Transcatheter Aortic Valve Replacement in Women
Confirming Translational Science or More Confounded Clinical Research?*

Joseph E. Bavaria, MD, Jay Giri, MD, MPH

In both animals and humans, differences have been noted in left ventricular remodeling between sexes when faced with the chronic pressure overload of aortic valve stenosis (AS) (1,2). Women, on average, develop thicker ventricles with greater preservation of ejection fraction and less fibrosis than male counterparts when faced with AS (2,3). Hence, despite being labeled uniformly as “severe AS patients” when presenting for aortic valve replacement (AVR), women and men may demonstrate markedly different phenotypes when considering the complete physiological picture of ventricular performance and valvuloarterial impedance. Initial studies suggest that left ventricular mass regression occurs more frequently and rapidly in women after AVR, allowing them to normalize ventricular biomechanics to a greater degree than men (4). Taken in sum, these translational data lead to the intriguing possibility that women may exhibit a differential and positive treatment response to AVR for AS. To date, clinical outcomes data addressing this question have been mixed (5–7).

In this issue of the Journal, O’Connor et al. (8) add to the relatively sparse existing published data on this topic by comparing outcomes between the sexes in high and extreme surgical risk patients that underwent transcatheter aortic valve replacement (TAVR). The authors should be congratulated for working collaboratively to share the data necessary to perform this analysis. In this era of big data with standardized outcome assessments and the ability for carefully planned post-hoc statistical analysis, such observational efforts are a vital window into broad practice patterns and outcomes. As President Harry S. Truman noted, “It is amazing what you can accomplish if you don’t care who gets the credit” (9).

As opposed to a traditional meta-analysis that relies only on published data to compute and combine effect sizes, the current investigation pools patient-level data from the original study datasets to create a new, larger database for de novo analysis. In situations like the present study, where the criteria for subject inclusion in the source trials do not exhibit much heterogeneity, this design may allow for more granular assessment of potential predictors of post-TAVR outcomes than that of a traditional meta-analysis. Despite these methodological strengths, the study design remains observational, and the analysis is potentially victim to treatment-selection bias.

The authors note several key findings. As has been seen in previous research, women were more likely to experience vascular complications and less likely to be left with post-procedural aortic insufficiency (10,11). The most important findings, though, related to mortality. In comparison to men, women experienced similar 30-day mortality rates but had significantly lower 1- and 2-year mortality after TAVR.
Do these findings reflect a true differential treatment response to TAVR in women, confirming the theorized pathophysiological mechanisms noted in the preceding text? Or are the present findings actually a result of comparing a lower-risk cohort of women with a higher-risk cohort of men?

In order to address this, we must ask ourselves why lower-risk women may have been systematically chosen for inclusion in these seminal trials and registries of high- and extreme-risk TAVR. An answer lies in the inclusion criteria of the studies, which relied heavily upon quantitative Society of Thoracic Surgeons or EuroScore risk scores that were designed to predict 30-day mortality after surgical aortic valve replacement (SAVR). The risk models were used during surgical evaluations and heart team meetings as a quantitative metric to characterize patients as high risk for SAVR, thus justifying utilization of TAVR. The risk models code female sex as an independent risk factor for mortality. Thus, other factors being equal, women have higher Society of Thoracic Surgeons and EuroScore risk scores for 30-day mortality after SAVR than their male counterparts. And, in fact, these predictive scores proved fairly accurate in distinguishing early operative risk among the male and female cohorts when applied to TAVR patients in the present study (8). Predicted risks between the cohorts were similar, and the actual 30-day mortalities were nearly identical. The implication is that women, despite entering the TAVR procedure with significantly fewer comorbidities than their male counterparts, sustained a similar periprocedural mortality rate as was predicted by pre-operative SAVR risk models.

However, notable differences over longer-term follow-up were seen, with women achieving higher survival rates at 1 year and mortality curves continuing to diverge to 2 years, with women increasing their survival advantage over time. From a valvular standpoint, it is well known that para-avalvular aortic insufficiency was the “Achilles’ heel” of the early-generation TAVR technology used in these datasets. Aortic insufficiency was less common in women because of their smaller average annular sizes and the resultant more aggressive TAVR bioprosthesis oversizing. Importantly, the authors controlled for differences in moderate or greater aortic insufficiency in their Cox proportional hazard model, ideally mitigating some of its influence on the observed results (8). However, mild aortic insufficiency, which was also more common in men in these studies, represents a confounder, given its known relationship with late mortality (12). Nevertheless, the differences in rates of aortic insufficiency do not seem nearly great enough to account for such significant observed mortality differences.

Apart from the issue of paravalvular aortic insufficiency, TAVR devices have proven to be quite durable technically, with late death rates largely attributed to the high burdens of comorbid conditions, both cardiac and noncardiac, present in those treated with the procedure (13,14). In the current study (8), observed comorbidities were far greater in men, and attempts were made to control for these measured covariates. However, it is a distinct possibility that this widespread and significantly higher general burden of comorbid illness among men served as a marker for other unmeasured confounding conditions that placed them at higher risk of late death post-TAVR. The initial selection of patients for TAVR was based upon scoring systems used to stratify patients into high-risk cohorts that were designed to predict only 30-day operative results and were not TAVR specific. Thus, the late mortality benefits seen in women may have emerged as a result of their relative absence of comorbid conditions, as opposed to a differentially positive response to TAVR over long-term follow-up.

We cannot be sure that the multivariable model used in the current analysis was adequate to control for baseline differences in risk between men and women. However, studies such as this raise important methodological issues for the conduct of future observational research with large prospective datasets. One way to more definitively assess for unmeasured confounding in observational studies is through the analysis of falsification endpoints (15). A falsification endpoint is an outcome highly unlikely to be influenced by the studied intervention that can be compared between matched groups to verify true absence of confounding, similar to a negative control experiment in a basic science investigation.

For prospective TAVR registries, consideration should be given to pre-specifying several falsification endpoints as part of data collection efforts. Examples of potential falsification endpoints for an analysis of late mortality after TAVR include incident hip fracture or urosepsis. Both are conditions that are unlikely to be associated with the initial procedure, but could raise concerns about unmeasured confounding if they were found to be unbalanced in purportedly “adjusted” cohorts during post-hoc observational comparative analyses.

Clearly, characterizing the influence of sex upon TAVR outcomes is not straightforward. Sex is an immutable trait, hence treatment-selection bias
cannot be mitigated through randomization. As such, the best way to determine effects of sex on TAVR outcomes will be to develop a TAVR-specific long-term mortality risk model utilizing a truly unselected patient population. It would be interesting to see whether female sex emerges as a “protective” factor in such a model. Given the rapid promulgation of TAVR into lower-risk cohorts and the careful prospective data accumulation taking place in multiple national registries, such a goal may be achievable in the not-too-distant future.

REPRINT REQUESTS AND CORRESPONDENCE: Dr. Joseph E. Bavaria, Cardiovascular Surgery Division, Hospital of the University of Pennsylvania, 6 Silverstein, 3400 Spruce Street, Philadelphia, Pennsylvania 19104. E-mail: joseph.bavaria@uphs.upenn.edu.

REFERENCES

KEY WORDS: aortic stenosis, gender, sex, transcatheter aortic valve implantation, transcatheter aortic valve replacement