



Published in final edited form as:

JACC Cardiovasc Interv. 2021 October 11; 14(19): 2112–2123. doi:10.1016/j.jcin.2021.08.006.

Applicability of Transcatheter Aortic Valve Replacement Trials to Real-World Clinical Practice: Findings from EXTEND-CoreValve

Neel M. Butala, MD, MBA^{a,b}, Eric Secemsky, MD, MSc^a, Dhruv S. Kazi, MD, MS^a, Yang Song, MS^c, Jordan B. Strom, MD, MSc^a, Kamil F. Faridi, MD^d, J. Matthew Brennan, MD^e, Sammy Elmariah, MD, MPH^b, Changyu Shen, PhD^a, Robert W. Yeh, MD, MSc^a

^aRichard A. and Susan F. Smith Center for Outcomes Research in Cardiology, Division of Cardiovascular Medicine, Beth Israel Deaconess Medical Center, Boston MA

^bCardiology Division, Department of Medicine, Massachusetts General Hospital, Boston, MA

^cBaim Institute for Clinical Research, Boston, MA

^dSection of Cardiology, Department of Medicine, Yale School of Medicine, New Haven, CT

^eDuke Clinical Research Institute, Duke University School of Medicine, Durham, NC

Abstract

Objectives: We examined the applicability of pivotal transcatheter aortic valve replacement (TAVR) trials to the real-world population of Medicare patients receiving TAVR.

Background: It is unclear whether randomized controlled trial results of novel cardiovascular devices apply to patients encountered in clinical practice.

Methods: We compared characteristics of patients enrolled in the US CoreValve pivotal trials to the population of Medicare beneficiaries who received TAVR in US clinical practice between 11/2/2011 and 12/31/2017. We employed inverse-probability weighting to reweight the trial cohort based on Medicare patient characteristics and estimated a “real-world” treatment effect.

Results: A total of 2026 patients received TAVR in the US CoreValve pivotal trials and 135,112 patients received TAVR in the Medicare cohort. Trial patients were mostly similar to real-world patients at baseline, though trial patients were more likely to have hypertension (50% vs 39%) and coagulopathy (25% vs 17%), whereas real-world patients were more likely to have congestive heart failure (75% vs 68%) and frailty. The estimated real-world treatment effect of TAVR was a 11.4% absolute reduction in death or stroke (95% confidence interval [CI]: 7.50%, 14.92%) and an 8.7% absolute reduction in death (95% CI: 5.20%, 12.32%) at 1 year with TAVR compared to conventional therapy (surgical aortic valve replacement for intermediate/high risk patients and medical therapy for extreme risk patients).

Conclusions: Trial and real-world populations were mostly similar, though had some notable differences. Nevertheless, the extrapolated real-world treatment effect was at least as high as the

observed trial treatment effect, suggesting that the absolute benefit of TAVR in clinical trials is similar to the benefit of TAVR in the US real-world setting.

Tweet:

Reweighting of Corevalve trials based on Medicare patients suggests that the benefits of TAVR in clinical trials extend to patients receiving TAVR in the US real-world setting.

Condensed abstract

It is unclear whether randomized controlled trial results of novel cardiovascular devices apply to patients encountered in clinical practice. We examined the applicability of the US CoreValve pivotal trials to the Medicare population receiving TAVR. We used inverse-probability weighting to reweight the trial cohort based on Medicare patient characteristics and estimated “real-world” treatment effects. Trial and real-world populations were mostly similar, though had some notable differences. Nevertheless, the extrapolated real-world treatment effect was at least as high as the observed trial treatment effect, suggesting that the absolute benefit of TAVR in clinical trials is similar to the benefit of TAVR in the US real-world setting.

Keywords

generalizability; TAVR; real-world

Introduction

Randomized controlled trials (RCTs) are the gold standard to understand effectiveness of an intervention, but whether trial results based on highly selected populations apply to patients encountered in clinical practice is often unknown. Although there has been a recent move towards more pragmatic clinical trials to better reflect patient populations and treatment decisions in practice, novel technologies are often still evaluated in pivotal clinical trials with narrow inclusion criteria and few pragmatic elements in order to give a particular intervention the best chance of demonstrating an effect.(1,2) Differences between trial participants and the broader population may affect whether the safety and efficacy of a treatment observed in a clinical trial translate into clinical practice.(3) Understanding the generalizability of clinical trials is essential to understanding how their results extend to patients treated in the community.

Transcatheter aortic valve replacement (TAVR) has revolutionized the treatment of aortic valve disease over the past decade, driven by the results of large, high-quality RCTs.(4) These trials resulted in TAVR being incorporated into major societal guidelines as first-line therapy in patients with severe aortic stenosis who meet criteria,(5) and are touted as an example of rapid translation of bench concepts to clinical practice.(6) However, there have been concerns that these trials may not have adequately represented important patient subgroups, thus raising questions as to their generalizability.(7,8) Given the rapid growth of TAVR, estimating its treatment effect in a real-world setting is critically important to informing the evidence-based dissemination of this technology.

We aimed to answer the following questions: 1) is the US real-world TAVR population different from the population enrolled in pivotal TAVR trials; 2) how can we estimate an event rate in a more representative population using trial event rates; and 3) what would be the anticipated TAVR treatment effects if real-world patients were included in trials and had similar outcomes, conditional on their baseline characteristics.

Specifically, we applied inverse probability weighting (IPW) to reweight event rates in the trial population based on real-world population characteristics, which, under certain assumptions, allows the extrapolation of what trial results would have been had the trial been performed in the real-world population. Answering these questions not only provides important insights about how TAVR trial results compare to anticipated treatment effects encountered in real-world clinical practice, but also provides a useful example of a methodology for transporting clinical trial results to real-world population.

Methods

Study population

We included all patients aged ≥ 65 in the US CoreValve Extreme Risk, High Risk, and SURTAVI pivotal trials who could be successfully linked to the Centers for Medicare and Medicaid Services (CMS) Medicare Provider and Review (MedPAR) database. The MedPAR database includes a 100% sample of inpatient discharge claims for Medicare fee-for-service beneficiaries and has been used extensively for outcomes research.⁽⁹⁾ The High Risk trial and SURTAVI were RCTs comparing the self-expanding Medtronic™ CoreValve with surgical aortic valve replacement (SAVR) that enrolled patients in high and intermediate STS predicted risk of mortality (STS-PROM) categories, respectively, ^(10,11) whereas the Extreme Risk trial was a nonrandomized comparison of the CoreValve with an objective performance measure that enrolled patients in the extreme STS-PROM risk category (deemed to be at prohibitive risk for surgery).⁽¹²⁾ TAVR and SAVR patients in the High Risk trial and SURTAVI were included, whereas all TAVR patients were included from the Extreme Risk trial.

We linked the CoreValve Pivotal Trials dataset and CMS database as a part of the Extending Trial-Based Evaluations of Medical Therapies Using Novel Sources of Data (EXTEND) Study.⁽¹³⁾ Using deterministic linkage rules, 76% (2026/2660) of US patients in the trials were successfully linked to the MedPAR database (Supplemental Methods). Trial patients successfully linked to Medicare data were generally similar to those who were not linked, though linked patients had a higher rate of pre-existing congestive heart failure and were less likely to have a pre-existing pacemaker or implantable cardiac defibrillator (Supplemental Table 1).

We additionally identified a non-nested “real-world” cohort of US patients receiving TAVR within the CMS dataset using ICD-9-PCS and ICD-10-PCS claims codes. These patients were aged ≥ 65 , received TAVR between 11/2/2011 and 12/31/2017, and were not enrolled in any of the CoreValve Pivotal studies (Extreme Risk, High Risk, SURTAVI and pre-approval Continued Access Studies). Real-world SAVR patients were not included as a comparison since many of these patients would not have been eligible for TAVR during the study period

due to low surgical risk or other exclusions not identifiable through claims data. This study was approved by the institutional review board at Beth Israel Deaconess Medical Center.

