12-overall summary score; other abbreviations as in Tables 1 and 2.