Variables

The primary outcomes were incidence of stroke or death at 1 year and incidence of death at 1 year. Secondary outcomes included incidence of stroke, aortic valve reintervention, and pacemaker implantation at 1 year. Date of death was identified via the Medicare Master Beneficiary Summary File. Aortic valve reintervention, stroke, and permanent pacemaker implantation were identified using validated ICD-9 and ICD-10 claims codes (Supplemental Table 2).(14)

Covariates examined included demographics (age, sex, race), Elixhauser comorbidities,(15) and percentile rank according to previously validated claims-based frailty indicators.(16,17) Although some of these variables were collected in the trial data collection form for trial participants, for the purposes of this study, all variables were assessed from claims data for both trial and real-world populations to ensure similar ascertainment between groups.

Statistical analysis

We first combined the trials in proportion to recreate the United States real-world ratio of 26% extreme risk patients, 59% high risk patients, and 15% intermediate risk patients observed in the Society of Thoracic Surgeons(STS)/American College of Cardiology TVT registry during the study period, which was assumed to be distribution of such patients in our real-world cohort.(18,19) Specifically, Extreme Risk patients were weighted to represent 26% of the combined trial cohort, High Risk trial patients were weighted to represent 59% of the combined trial cohort, and SURTAVI patients were weighted to represent 15% of the combined trial cohort. As such, the distribution of patients across the extreme, high, and intermediate risk STS-PROM categories was similar across the trial and real-world cohorts by design. This proportionally combined sample was used to generate all subsequent ‘trial cohort’ estimates. Trials were combined given that surgical risk in the real-world population is a continuous distribution, and categorization into discrete risk categories may differ by center and over time, especially as the STS-PROM risk score calculation itself changes by year,(20,21) so a combined assessment provided the most reliable estimates.

We compared the demographics, comorbidities, and frailty scores between trial and real-world patients using standardized differences, as well as Student’s t tests for continuous variables and Fisher’s Exact tests for categorical variables.

We then combined the trial and real-world patients into a single cohort and created a propensity score model to predict the likelihood of being a trial patient (vs a real-world patient) using demographic, comorbidity, and frailty characteristics. We compared the distribution and degree of overlap of propensity scores between trial and real-world cohorts.

Next, in order to extrapolate trial findings to the broader population of patients who received TAVR, we estimated the incidence of outcomes that would have been observed in the trial among patients who received TAVR if trial participants had the same demographic and comorbidity distribution as the TAVR patients in the general community.

The idea underlying this methodology is to up-weight individuals in the trial cohort with characteristics more common in the real world and down-weight individuals in the trial cohort with characteristics less common in the real world (Central Illustration). Specifically, we employed an IPW method to reweight the cumulative incidence of events observed in the trial cohort who received TAVR, based on the distribution of characteristics in the real-world TAVR cohort. In order to extrapolate trial findings to the broader population of patients who received TAVR (i.e. for trial findings to be transportable), this method assumes that all members of the real-world cohort could have been included in the trials (positivity) and that real-world patients would have had similar outcomes had they been included in the trials, conditional on their baseline characteristics (conditional exchangeability) (22).

We then estimated the projected real-world treatment effect of TAVR vs. the comparator arm for death or stroke and for death alone at 1 year. We created a trial comparator arm cohort of SURTAVI and High Risk trial patients who received SAVR. We additionally included Extreme Risk trial patients who received TAVR in the trial comparator arm cohort, but assigned an incidence rate of stroke or death of 43% and an incidence of death of 42%. These rates were based on the methodology originally employed to compute the event rates used in the historical non-surgical comparator arm in the Extreme Risk trial, derived from the observed event rates for standard therapy in a similar randomized trial of TAVR in extreme risk patients.(23) We employed the same IPW method to estimate the cumulative incidence of outcomes that would have been observed for trial comparator arm patients if they had the same demographic and comorbidity distribution as observed for the real world TAVR patients. We calculated the estimated real-world treatment effect by subtracting the reweighted trial cumulative incidence of death or stroke for the trial comparator arm from the reweighted trial cumulative incidence of death or stroke for trial TAVR patients. Confidence intervals were calculated using bootstrapping with 500 iterations.

The valid reweighting of trial results to reflect real-world treatment effects requires that there are no unmeasured differences between the trial and real-world populations. We therefore tested whether the propensity score appropriately accounted for variables that could have affected the outcome in the real-world cohort by comparing the reweighted trial cumulative incidence of events based on real-world patient characteristics with the observed cumulative incidence of events in the real-world cohort using a log-rank test. Additionally, because the treatment effect combined across 3 trials was dependent on the historical event rate for non-surgical treatment for the Extreme Risk trial, we estimated a range of treatment effects using the upper and lower bounds of the 95% confidence interval for the historical event rates used in the Extreme Risk trial (death or stroke: [35.4%, 50.3%], death: [34.8%, 49.7%]) (12). Finally, we considered an alternative approach to extrapolate results of a trial to a broader population by multiplying the observed TAVR real-world event rates by the proportionally combined relative risks derived from the trials to calculate absolute risk differences. All analyses were conducted in SAS v 9.4 (SAS Institute, Cary, NC). We defined significance as a two-tailed $p < 0.05$.

Results

Similarity Between Trial and Real-World Populations

A total of 2,026 patients were included in the trial cohort and 135,112 patients were included in the real-world cohort. Based on the distribution of risk in the real-world TVT population at the time of the study, the 421 patients from the Extreme risk study were weighted to represent 26% of the combined trial cohort, the 600 patients from the high risk study were weighted to represent 59% of the combined trial cohort, and the 1005 patients from SURTAVI were weighted to represent 15% of the combined trial cohort (Supplemental Figure 1). Combined trial cohort patients were generally similar to real-world patients with regards to the majority of characteristics examined (Table 1, Figure 1). Nevertheless, trial patients were more likely to have certain important comorbidities such as hypertension (50% vs 39%), coagulopathy (25% vs 17%), fluid/electrolyte disorders (32% vs 20%), and weight loss (7.6% vs 4.2%), whereas real-world patients were more likely to have congestive heart failure (75% vs 68%), renal failure (37% vs 32.3%), metastatic cancer (0.6% vs 0.0%), solid tumor without metastasis (2.3% vs 0.8%), obesity (17% vs 13.3%), alcohol abuse (0.8% vs 0.1%), and a higher frailty index percentile (Table 1, Figure 1). Although there are some meaningful differences with respect to certain important comorbidities, the magnitude of the differences between these populations was generally modest, and the distribution of propensity scores between trial and real-world populations had a substantial degree of overlap (Figure 2), indicating similarity between these populations.

TAVR 1-Year Outcome Rates After Reweighting Trial Populations to Mimic Real-World TAVR Patients

To assess whether imbalances in the distribution of characteristics between the trial and real-world cohort would affect anticipated cumulative incidence of outcomes in the real-world population, the trial cohort was reweighted to represent the characteristics of the real-world population. The cumulative incidence of death or stroke at 1 year among patients undergoing TAVR in the trial cohort (20.9% [95% CI: 18.2%, 23.6%]) was similar to the incidence of death or stroke after reweighting the trial cohort based on characteristics of the real-world cohort (19.6% [95% CI: 16.6%, 22.7%]; Table 2 and Figure 3). Additionally, the cumulative incidence of death at 1 year among patients undergoing TAVR in the trial cohort (18.2% [95% CI: 15.6%, 20.7%]) was similar to the incidence of death after reweighting the trial cohort based on characteristics of the real-world cohort (17.0% [95% CI: 14.1%, 19.8%]). Cumulative incidence of stroke (5.0%), aortic valve reintervention (1.1%), and pacemaker (20.2%) at 1 year among patients undergoing TAVR in the trial cohort were similar to the incidence of stroke (4.8%), reintervention (1.1%), and pacemaker (19.9%) at 1 year after reweighting the trial cohort based on characteristics of the real-world cohort (Table 2, Figure 3, and Supplemental Figure 2).

TAVR Treatment Effect After Reweighting Trial Populations to Mimic Real-World Patients

We found that the estimated real-world TAVR treatment effect was similar to or greater than the trial TAVR treatment effect (Table 3). For the outcome of death or stroke at 1 year, the estimated real-world treatment effect of TAVR compared to conventional therapy was a 11.37% (95% confidence interval: 7.50%, 14.92%) absolute reduction, whereas the

trial treatment effect was an 8.38% absolute reduction (95% confidence interval: 4.62%, 11.91%, difference in risk differences 2.99%, 95% CI: 0.82%, 5.53%). For the outcome of death at 1 year, the estimated real-world treatment effect of TAVR compared to conventional therapy was an 8.74% absolute reduction (95% CI: 5.20%, 12.32%), whereas the trial treatment effect was a 6.95% absolute reduction (95% CI: 3.73%, 10.31%; difference in risk differences 1.79%, 95% CI: -0.43%, 4.12%).

Supplemental Analysis

Rewighted trial estimates of death or stroke, death, stroke, and aortic reintervention, among patients receiving TAVR based on characteristics of the real-world cohort were similar to the observed rates of these outcomes in the real-world cohort. Although rates of pacemaker placement were different (reweighted trial 19.9% [95% CI: 16.8%, 23.0%] vs. observed real world 14.1% [95% CI: 13.9%, 14.2%]); Supplemental Table 3), presence of pre-existing pacemaker, which is an important determinant of this particular outcome, was not included in the propensity score model due to inconsistent capture in claims data. Overall, this suggests that there were minimal unobserved differences in baseline characteristics between the trial and real-world cohorts affecting the majority of trial outcomes that were not accounted for by the covariates included in the propensity score model.

The estimated real-world treatment effect for reducing death and stroke was similar to the trial treatment effect when using the lower bound historical extreme risk non-surgical treatment event rate (difference between trial and estimated real-world treatment effect 0.34%, 95% CI: -1.81%, 2.69%), but was greater than the trial treatment effect when using the upper bound historical extreme risk non-surgical treatment event rate (difference between trial and estimated real-world treatment effect 3.29%, 95% CI: -0.97%, 5.75%; Supplemental Table 4).

The estimated real-world treatment effects derived from an alternative relative risks-based approach were an 8.59% reduction in death and an 10.16% reduction in death or stroke with TAVR compared to conventional therapy, both which were higher than trial treatment effects, but lower than estimated real-world treatment effects using the reweighting approach.

Discussion

In this study examining TAVR trials and a Medicare population of patients undergoing TAVR in routine clinical practice, we found that the trial and real-world populations were generally similar with respect to baseline characteristics, though had some meaningful differences with respect to certain important comorbidities. Nevertheless, trial TAVR outcome rates and treatment effects were similar to reweighted trial TAVR outcome rates and estimated treatment effects, supporting the notion that the treatment effects in the original TAVR pivotal trials reflect treatment effects anticipated in the real-world population of patients receiving TAVR in the United States. These results not only have implications for TAVR trial applicability, but also for estimating trial effects in real-world populations more broadly.

Comparison with previous studies

Prior analyses in real-world populations have questioned whether the results of TAVR clinical trials extend to the broader population of patients receiving TAVR.(7,8) Although one registry-based study found similar results to clinical trials,(24) 30% of the real-world TAVR population was still excluded from this analysis given that they had characteristics thought to strongly favor receipt of TAVR or SAVR. In our study of Medicare patients receiving TAVR, we found that the estimated real-world treatment effect for TAVR compared to conventional therapy was greater than the treatment effect observed across all trials in reducing death or stroke and similar to the treatment effect observed across all trials in reducing death alone. This extrapolation assumes that the Medicare TAVR population could have been included in TAVR trials and would have had similar outcomes, conditional on their baseline characteristics. Given that we found similar trial death and stroke rates for TAVR before and after reweighting, a greater estimated real-world treatment effect for death or stroke was driven by the selection of real-world patients for TAVR who would have otherwise had a higher incidence of events had they undergone SAVR or medical therapy. Taken together, these data support the applicability of TAVR trial estimates and overall treatment effects of TAVR to the real-world Medicare population of patients receiving TAVR.

Implications

This study has implications for the use of real-world data to assess the real-world impact of new cardiovascular therapies. New medical devices, such as TAVR, are tested in pivotal clinical trials to demonstrate efficacy and safety, as well as meet rigorous standards for regulatory approval, but there is often concern that such trials can be poor predictors of effectiveness in a real-world setting.(1,2) The 21st Century Cures Act explicitly encourages the use of real-world data to support regulatory decision making, including approval of new indications and post-approval evaluation.(25) However, there is concern that biases inherent in observational studies of real-world data may lead to misleading conclusions.(26) We use real-world data in conjunction with randomized trial data to calculate an estimate of population-level effectiveness while maintaining the benefits of randomization. Instead of simply comparing estimates between trial and real-world populations, we reweighted the trial population to represent the real-world population based on observable characteristics, which allowed the generation of trial estimates more applicable to the real-world population while maintaining the benefits of randomization in the trial.

This study also has implications for methodologies for transporting the results of other cardiovascular trials to real-world populations. Many studies have compared characteristics of trial and real-world populations or assessed the proportion of real-world patients meeting trial inclusion criteria.(27–30) However, these methods do not provide insight into what the estimated treatment effect is expected to be in the real-world population. In this study, the application of IPW methods offered an opportunity to extrapolate trial results to real-world populations,(22,31–35) a method which has had only limited prior application in large clinical trials.(22,36,37)

While these methods can be applied more broadly, they may not be as suitable (i.e. trial results may not be as transportable) in certain contexts. First, the estimated real-world treatment effect is based on those that actually received treatment in real-world practice, and thus does not account for patients who may have been eligible for an intervention, but for whom an intervention was not pursued. For instance, our real-world population of TAVR patients has notably few black individuals, which is consistent with other studies.⁽³⁸⁾ Nevertheless, given that TAVR outcomes are generally similar across racial groups,⁽³⁸⁾ such underrepresentation may be unlikely to influence the real-world treatment effect in our study. Second, these methods cannot estimate a real-world treatment effect for patients that would have been completely excluded from a trial, as such patients would not have corresponding trial patients with similar characteristics that could be upweighted. While the contribution of such patients is likely to be small for our study given the relatively broad inclusion criteria across the multiple TAVR trials in our sample and the stringent regulation of TAVR in the real-world during the study period, the impact of real-world patients that would have been completely excluded from trials may be larger in applying these methods to other contexts. Thus, although these methods can help bridge the gap between clinical trials and real-world practice, clinical trials should still strive to include broadly representative populations in order to provide generalizable results. Third, the real-world population for which trial results are extrapolated must be well-represented by the real-world dataset used. In the case of TAVR, the vast majority of the U.S. real-world TAVR patients receive insurance through Medicare. Attempts to replicate this approach for other technologies using Medicare data may provide insight into only a subset of the intended target population.

Limitations

The findings of this study must be interpreted in context of its limitations. First, our results represent applicability of TAVR trials to the population of Medicare patients at a time when TAVR was only approved to be performed in extreme, high, and intermediate risk individuals. Second, it is possible that the distribution of STS-PROM risk in our real-world cohort is different from the distribution of STS-PROM risk in the TVT registry during the study period. However, this is highly unlikely given that the majority of patients receiving TAVR during this time period would have been eligible for Medicare and the distribution of STS-PROM risk was similar between linked and unlinked patients in our trial cohort. Third, given that real-world operators may be distinct from trial operators and that we could not ascertain certain important variables such as valve type, prior pacemaker, or aortic valve anatomy (i.e. bicuspid vs tricuspid) from Medicare claims data, there may be residual confounding that is not accounted for in our propensity score model, despite our use of several validated comorbidity measures. However, we found that reweighted trial event rates based on real-world population characteristics were similar to observed event rates in the real-world population for most outcomes, suggesting that the impact of unobserved patient and procedural differences may be minimal. Fourth, it is possible that administrative coding practices may differ between trial and real-world patients which could affect the distributions of baseline characteristics in our sample, though this may be unlikely given that administrative coding generally operates in a parallel workstream to clinical trial data collection. Finally, given that we do not perform separate reweighting analyses for the extrapolated benefit of TAVR against surgery (for the high and intermediate risk

groups) and against medical therapy (for the extreme risk group), the combined real-world treatment effect calculated in this study may not apply directly to a particular patient's clinical decision-making. However, these results are still useful in assessing whether the TAVR trial results in aggregate extrapolate to the real-world setting and can inform similar assessment of other novel technologies.

Conclusions

Our study is the first to estimate the real-world treatment effect of TAVR trials. Although trial and real-world populations were similar with respect to many baseline characteristics, there were key differences in certain some important baseline characteristics. Nevertheless, event rates and treatment effects in trials were similar whether or not trial populations were reweighted to better reflect real-world populations. The methods used in this study can be used to estimate trial effects in real-world populations to evaluate how novel therapies are introduced into clinical practice more broadly.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author funding/disclosures

The study was funded by NHLBI 1R01HL136708 (Yeh). Dr. Butala is funded by the John S. LaDue Memorial Fellowship at Harvard Medical School, Boston, MA and reports consulting fees and ownership interest in HiLabs, outside the submitted work. Dr. Secemsky receives grants from AstraZeneca, BD Bard, Boston Scientific, Cook Medical, CSI, Medtronic, Philips, and UCSF. He consults for CSI, Medtronic, and Philips and is on the speaking bureau of BD Bard, Cook Medical and Medtronic. Dr. Strom is funded by a grant from the NIH/NHLBI (1K23HL144907). Dr. Brennan holds an Innovation in Regulatory Science Award from Burroughs Wellcome Fund (1014158), a Food and Drug Administration grant (1U01FD004591-01), and consulting for Edwards Lifesciences and Atlicure. Dr. Elmariah is funded by grants from the American Heart Association (19TPA34910170) and the US Department of Defense (W81XWH1810080). He also receives grant support from Edwards Lifesciences and Svelte Medical and consulting fees from AstraZeneca and Medtronic, outside the submitted work. Dr. Yeh reports additional grant support from Abiomed, Astra Zeneca and Boston Scientific and, and consulting fees from Abbott, Boston Scientific, Medtronic, and Teleflex, outside the submitted work. Dr. Strom, Dr. Faridi, Dr. Kazi, Mr. Song, and Dr. Chen have no relationships with industry.

Abbreviations

RCT	randomized controlled trial
TAVR	transcatheter aortic valve replacement
CMS	Centers for Medicare and Medicaid Services
MedPAR	Medicare Provider and Review
SAVR	surgical aortic valve replacement
EXTEND	Evaluations of Medical Therapies Using Novel Sources of Data
IPW	inverse probability weighting

References

1. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Clin Epidemiol* 2009;62:499–505. [PubMed: 19348976]
2. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials* 2009;10:37. [PubMed: 19493350]
3. Fonarow GC. Randomization—There Is No SubstituteRandomization—There Is No SubstituteEditorial. *JAMA Cardiology* 2016;1:633–635. [PubMed: 27439153]
4. Otto CM. Informed Shared Decisions for Patients with Aortic Stenosis. *N Engl J Med* 2019;380:1769–1770. [PubMed: 31042831]
5. Otto CM, Nishimura RA, Bonow RO et al. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J Am Coll Cardiol* 2021;77:450–500. [PubMed: 33342587]
6. Cribier A The development of transcatheter aortic valve replacement (TAVR). *Global cardiology science & practice* 2016;2016.
7. Jilaihawi H, Chakravarty T, Weiss RE, Fontana GP, Forrester J, Makkar RR. Meta-analysis of complications in aortic valve replacement: Comparison of Medtronic-Corevalve, Edwards-Sapien and surgical aortic valve replacement in 8,536 patients. *Catheterization and Cardiovascular Interventions* 2012;80:128–138. [PubMed: 22415849]
8. Mohr FW, Board ftGE, Holzhey D et al. The German Aortic Valve Registry: 1-year results from 13 680 patients with aortic valve disease†. *European Journal of Cardio-Thoracic Surgery* 2014;46:808–816. [PubMed: 25079769]
9. Centers for Medicare and Medicaid Services. Medicare provider analysis and review (MEDPAR) file. Available at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/MEDPARLDSHospitalNational>. Accessed 7/1/2019.
10. Reardon MJ, Van Mieghem NM, Popma JJ et al. Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk Patients. *N Engl J Med* 2017;376:1321–1331. [PubMed: 28304219]
11. Adams DH, Popma JJ, Reardon MJ et al. Transcatheter aortic-valve replacement with a self-expanding prosthesis. *N Engl J Med* 2014;370:1790–8. [PubMed: 24678937]
12. 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;63:1972–81. [PubMed: 24657695]
13. Strom JB, Tamez-Aguilar H, Zhao MY et al. Validating the Use of Registries and Claims Data to Support Randomized Trials: Rationale and Design of the Extending Trial-Based Evaluations of Medical Therapies Using Novel Sources of Data (EXTEND) Study. *American Heart Journal* 2019.
14. Strom JB, Zhao Y, Faridi KF et al. Comparison of Clinical Trials and Administrative Claims to Identify Stroke Among Patients Undergoing Aortic Valve Replacement: Findings From the EXTEND Study. *Circulation Cardiovascular interventions* 2019;12:e008231. [PubMed: 31694411]
15. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Medical care* 1998;8–27. [PubMed: 9431328]
16. 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* 2017;55:716–722. [PubMed: 28437320]
17. Gilbert T, Neuburger J, Kraindler J et al. Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study. *Lancet (London, England)* 2018;391:1775–1782.
18. Holmes DR, Nishimura RA, Grover FL et al. Annual Outcomes With Transcatheter Valve Therapy: From the STS/ACC TVT Registry. *Ann Thorac Surg* 2016;101:789–800. [PubMed: 27175453]
19. Bavaria JE. A View from the STS/ACC TVT Registry Steering Committee. *Cardiovascular Research Technologies Conference* 2018. 2018. Accessible at: <https://www.crtonline.org/presentation-detail/view-from-sts-acc-tvt-steering-committee-2>. Accessed 10/31/2019.

20. Rogers T, Koifman E, Patel N et al. Society of Thoracic Surgeons Score Variance Results in Risk Reclassification of Patients Undergoing Transcatheter Aortic Valve Replacement. *JAMA Cardiol* 2017;2:455–456. [PubMed: 27851845]
21. Kumar A, Sato K, Narayanswami J et al. Current Society of Thoracic Surgeons Model Reclassifies Mortality Risk in Patients Undergoing Transcatheter Aortic Valve Replacement. *Circulation Cardiovascular interventions* 2018;11:e006664. [PubMed: 30354591]
22. Dahabreh IJ, Robertson SE, Tchetgen EJ, Stuart EA, Hernan MA. Generalizing causal inferences from individuals in randomized trials to all trial-eligible individuals. *Biometrics* 2019;75:685–694. [PubMed: 30488513]
23. Leon MB, Smith CR, Mack M et al. Transcatheter Aortic-Valve Implantation for Aortic Stenosis in Patients Who Cannot Undergo Surgery. *New England Journal of Medicine* 2010;363:1597–1607.
24. Brennan JM, Thomas L, Cohen DJ et al. Transcatheter Versus Surgical Aortic Valve Replacement: Propensity-Matched Comparison. *J Am Coll Cardiol* 2017;70:439–450. [PubMed: 28728688]
25. 21st Century Cures Act, H.R. 34, 114th Cong. (2015).
26. Collins R, Bowman L, Landray M, Peto R. The Magic of Randomization versus the Myth of Real-World Evidence. *N Engl J Med* 2020;382:674–678. [PubMed: 32053307]
27. Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* 2015;16:495. [PubMed: 26530985]
28. Sen A, Goldstein A, Chakrabarti S et al. The representativeness of eligible patients in type 2 diabetes trials: a case study using GIST 2.0. *Journal of the American Medical Informatics Association* 2017;25:239–247.
29. Sen A, Chakrabarti S, Goldstein A, Wang S, Ryan PB, Weng C. GIST 2.0: A scalable multi-trait metric for quantifying population representativeness of individual clinical studies. *Journal of biomedical informatics* 2016;63:325–336. [PubMed: 27600407]
30. Seprehvand N, Alemayehu W, Das D et al. Trends in the Explanatory or Pragmatic Nature of Cardiovascular Clinical Trials Over 2 Decades. *JAMA Cardiology* 2019.
31. Nguyen TQ, Ackerman B, Schmid I, Cole SR, Stuart EA. Sensitivity analyses for effect modifiers not observed in the target population when generalizing treatment effects from a randomized controlled trial: Assumptions, models, effect scales, data scenarios, and implementation details. *PloS one* 2018;13:e0208795. [PubMed: 30533053]
32. Buchanan AL, Hudgens MG, Cole SR et al. Generalizing Evidence from Randomized Trials using Inverse Probability of Sampling Weights. *Journal of the Royal Statistical Society Series A, (Statistics in Society)* 2018;181:1193–1209.
33. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. *American journal of epidemiology* 2010;172:107–115. [PubMed: 20547574]
34. Chung JW, Bilimoria KY, Stulberg JJ, Quinn CM, Hedges LV. Estimation of Population Average Treatment Effects in the FIRST Trial: Application of a Propensity Score-Based Stratification Approach. *Health services research* 2018;53:2567–2590. [PubMed: 28833067]
35. Lu H, Cole SR, Hall HI et al. Generalizing the per-protocol treatment effect: The case of ACTG A5095. *Clinical trials (London, England)* 2019;16:52–62.
36. Hong JL, Jonsson Funk M, LoCasale R et al. Generalizing Randomized Clinical Trial Results: Implementation and Challenges Related to Missing Data in the Target Population. *Am J Epidemiol* 2018;187:817–827. [PubMed: 29020193]
37. Berkowitz SA, Sussman JB, Jonas DE, Basu S. Generalizing Intensive Blood Pressure Treatment to Adults With Diabetes Mellitus. *J Am Coll Cardiol* 2018;72:1214–1223. [PubMed: 30189998]
38. Alkhoul M, Holmes DR Jr, Carroll JD et al. Racial Disparities in the Utilization and Outcomes of TAVR: TVT Registry Report. *JACC Cardiovascular interventions* 2019;12:936–948. [PubMed: 31122351]

Perspectives

What is known:

Randomized controlled trials (RCTs) are the gold standard to understand effectiveness of an intervention, but whether trial results apply to patients encountered in clinical practice is often unknown.

What is new:

TAVR trial and real-world populations receiving TAVR were mostly similar, though had some notable differences. Nevertheless, the extrapolated real-world treatment effect of TAVR was at least as high as the observed trial treatment effect, suggesting that the absolute benefit of TAVR in clinical trials is similar to the benefit of TAVR in the US real-world setting.

What is next:

Future research can use the methods in this manuscript to examine the applicability of other cardiovascular clinical trials to real-world practice.

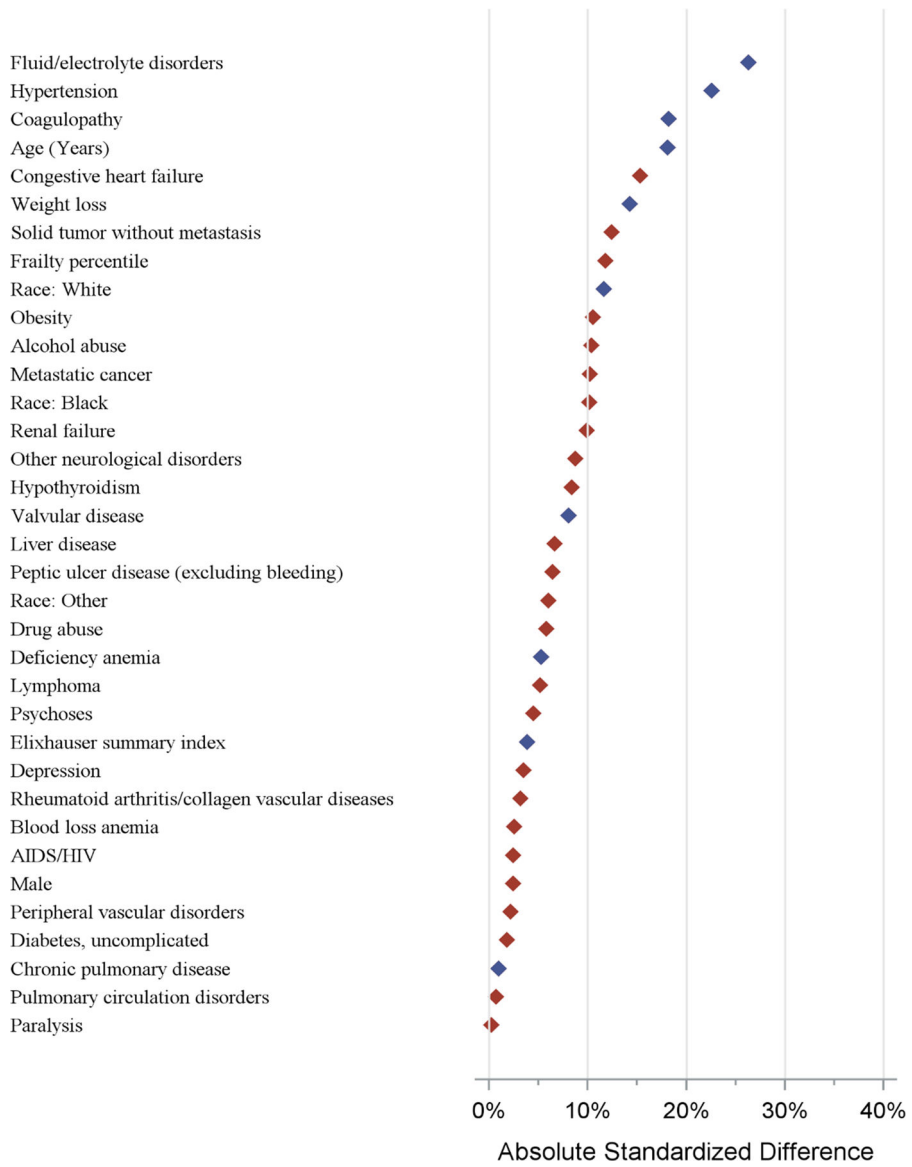


Figure 1. Standardized mean differences in patient characteristics between trial and real-world cohorts. Blue diamond covariates represent positive standardized difference (more frequent in trial), and the red diamond covariates represent negative standardized differences (more frequent in non-trial).

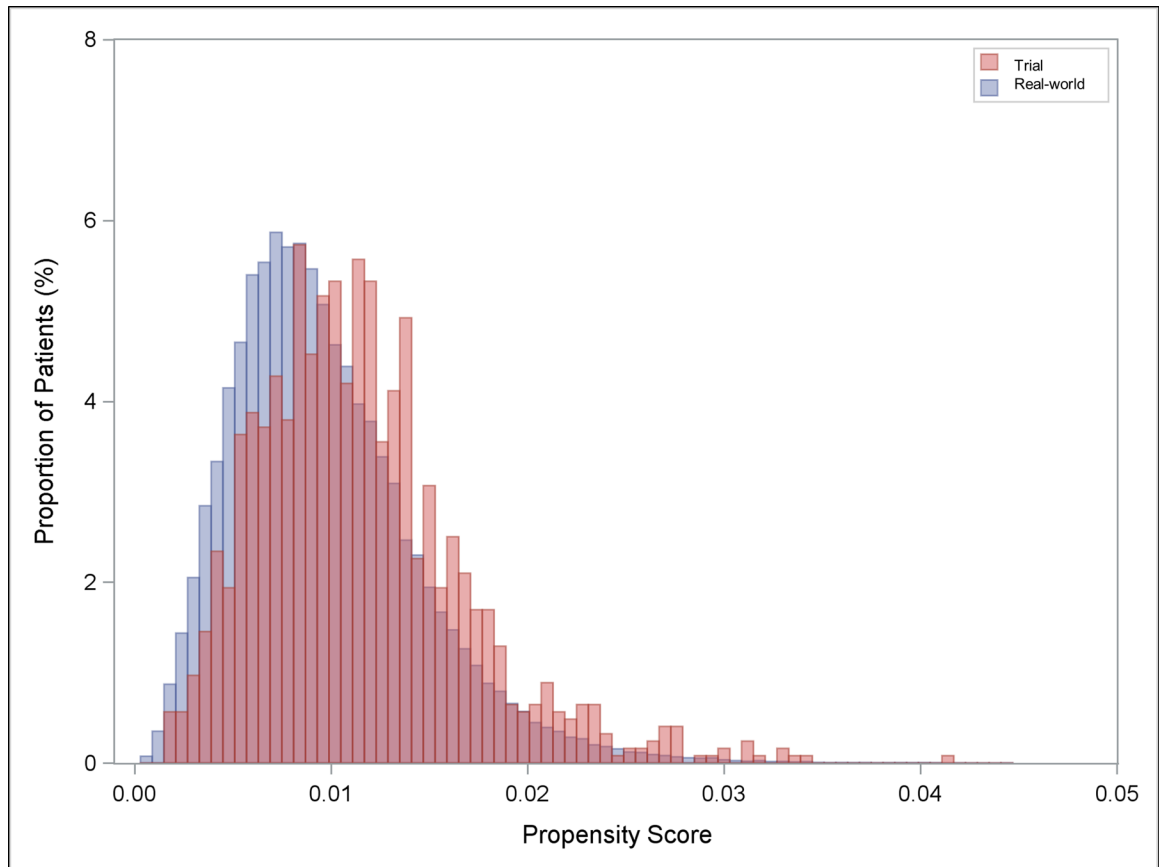
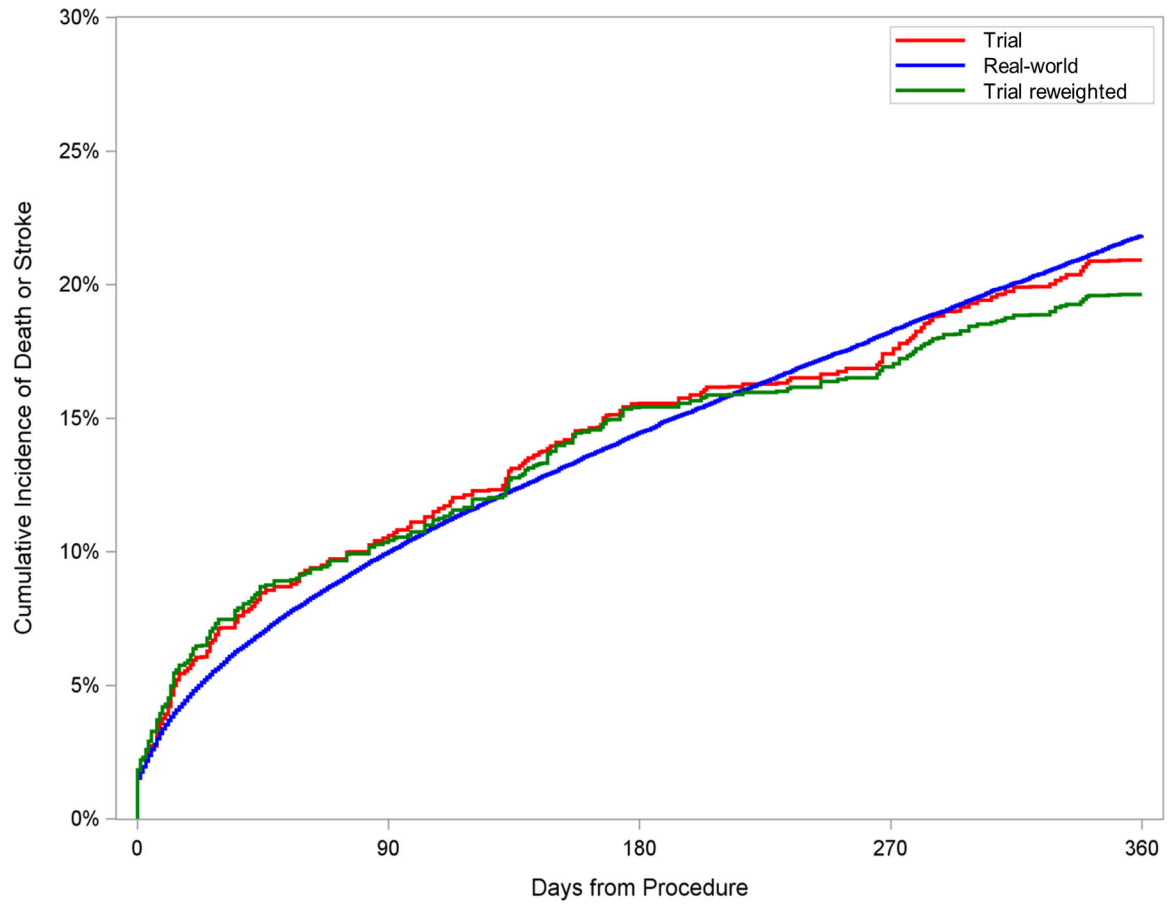
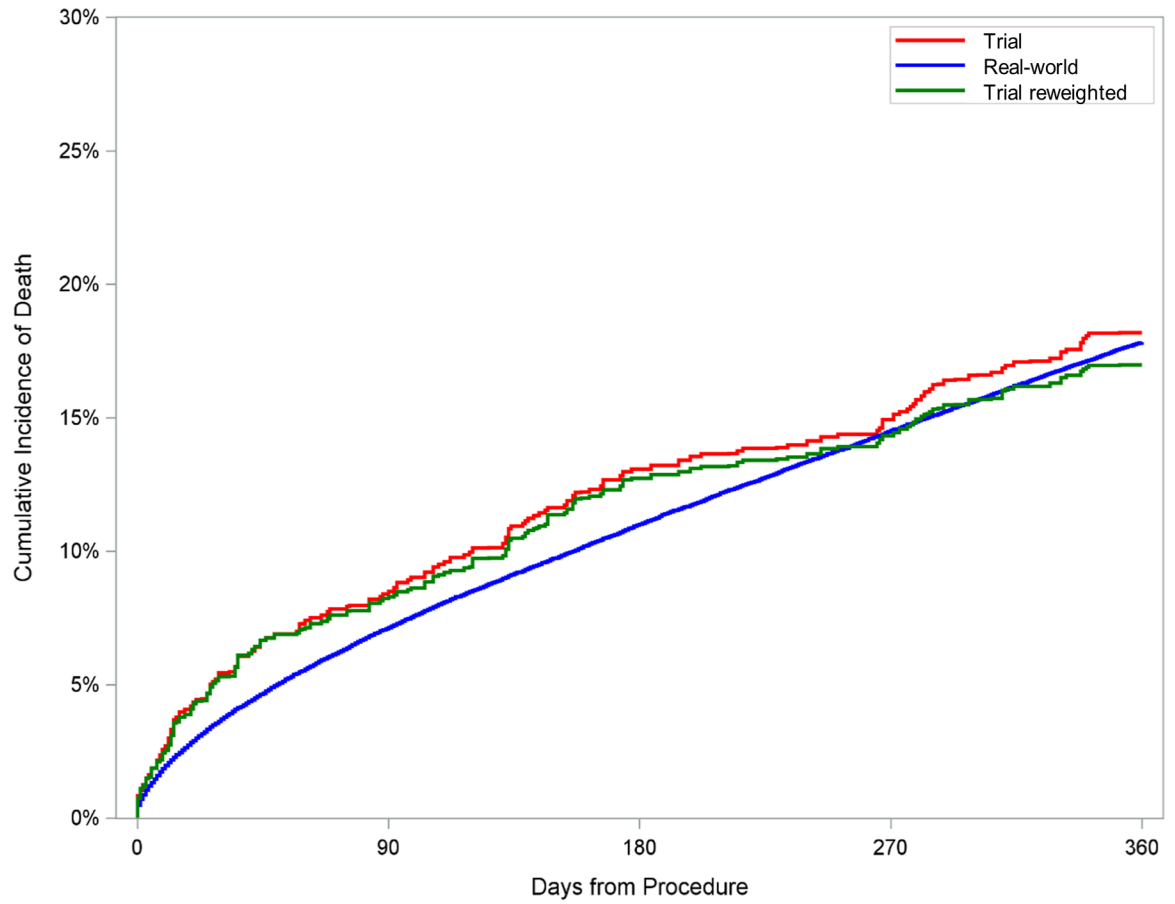


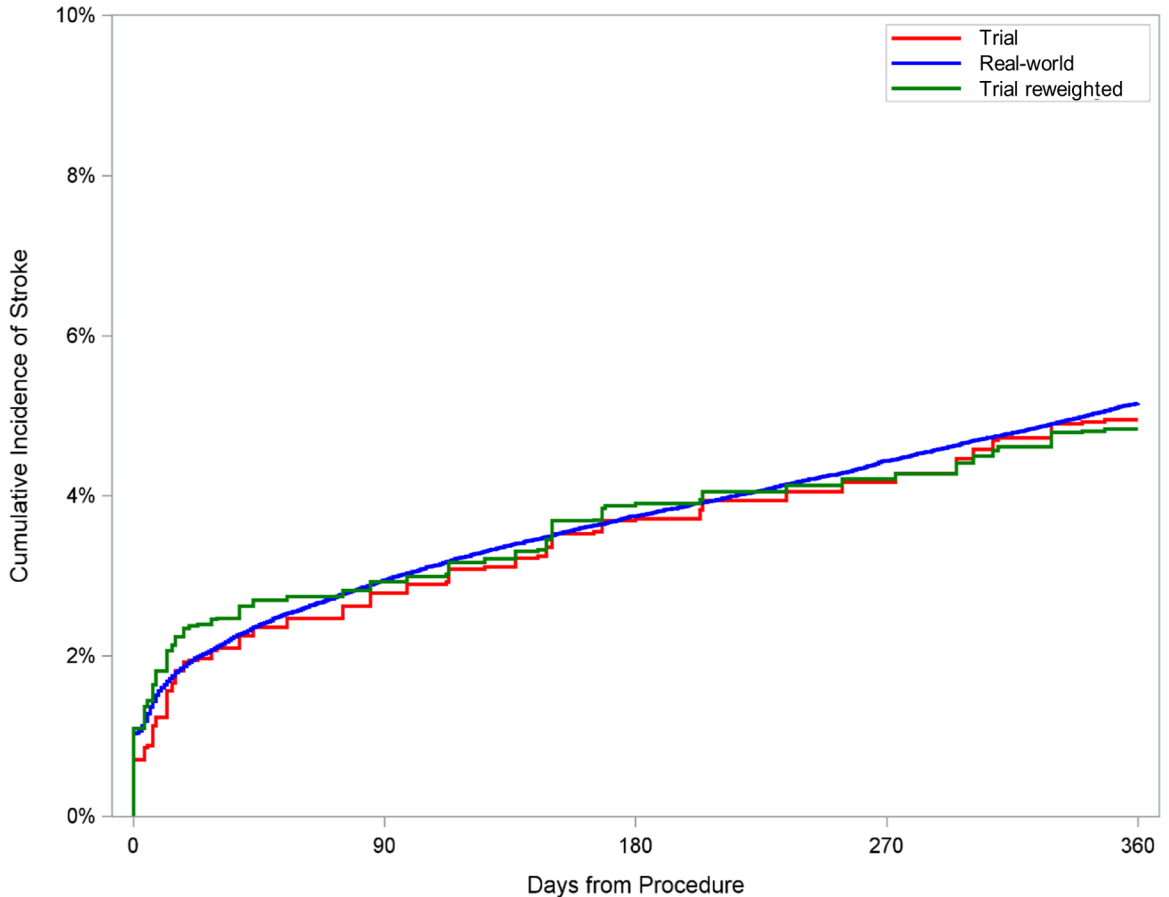
Figure 2. Distribution of propensity scores in pooled trial and real-world participants. Propensity score represents the predicted probability of inclusion in trial.



	0	90	180	270	360
Trial	1297.5	1178.3	1113.0	1088.6	1042.2
Real-world	133054.0	111833.0	97073.0	83974.0	72212.0
Trial Reweighted	1290.0	1177.2	1111.6	1091.7	1056.1

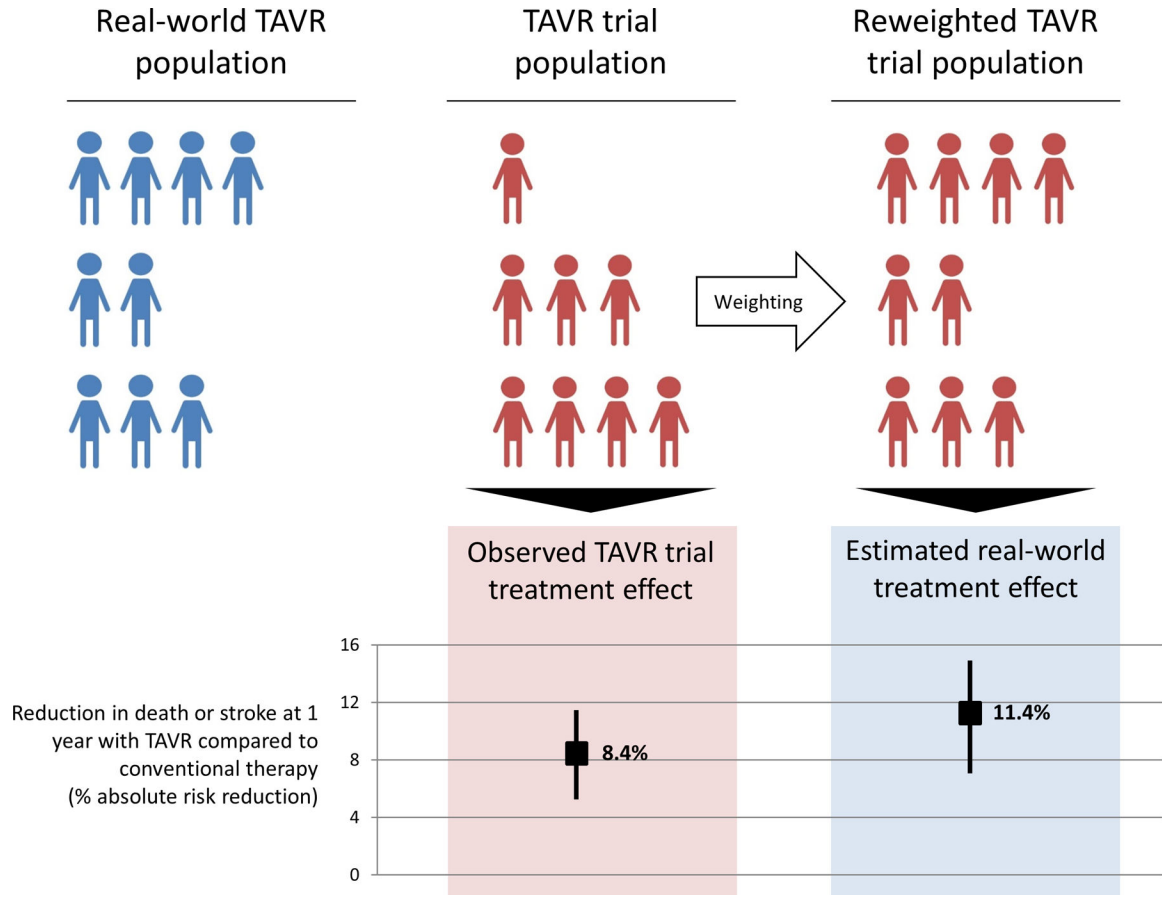


	0	90	180	270	360
Trial	1306.8	1205.8	1145.7	1121.2	1078.3
Real-world	134454.0	115459.0	101177.0	88166.0	76387.0
Trial Reweighted	1304.4	1205.3	1146.8	1125.8	1091.0



	0	90	180	270	360
Trial	1302.7	1178.3	1113.0	1088.6	1042.2
Real-world	133054.0	111833.0	97073.0	83974.0	72212.0
Trial Reweighted	1293.7	1177.2	1111.6	1091.7	1056.1

Figure 3. Cumulative incidence plot of outcomes for trial cohort, reweighted trial cohort, and observed real-world cohort among patients receiving TAVR. The trial cohort curve represents cumulative incidence of outcomes after combining CoreValve Extreme Risk, High Risk, and SURTAVI trials in proportion according to the distribution of extreme, high, and intermediate risk patients in the TVT registry. The reweighted trial cohort curve represents the cumulative incidence of outcomes after reweighting the trial cohort to represent the distribution of characteristics among real-world patients receiving TAVR. The observed real-world cohort curve represents the cumulative incidence of outcomes observed in claims data among real-world patients receiving TAVR. A. Death or stroke. B. Death. C. Stroke.



Central Illustration.

The CoreValve Extreme Risk, High Risk, and SURTAVI trial populations were reweighted to resemble the contemporary real-world population of patients undergoing TAVR with regards to all measured covariates. The CoreValve trial treatment effects were then re-estimated in this reweighted sample. The re-weighted trial results represent the cumulative incidence of outcomes expected among patients in the real-world cohort if they had all been in the trials. Conventional therapy refers to surgical aortic valve replacement for intermediate or high risk patients and medical therapy for extreme risk patients.

Table 1.

Baseline characteristics of CoreValve trial participants and real-world patients receiving TAVR

Subject Characteristic	Trial [95% CI] (N = 2026 Subjects)	Real-world [95% CI] (N = 135112 Subjects)	Standardized Difference	p-value*
Demographics				
Age (yrs) †				
Mean ± SD	83.3±6.6 (2026)	82.0±7.5 (135112)	0.18	<.001
95% CI	[83.0,83.6]	[82.0,82.1]		
Male	51.4% [49.2%,53.6%]	52.6% [52.3%,52.9%]	-0.02	0.279
Race				<.001
White	95.7% [94.7%,96.6%]	93.0% [92.9%,93.2%]	0.12	
Black	2.1% [1.5%,2.8%]	3.8% [3.7%,3.9%]	-0.10	
Other	2.2% [1.6%,2.9%]	3.1% [3.1%,3.2%]	-0.06	
Comorbidities				
Congestive heart failure	68.0% [66.0%,70.1%]	74.9% [74.7%,75.2%]	-0.15	<.001
Valvular disease	99.5% [99.1%,99.8%]	98.8% [98.7%,98.8%]	0.08	0.003
Pulmonary circulation disorders	20.2% [18.5%,22.0%]	20.5% [20.3%,20.7%]	-0.01	0.761
Peripheral vascular disorders	27.3% [25.4%,29.3%]	28.3% [28.0%,28.5%]	-0.02	0.337
Hypertension	50.1% [47.9%,52.3%]	39.0% [38.7%,39.2%]	0.23	<.001
Paralysis	2.3% [1.7%,3.0%]	2.3% [2.2%,2.4%]	-0.00	0.925
Other neurological disorders	5.7% [4.7%,6.8%]	7.9% [7.8%,8.1%]	-0.09	<.001
Chronic pulmonary disease	28.9% [26.9%,30.9%]	28.5% [28.2%,28.7%]	0.01	0.678
Diabetes, Combined uncomplicated and complicated	36.1% [34.0%,38.2%]	37.0% [36.7%,37.2%]	-0.02	0.419
Hypothyroidism	18.2% [16.5%,19.9%]	21.5% [21.3%,21.7%]	-0.08	<.001
Renal failure	32.3% [30.2%,34.4%]	37.0% [36.7%,37.2%]	-0.10	<.001
Liver disease	1.6% [1.1%,2.3%]	2.6% [2.5%,2.6%]	-0.07	0.007
Peptic ulcer disease (excluding bleeding)	0.1% [0.0%,0.4%]	0.5% [0.4%,0.5%]	-0.06	0.022
AIDS/HIV	0.0% [0.0%,0.2%]	0.0% [0.0%,0.0%]	-0.02	0.438
Lymphoma	0.6% [0.3%,1.1%]	1.1% [1.0%,1.2%]	-0.05	0.041
Metastatic cancer	0.0% [.,0.2%]	0.6% [0.5%,0.6%]	-0.10	<.001
Solid tumor without metastasis	0.8% [0.5%,1.3%]	2.3% [2.3%,2.4%]	-0.12	<.001
Rheumatoid arthritis/collagen vascular diseases	4.2% [3.4%,5.2%]	4.9% [4.8%,5.0%]	-0.03	0.170
Coagulopathy	24.6% [22.7%,26.5%]	17.2% [17.0%,17.4%]	0.18	<.001
Obesity	13.3% [11.8%,14.8%]	17.0% [16.8%,17.2%]	-0.11	<.001
Weight loss	7.6% [6.5%,8.8%]	4.2% [4.1%,4.3%]	0.14	<.001
Fluid/electrolyte disorders	31.5% [29.5%,33.5%]	20.1% [19.9%,20.3%]	0.26	<.001
Blood loss anemia	1.0% [0.6%,1.5%]	1.3% [1.2%,1.3%]	-0.03	0.281
Deficiency anemia	25.6% [23.7%,27.5%]	23.3% [23.1%,23.5%]	0.05	0.017

Subject Characteristic	Trial [95% CI] (N = 2026 Subjects)	Real-world [95% CI] (N = 135112 Subjects)	Standardized Difference	p-value *
Alcohol abuse	0.1% [0.0%,0.4%]	0.8% [0.8%,0.9%]	-0.10	<.001
Drug abuse	0.0% [.,0.2%]	0.2% [0.2%,0.2%]	-0.06	0.058
Psychoses	0.4% [0.2%,0.8%]	0.7% [0.7%,0.8%]	-0.04	0.078
Depression	6.7% [5.6%,7.8%]	7.5% [7.4%,7.7%]	-0.03	0.130
Frailty Index [†]				
Mean ± SD	-0.1±1.0 (2026)	0.0±1.0 (135112)	-0.12	<.001
95% CI	[-0.2,-0.1]	[-0.0,0.0]		
Frailty Percentile [†]				
Mean ± SD	47.2±28.6 (2026)	50.5±28.9 (135112)	-0.12	<.001
95% CI	[45.9,48.4]	[50.4,50.7]		

The 3 CoreValve pivotal trials (Extreme Risk, High Risk, and SURTAVI) were pooled together based on the real-world proportion of TAVR volume in each of these risk categories.

CI=confidence interval; AIDS=acquired immunodeficiency syndrome; HIV=human immunodeficiency virus

* Student's T-tests for continuous variables and Pearson Chi-square tests for categorical variables.

[†] Mean (standard deviation) presented.

Table 2.

Cumulative incidence of outcomes at 1 year in trial participants receiving TAVR and reweighted trial participants representing real-world patients receiving TAVR

Outcome	Trial [95% CI]	Reweghted trial representing real-world patient characteristics* [95% CI]
Death or Stroke	20.93% [18.24%, 23.62%]	19.63% [16.60%, 22.66%]
Death	18.19% [15.64%, 20.74%]	16.98% [14.11%, 19.84%]
Stroke	4.95% [3.46%, 6.45%]	4.83% [3.14%, 6.53%]
Aortic valve reintervention	1.07% [0.36%, 1.78%]	1.14% [0.30%, 1.98%]
Pacemaker placement	20.18% [17.48%, 22.87%]	19.90% [16.80%, 22.99%]

* Reweghted trial outcomes are calculated by employing an inverse-probability weighting (IPW) method to reweight the cumulative incidence of events observed in the trial cohort who received TAVR, based on the distribution of characteristics in the real-world TAVR cohort.

CI = confidence interval

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Comparison of trial and estimated real-world treatment effects

Outcome	Trial treatment effect (95% CI)	Estimated real-world treatment effect (95% CI)	Difference (95% CI)
Death or stroke	8.38% [4.62%, 11.91%]	11.37% [7.50%, 14.92%]	2.99% [0.82%, 5.53%]
Death	6.95% [3.73%, 10.31%]	8.74% [5.20%, 12.32%]	1.79% [-0.43%, 4.12%]

Treatment effect is based on comparison with SAVR for intermediate and high risk patients and comparison with medical therapy for extreme risk patients.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript