



RESEARCH ARTICLE

REVISED Patient-centered benefit-risk analysis of transcatheter aortic valve replacement [version 5; peer review: 3 approved]

Kevin Marsh ¹, Natalia Hawken², Ella Brookes¹, Carrie Kuehn³, Barry Liden ³

¹Patient-Centered Research, Evidera Inc, London, UK
²Modus Outcomes, London, UK
³Patient Engagement, Edwards Lifesciences, Washington D.C., USA

v5 First published: 08 Apr 2019, 8:394
<https://doi.org/10.12688/f1000research.18796.1>
 Second version: 14 May 2019, 8:394
<https://doi.org/10.12688/f1000research.18796.2>
 Third version: 21 Jun 2019, 8:394
<https://doi.org/10.12688/f1000research.18796.3>
 Fourth version: 05 May 2020, 8:394
<https://doi.org/10.12688/f1000research.18796.4>
 Latest published: 08 Feb 2021, 8:394
<https://doi.org/10.12688/f1000research.18796.5>

Abstract

Background: Aortic stenosis (AS) treatments include surgical aortic valve replacement (SAVR) and transcatheter aortic valve replacement (TAVR). Choosing between SAVR and TAVR requires patients to trade-off benefits and risks. The objective of this research was to determine which TAVR and SAVR outcomes patients consider important, collect quantitative data about how patients weigh benefits and risks, and evaluate patients' preferences for SAVR or TAVR.
Methods: Patients were recruited from advocacy organization databases. Patients self-reported as being diagnosed with AS, and as either having received AS treatment or as experiencing AS-related physical activity limitations. An online adapted swing weighting (ASW) method – a pairwise comparison of attributes – was used to elicit attribute trade-offs from 219 patients. Survey data were used to estimate patients' weights for AS treatment attributes, which were incorporated into a quantitative benefit-risk analysis (BRA) to evaluate patients' preferences for TAVR and SAVR.
Results: On average, patients put greater value on attributes that favored TAVR than SAVR. Patients' valuation of the lower mortality rate, reduced procedural invasiveness, and quicker time to return to normal quality of life associated with TAVR, offset their valuation of the time over which SAVR has been proven to work. There was substantial heterogeneity in patients' preferences. This was partly explained by age, with differences in preference observed between patients <60 years to those ≥60 years. A Monte Carlo Simulation found that 79.5% of patients prefer TAVR.
Conclusions: Most AS patients are willing to tolerate sizable increases in clinical risk in exchange for the benefits of TAVR, resulting in a large

Open Peer Review

Reviewer Status

	Invited Reviewers		
	1	2	3
version 5 (revision) 08 Feb 2021			 report
version 4 (update) 05 May 2020			 report
version 3 (revision) 21 Jun 2019			
version 2 (revision) 14 May 2019	 report 	 report 	
version 1 08 Apr 2019	 report	 report	

1. **Ross Jaffe**, Versant Ventures, San Francisco, USA
 2. **Megan Coylewright**, Dartmouth-Hitchcock Medical Center, Lebanon, USA
 3. **Juan Marcos Gonzalez** , Duke University School of Medicine, Durham, USA

Any reports and responses or comments on the

proportion of patients preferring TAVR to SAVR. Further work should be undertaken to characterize the heterogeneity in preferences for AS treatment attributes. Shared decision-making tools based on attributes important to patients can support patients' selection of the procedure that best meets their needs.

.....
article can be found at the end of the article.

Keywords

TAVR, aortic valve, transcatheter, patient preference, benefit-risk analysis

Corresponding author: Kevin Marsh (Kevin.Marsh@evidera.com)

Author roles: **Marsh K:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Hawken N:** Conceptualization, Investigation, Methodology, Project Administration, Supervision, Validation, Writing – Original Draft Preparation; **Brookes E:** Conceptualization, Project Administration, Visualization, Writing – Original Draft Preparation; **Kuehn C:** Conceptualization, Funding Acquisition, Methodology, Resources, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Liden B:** Conceptualization, Funding Acquisition, Methodology, Resources, Validation, Writing – Review & Editing

Competing interests: Barry Liden and Carrie Kuehn are employees of Edwards Lifesciences. Kevin Marsh and Ella Brookes are salaried employees of Evidera and are not allowed to accept remuneration from any clients for their services. Natalia Hawken was a salaried employee of Evidera at the time of carrying out this study. Evidera received funding from Edwards Lifesciences to conduct the study and develop this manuscript.

Grant information: Funding for this project was provided by Edwards Lifesciences, Washington D.C., US.

Copyright: © 2021 Marsh K *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Marsh K, Hawken N, Brookes E *et al.* **Patient-centered benefit-risk analysis of transcatheter aortic valve replacement [version 5; peer review: 3 approved]** F1000Research 2021, 8:394 <https://doi.org/10.12688/f1000research.18796.5>

First published: 08 Apr 2019, 8:394 <https://doi.org/10.12688/f1000research.18796.1>

UPDATE Updates from Version 4

The present version of this manuscript updates previous versions to incorporate feedback provided through peer review, including updates to the discussion section to clarify some of the assumptions made, and to include more reflections on the potential challenges associated with the adapted swing weighting methodology.

Any further responses from the reviewers can be found at the end of the article

Introduction

Aortic stenosis (AS) is a progressive cardiovascular condition resulting from narrowing (or stenosis) of the aortic valve. This narrowing prevents the valve from fully opening, decreasing blood flow out of the heart and forcing it to work harder to maintain sufficient blood flow. If left untreated, AS can lead to severe cardiovascular complications and death¹. As of 2015, 12.4% of the United States population over age 75 (nearly 2.5 million people) were reportedly diagnosed with AS^{2,3}. More than one in eight people (13.3%) over the age of 75 in the US have moderate-to-severe AS⁴. Patients with AS may be asymptomatic for many years until the valve is narrowed severely enough to cause symptoms. Symptoms of AS include chest pain and angina, syncope, dyspnea, fatigue, and palpitations, all of which are exacerbated by physical activity⁵. Under-treatment of AS patients is common, with more than half of patients referred to cardiologists failing to receive surgical treatment⁶. Once symptoms appear, often between the ages of 70 and 80 years old, the prognosis of untreated patients is poor⁷. Among untreated patients, average survival is 50% at two years and 20% at five years after the onset of AS symptoms.

The traditionally recommended treatment for AS is surgical aortic valve replacement (SAVR). Such invasive surgery, involving a large incision in the chest, may not be suitable for all patients, especially those with comorbidities. The alternative, transcatheter aortic valve replacement (TAVR), is a minimally invasive, catheter-based procedure to replace the aortic valve in patients with AS. TAVR is the first-line therapy for inoperable patients with severe AS and an alternative to SAVR in operable high-risk patients. Among patients who are at intermediate surgical risk, both TAVR and SAVR are associated with improvements in disease-specific and generic health status⁸. However, TAVR is associated with a reduced rate of complications and a quicker recovery time, with patients returning to a normal quality of life more quickly⁹. When first available, the benefits of TAVR were offset by reportedly increased risks of stroke and the need for a pacemaker^{10–12}. Recent clinical data reveal similar, if not improved rates of stroke and need for pacemaker among TAVR patients¹³. Furthermore, at a median two-year follow-up, all-cause mortality for patients undergoing TAVR was 20.2% compared with 21.9% for patients undergoing SAVR¹⁴.

The choice of TAVR or SAVR involves patients making trade-offs between multiple treatment attributes, including invasiveness, speed of recovery, mortality rates, and risks of complications. Given the challenging nature of this decision,

tools have been developed to support patient decision-making¹⁵. However, little is known about the weights that patients assign to each attributes, how they make trade-offs between these attributes, and whether and how these preferences vary between patients. The objective of this research was to determine which outcomes associated with TAVR and SAVR patients consider most important, collect quantitative data about how patients weigh the benefits and risks associated with TAVR and SAVR, and to use this data to evaluate patients' preferences for SAVR or TAVR.

Interim results from the first 93 participants enrolled into this study were published on 21 June 2019 as Version 3. This manuscript reports the final results from this study, based on a sample of 219 participants.

Methods

Overview

Patients' preferences for TAVR or SAVR were assessed using a quantitative benefit-risk assessment (BRA). This involved identifying attributes that distinguish TAVR and SAVR, measuring TAVR and SAVR performance on these attributes, eliciting patients' preferences for these attributes, and combining performance and preference in a BRA to determine what proportion of patients prefer TAVR or SAVR. The patient-preference survey upon which the BRA is based was fielded from July 2018 – January 2019.

Attribute selection

A long list of potential attributes was identified by reviewing the attributes highlighted in previous patient preference studies for heart valve surgical interventions, published meta-analyses and clinical studies, and regulators' assessments of related products. Additional attributes were identified based on consultation with TAVR and SAVR clinical experts and from patient input. The final attributes used in the study were selected based on clinical and regulatory relevance, whether or not the attribute distinguishes between TAVR and SAVR, and to comply with the attribute set properties required of an additive BRA¹⁶. For example, to avoid overlap with 'mortality', the 'stroke' attribute was defined as 'disabling, non-fatal strokes.'

Descriptions of the final attributes included in the BRA are summarized in [Table 1](#).

Performance measurement

TAVR and SAVR performance against the final attributes were identified from the published literature and from clinical data ([Table 2](#)), focusing on data that had a high degree of reference and use within the clinical community. Available data for stroke risk, defined as "all stroke" (both fatal and non-fatal) in the literature, and independence, defined by Kansas City Cardiomyopathy Questionnaire (KCCQ)¹⁷ score, required transformation to estimate performance as defined by the final study attributes. To estimate the risk of non-fatal stroke only, available stroke risk data was adjusted using the mortality rate among patients with severe aortic stenosis enrolled in the PARTNER trial who suffered a stroke compared to the mortality rate among those in the trial who did not suffer a stroke¹⁸. 'Independence' was defined as the probability of achieving relief from AS

Table 1. Selected attributes for the patient-centered benefit-risk analysis of transcatheter aortic valve replacement vs surgical aortic valve replacement.

Attribute	Definition	Description provided to participants
Type of procedure	Type of procedure	The invasiveness of the procedure is described by three characteristics: The length and depth of the incision Whether your heart is stopped The number of days that you will need to be in the hospital following a procedure. There are two types of procedure: A minimally invasive procedure requiring, on average, 8 days in hospital. A small incision is made near your groin, and a valve is inserted and guided to your heart using a long tube through an artery. The tube is used to implant a new valve in the heart to replace the diseased aortic valve. An invasive procedure, requiring, on average, 12 days in hospital. A large cut about 25 cm long is made in your chest to access your heart. Then, your heart is stopped while a machine takes over your heart and lung function. A new valve is implanted to replace the damaged valve. Your heart is started again, and your chest is stitched closed.
Mortality	Number of patients out of 100 who will die within 1 month	The numbers of patients who will die from any cause within 1 month of having the procedure. Death could be due to complications from the procedure, from complications of aortic stenosis, or as a result of disabling stroke.
Disabling non-fatal stroke	The number of patients out of 100 who will experience a non-fatal disabling stroke within 1 month	The number of patients who will experience a non-fatal but disabling stroke within 1 month of having the procedure. If you experience a stroke, you will be hospitalized. If the stroke is severe, it may lead to temporary or permanent disability, such as paralysis, reduced mobility, and problems with thinking, memory and speech.
Independence	Number of patients out of 100 who experience greater independence within 1 month	The number of patients who experience improvement in daily activities (greater independence) following relief from aortic stenosis symptoms within 1 month of the procedure. The symptoms of aortic stenosis (shortness of breath, fatigue, chest pain, and dizziness), physical function, and quality of life are improved to an extent that you experience improvements in your degree of independence and ability to engage in activities of daily living.
New permanent pacemaker	The number of patients out of 100 who will require a pacemaker within 1 year	The number of patients that will need to have a pacemaker permanently implanted as a result of the procedure. Typically, a pacemaker is implanted under local anesthetics and you may be discharged the same day if you get your pacemaker in the morning.
Requirement for dialysis	The number of patients out of 100 who will require dialysis within 1 year	The number of patients that will experience kidney function damage that will need dialysis as a result of the procedure. A machine is used to do the kidney's job of cleaning the blood. If you need dialysis, you will need to go to the hospital three times a week, with each visit lasting 4 hours.
Time over which the procedure has been proven to work	The number of years for which your procedure has been available and proven to work	The number of years the procedure has been available and is proven to keep symptoms of aortic stenosis from coming back. Following this period, it is currently not known whether you will experience aortic stenosis symptoms again.

Table 2. Transcatheter aortic valve replacement (TAVR) and surgical aortic valve replacement (SAVR) performance against benefit-risk analysis attributes.

Attribute	Measure	TAVR		SAVR		Performance Range
		Mean	95% CI	Mean	95% CI	
Mortality (all cause) ¹⁹	One-month risk	0.011	0.005-0.017	0.040	0.028-0.053	0.005-0.053
Disabling non-fatal stroke ¹⁹	One-month risk	0.008	0.002-0.013	0.033	0.021-0.044	0.002-0.044
Independence ⁸	Probability of having relief from AS symptoms that have an effect on daily life	0.479	0.454-0.500	0.249	0.227-0.276	0.227-0.500
New permanent pacemaker ¹⁹	One-year risk	0.123	0.103-0.142	0.090	0.072-0.108	0.072-0.142
Requirement for dialysis ²⁰	One-year risk	0.032	0.021-0.042	0.047	0.034-0.060	0.021-0.060
Time over which the Procedure has been Proven to Work	Years	10		20		5 - 30

CI, confidence interval; AS, aortic stenosis.

symptoms within a month of a procedure. Given available data, this was estimated as the probability of achieving a total score of 75 on KCCQ¹⁷. The KCCQ is a standard patient reported outcome measure used in clinical trials of surgical and transcatheter heart valves²¹. Estimated mean KCCQ score and variation in this measure were transformed into the proportion of patients achieving a KCCQ total score of 75 at 1 month using procedures previously described in Marchini²².

Survey methodology

An adapted swing weighting (ASW) exercise was administered online to elicit patients' preferences for treatment attributes²³. The objective of the ASW exercise was to identify the level of change in an attribute that patients would be willing to accept in exchange for improving their procedure from 'invasive' to 'minimally invasive' (see Table 1 for definitions). The ASW exercise consisted of a series of pairwise comparisons of attributes—the 'invasiveness' attribute and one other attribute. Participants were shown 'current' and 'improved' levels on each attribute and asked which improvement they would prefer to make (an example choice question is shown in Figure 1). The 'current' levels were chosen to reflect the attribute performance levels of TAVR and SAVR (Table 2), to ensure they had credibility with patients, adjusted to ensure that patients had sufficient range to indicate the change in the attribute that would have the equivalent value as reducing invasiveness. Therefore, the exercises were not designed to directly elicit patients' willingness to tolerate the actual change observed with TAVR.

Each set of pairwise comparisons included three iterations of the choice question. Depending on the answer to the choice question, the level of improvement offered on the non-procedure attribute was updated in the direction that made the value of improvements on the two attributes more similar than in the previous question. The algorithm used to identify

participants' indifference point included the levels used in each of the three iterations and how these changed dependent on the participants' previous responses is shown in extended data, Appendix 1²⁴.

A pilot of the survey among five AS patients, carried out over a 4-week period in June 2018, ensured acceptable cognitive burden, clarity of the instructions, and ease of use of the elicitation software. Before completing the ASW exercises, participants were introduced to the attributes and their definitions. Participants also completed two ASW practice questions; interpretation of their response to these practice questions was tested to ensure participants understood how to complete each of the pairwise comparisons. Only eighteen participants (8.2%) incorrectly interpreted the meaning of their response to the first practice question, and no participants incorrectly interpreted their response to the second practice question. Each participant completed a proportion of the possible pairwise question—either 3 or 4 sets of pairwise comparison questions. Participants also completed clinical and demographic questions, and health literacy and numeracy questionnaires, available as extended data²⁴.

Participants

Potentially eligible participants were recruited by M3 Global Research via email and from the membership of Heart Valve Voice and Mended Hearts patient organizations through e-mails and advertisements on social media platforms. Patients from the American Heart Association membership who had given prior permission to receive mailings were also invited to participate. Potential participants were directed to an online screening tool, where their eligibility for the survey was determined. Participants had to meet specific inclusion/exclusion criteria to participate (Box 1). There were no predefined enrollment targets or stratification by other demographic characteristics. Following completion of the online screening tool, eligible participants

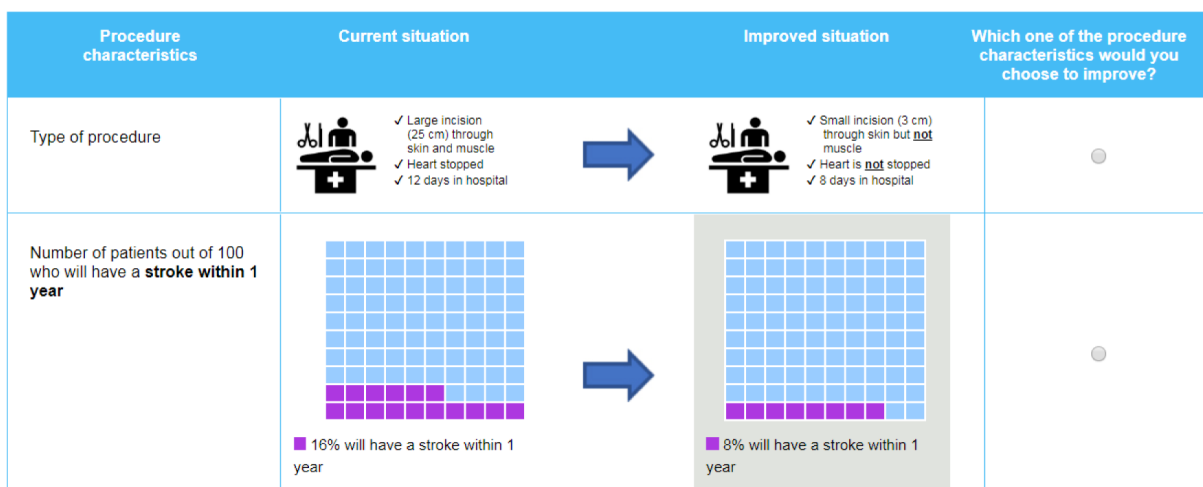


Figure 1. Example adapted swing weighting question presented to study participants. 'Type of Procedure' was used in every pairwise comparison. Only the comparison attribute was varied across different attributes and performance levels.

were directed to the web survey. They were also sent a link to the web survey, available as **extended data**²⁴, via e-mail so that they were able to complete the study at a time that was convenient for them.

Box 1. Eligibility criteria

Inclusion criteria

- Be at least 19 years of age
- Have a self-reported diagnosis of aortic valve disease
- Able to read and understand English
- Willing and able to complete an online survey
- Willing and able to provide (electronic) consent to participate in study
- Be a resident of the United States
- Have been treated procedurally (TAVR or SAVR) for AS in the past 10 years

or

- If untreated, must have experienced at least one of the following symptoms of AS:
 - Breathlessness/Shortness of breath
 - Tiredness/Lack of energy/Fatigue
 - Chest pain
 - Dizziness or blackouts/Syncope
- Be willing and able to participate in a telephone interview, and to be audio-recorded (qualitative pilot only)

Exclusion criteria

- Have a cognitive impairment, hearing difficulty, visual impairment, acute psychopathology
- Have insufficient knowledge of English, which could interfere with the patient’s ability to provide written consent and complete the web survey
- Are not experiencing at least one of the symptoms of AS, as described in the inclusion criteria

Analysis

Participant responses to the ASW questions identified the level of change in an attribute that patients would be willing to accept in exchange for improving their procedure from ‘invasive’ to ‘minimally invasive.’ For instance, a participant might assign the same value to a 2% reduction in mortality risk as reducing the invasiveness of the procedure. Inverting this relationship, responses were used to estimate the maximum acceptable increase in risks (MIR) or the maximum acceptable reduction in benefits (MRB) that participants would be willing to tolerate in exchange for moving from an invasive procedure to a minimally invasive procedure. In the above example, the participant would be willing to tolerate a MIR of 2% in mortality risk to reduce the invasiveness of the procedure.

In cases of incomplete or missing data, no data imputation was performed, and unanswered questions were coded as missing. Data from participants who completed only a portion of the

survey were included in the analyses if they had completed all the questions in at least one ASW exercise.

Participant preferences were incorporated into the following benefit-risk model constructed in Microsoft Excel.

$$U(x) = \sum_{i=1}^n w_i v_i(x_i) \quad \text{(Equation 1)}$$

Where

$U(x)$ is the overall value generated by procedure x .

w_i is the weight associated with attribute i ,

v_i is the partial value function for attribute i , which was assumed to be linear

x_i is the performance of performance x on attribute i

Participants’ MIR/MRB was converted into an attribute weight (w_i) by setting the weight for ‘type of procedure’ to 1, and then dividing the range in performance between TAVR and SAVR on that attribute (Table 2) by the MIR/MRB. For instance, if patients were willing to tolerate a 2% increase in mortality risk in exchange for a reduction in invasiveness, and reducing invasiveness is given a weight of 1, then the 4.8% change in mortality risk covered by the benefit-risk model (0.5%–5.3% range identified in Table 2) would be given a weight of 2.4 (4.8%/2%). Weights across all attributes in the model were then normalized to sum to 100.

Four outputs were generated from the benefit-risk model to evaluate patients’ preferences for TAVR and SAVR. First, the incremental overall value generated by TAVR:

$$U(\text{increment}) = U(\text{TAVR}) - U(\text{SAVR}) \quad \text{(Equation 2)}$$

Second, the incremental partial value on each attribute generated by TAVR:

$$PV_i(\text{increment}) = w_i v_i(x_{\text{TAVR}}) - w_i v_i(x_{\text{SAVR}}) \quad \text{(Equation 3)}$$

Where

x_{TAVR} is the performance of TAVR on attribute i

x_{SAVR} is the performance of SAVR on attribute i

Third, a threshold analysis was undertaken, estimating the level of performance (minimum acceptable benefit or maximum acceptable risk) on each attribute that would leave patients indifferent between TAVR and SAVR. More specifically:

$$MAR_a \text{ or } MAB_a = TAVR_a + ((U(\text{TAVR}) - U(\text{SAVR})) / w_i v_i(x_{1, \text{unit } a})) \quad \text{(Equation 4)}$$

Where

MAR_a or MAB_a is the maximum acceptable risk or minimum acceptable benefit of attribute a

$TAVR_a$ is the performance of TAVR on attribute a (see [Table 2](#)).

$x_{1\ unit\ a}$ is a single unit of performance on attribute a

[Table 3](#) outlines the definitions of MIR, MRB, MAR, and MAB.

Fourth, a Monte Carlo Simulation (MCS) was run to explore uncertainty in model inputs. That is, the benefit-risk model was run 10,000 times for both TAVR and SAVR. In each instance, the model drew from the distribution around both performance and weight inputs. Specifically, performance inputs were drawn from the distribution around TAVR and SAVR performance data ([Table 2](#)). Weight inputs were drawn in a manner that reflected the probability that participants identified different MIR/MRBs in their responses to the survey. For each iteration of the MCS, TAVR and SAVR, we've ranked based on U ([Equation 2](#)) and the proportion of instances that TAVR ranked first was estimated.

Ethics

In accordance with ethical practice, Institutional Review Board (IRB) approval was obtained through Advarra (approval MOD00300398 and MOD00354863) to comply with human participants research requirements prior to initiation of participant recruitment or administration of measures in the pilot or main studies. Informed consent was recorded electronically via the online survey platform. Any change to the protocol and/or informed consent form was resubmitted to the IRB for review and approval prior to implementation. The study was available for monitoring, auditing, IRB review, and regulatory inspection as applicable.

Results

Demographic characteristics of participants

A total of 219 patients completed the survey over two rounds ([Table 4](#)). Raw data are available from [Open Science Framework](#)²⁴. The majority of patients were less than 60 years old ($n=132$, 60.3%). More than half of the respondents were female ($n=128$, 58.4%) and a majority were white ($n=173$, 79.0%). Over half of the sample had completed a college degree or

higher ($n=157$, 71.7%). Most of the participants lived with a partner/spouse, family, or friends ($n=164$, 74.9%). No participants reported experiencing severe limitations to their physical activity. Few patients demonstrated low health literacy ($n=23$, 10.5%) or numeracy ($n=22$, 10.0%), and only 7 patients (3.2%) were low on both scales.

Responses to ASW questions

When responding to the ASW questions, only a small proportion (8.68%) of participants 'straight-lined' on all questions—consistently choosing to improve either 'procedure' or the comparison attribute across all three iterations of the choice question. These responses may be a valid reflection of participants' preference—suggesting a strong preference either for avoiding an invasive procedure, or a strong preference to prioritize improving other procedure attributes. Thus, these responses were included in the analysis. The impact of excluding these data were tested, and it was determined that results of the BRA were not sensitive to whether these data were included or excluded.

Comparisons of TAVR and SAVR

[Table 5](#) shows the difference in performance of TAVR compared with SAVR on each attribute, and patients' willingness to accept this difference in exchange for the lesser invasiveness of TAVR. The increase in risks or the reduction in benefits that patients are, on average, willing to accept in exchange for reducing procedure invasiveness is reported in the middle three columns. For instance, patients would be willing to tolerate a 6.69% increase in the probability of experiencing disabling, non-fatal stroke in exchange for the reduction in invasiveness associated with receiving TAVR instead of SAVR. In each case, patients were on average willing to accept TAVR's performance on any attribute in exchange for its lower invasiveness. In the case of four attributes (mortality, disabling non-fatal stroke, independence, and dialysis), TAVR performs better than SAVR. Where SAVR performs better than TAVR (the need for new permanent pacemaker and time over which the procedure has been proven to work), patients would, on average, be willing to accept TAVR's performance given its lower invasiveness. For example, participants are willing to tolerate a 6.98% increase in the risk of a new permanent pacemaker, while the probability of

Table 3. Benefit-risk analysis definitions.

Outcome	Definition
MIR	Maximum Acceptable Increase in Risk: The maximum acceptable increase in risk for a single attribute that patients would tolerate in exchange for reducing procedure invasiveness.
MRB	Maximum Acceptable Reduction in Benefit: The maximum acceptable reduction in benefit for a single attribute that patients would tolerate in exchange for reducing procedure invasiveness.
MAR	Maximum Acceptable Risk: The maximum acceptable risk that would make patients indifferent between TAVR and SAVR.
MAB	Minimum Acceptable Benefit: The minimum acceptable benefit that would make patients indifferent between TAVR and SAVR.

TAVR, transcatheter aortic valve replacement; SAVR, surgical aortic valve replacement.

Table 4. Demographic and clinical characteristics of aortic stenosis patients who completed the preference elicitation survey (N=219).

Characteristics	Patients, n (%)
Gender	
Male	91 (41.6%)
Female	128 (58.4%)
Ethnic background	
Not Hispanic or Latino	209 (95.4%)
Age group, years	
19–39	58 (26.5%)
40–59	74 (33.8%)
60–74	55 (25.1%)
75–89	29 (13.2%)
90+	3 (1.4%)
Racial background	
White	173 (79.0%)
Black or African American	27 (12.3%)
Asian	13 (5.9%)
Native Hawaiian or other Pacific Islander	3 (1.4%)
American Indian or Alaskan Native	4 (1.8%)
Other	7 (3.2%)
Living Situation	
Living alone	54 (24.7%)
Living with a partner or spouse, family, or friends	164 (74.9%)
Employment Status	
Employed, full-time	93 (42.5%)
Employed, part-time	28 (12.8%)
Homemaker	8 (3.7%)
Student	6 (2.7%)
Unemployed	8 (3.7%)
Retired	71 (32.4%)
Disabled	19 (8.7%)
Highest level of education completed	
Secondary/high school	16 (7.3%)
Some college	46 (21.0%)
College degree	93 (42.5%)
Postgraduate degree	63 (28.8%)

Characteristics	Patients, n (%)
Other:	1 (0.5%)
General health within past week	
Very good	55 (25.1%)
Good	85 (38.8%)
Fair	65 (29.7%)
Poor	13 (5.9%)
Very poor	1 (0.5%)
Previously treated for AS	
Yes	176 (80.4%)
No	43 (19.6%)
NYHA Classification	
No limitation of physical activity	78 (35.6%)
Slight limitation of physical activity	101 (46.1%)
Marked limitation in physical activity	40 (18.3%)
Severe limitation in physical activity	0 (0.0%)
Health literacy and numeracy	
Low literacy (1)	23 (10.5%)
Low numeracy (2)	22 (10.0%)
Overall low literacy / numeracy (3)	7 (3.2%)

¹Responses were scored between 0 (always) and 4 (never). Each participants' scored responses were averaged for a composite score ranging from 0–4. A low score if ≤ 2 .

²Participants were given one point for each correctly answered question (maximum numeracy score = 5). A low score if given is ≤ 2 answered incorrectly. ³Overall low: individuals who scored low on both educational level and objective health literacy. NYHA, New York Heart Association.

needing a new permanent pacemaker only increases by 3.3% with TAVR.

The standard errors in patients' MIR/MRB suggests a substantial heterogeneity in patients' responses to the ASW exercise (see [extended data, Appendix S2²⁴](#) for more detail). Some of this heterogeneity was associated with participants' age. MIR/MRB for three attributes—probability of having a new permanent pacemaker, probability of requiring dialysis, and one month mortality risk—were associated with whether patients are over or under 60 years old. Older patients were more willing to tolerate increases in risks/reductions in benefit to avoid having to undergo an invasive procedure.

No other correlation was found between participant characteristics and MIR/MRB. This includes whether a participant reported having previously undergone treatment for their AS. While this might be expected to influence preferences, the ability of the analysis to identify this influence is limited by

Table 5. Patients Maximum Acceptable Increase in Risk (MIR)/Minimum Acceptable Reduction in Benefit (MRB) in Exchange for Reducing Procedure Invasiveness from 'Invasive' to 'Minimally Invasive'.

Attribute	Impact of TAVR compared with SAVR [†]	Maximum Acceptable Increase in Risk/Maximum Acceptable Reduction in Benefit [Mean, (SD, n)]				Proportion of participants MIR/MRB > TAVR impact
		Whole sample	<60yrs old	≥60yrs old	p-value	
Mortality (1 Month)	-2.9%	3.86% (2.95, 109)	3.4 (2.7, 65)	4.5 (3.2, 44)	0.0521*	100%
Disabling Non-Fatal Stroke	-2.5%	6.69% (5.73, 110)	6.5 (5.6, 67)	7.0 (6.0, 43)	0.6524	100%
Independence	+23.0%	13.94% (11.77, 131)	13.5 (11.5, 80)	14.6 (12.2, 51)	0.5949	100%
New Permanent Pacemaker	+3.3%	6.98% (5.71, 132)	5.9 (5.1, 76)	8.4 (6.2, 56)	0.0121**	69.8%
Requirement for Dialysis	-1.5%	6.21% (5.62, 131)	5.4 (5.1, 81)	7.5 (6.2, 50)	0.0351**	100%
Proven to Work	-10yrs	17.42 years (16.86, 131)	17.9 (16.7, 80)	16.6 (17.2, 51)	0.6668	47.3%

[†]See Table 2 for source of data

* p<0.1, **p<0.05

the relatively small sample size and the small proportion of participants who reported not having previously received AS treatment (19.6%).

There were a large proportion of participants whose individual MIR/MRB was greater than the change in attribute performance achieved with TAVR (Table 5). For attributes where performance is better with TAVR compared to SAVR (mortality, disabling non-fatal stroke, independence, and dialysis), 100% of patients would prefer the improved performance and reduced invasiveness of TAVR. For the two attributes on which attribute performance is better with SAVR compared to TAVR, 70% of patients would be willing to accept the increased risk of needing a new permanent pacemaker, and 47% of participants would be willing to accept the shorter period for which TAVR had been proven to work in order to experience TAVR's reduced invasiveness.

The above analysis compares the performance of TAVR on each attribute separately. A complete comparison of TAVR and SAVR should do so across all attributes simultaneously and take into account observed heterogeneity (in this case across age groups). This objective is accomplished by means of the benefit-risk model (see Equation 2 and Equation 3). Figure 2 and Figure 3 show the incremental value of TAVR (overall and by attribute) observed among patients 60 years old or older (Figure 2) and among patients less than 60 years old (Figure 3). These figures reveal that, overall, TAVR is of greater value to patients than SAVR. Specifically, patients placed greater value on TAVR based on a lower short-term mortality rate, reduced procedural invasiveness, and a quicker time to return to normal quality of life (independence) offsetting the value patients placed on longer period over which the procedure has been proven to work and reduced risk of needing a pacemaker generated by

SAVR. Similar patterns were observed among younger and older patients.

TAVR threshold analysis

Table 6 reports a threshold analysis, which shows the minimum amount of benefit that patients would accept before preferring TAVR, or the maximum amount of risk that patients would tolerate before preferring TAVR. For example, given the incremental value that patients attach to TAVR (as reflected in Figure 2 and Figure 3) they would be willing to tolerate a mortality risk of 12.6% following TAVR before they would be indifferent between TAVR and SAVR.

The MCS shows that 79.5% of patients would prefer TAVR over SAVR. When the analysis is run separately for patients less than and greater than 60 years old, the proportions of patients preferring TAVR are 80.8% and 78.2% respectively. Removing patients who 'one-lined' in response to ASW exercises does not impact the results of the MCS, with 80.7% of patients still preferring TAVR.

Discussion

The choice between TAVR and SAVR for the treatment of AS involves making trade-offs between: procedure invasiveness; the period over which the procedure has been proven to be effective; mortality, stroke and independence outcomes; and risks such as the need for a new pacemaker or dialysis. This study elicited patients' preferences for AS procedure attributes to determine how they make these trade-offs, and thus whether they prefer TAVR or SAVR. Results suggest that, given the potential benefits and risks of TAVR and SAVR, on average, patients attach more value to TAVR, and the majority of patients would prefer TAVR. Patients placed a greater value on the lower invasiveness, quicker speed of recovery, and reduced risk of

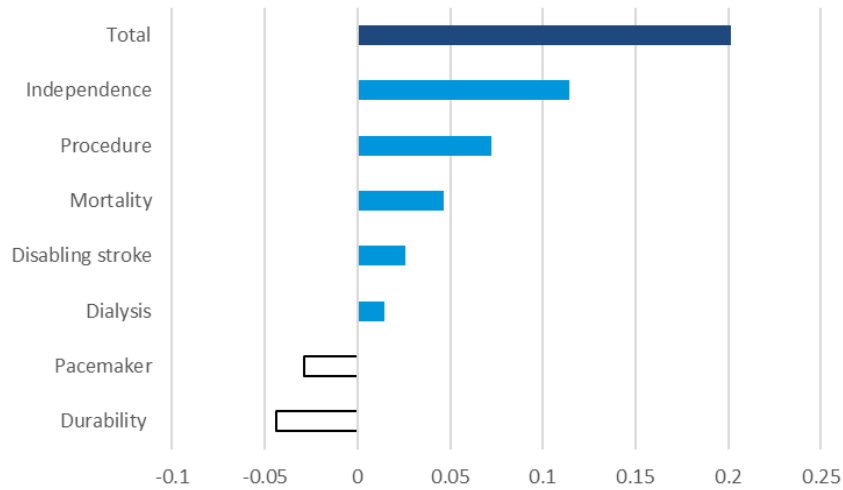


Figure 2. Incremental value of transcatheter aortic valve replacement vs. surgical aortic valve replacement in patients ≥ 60 years old.

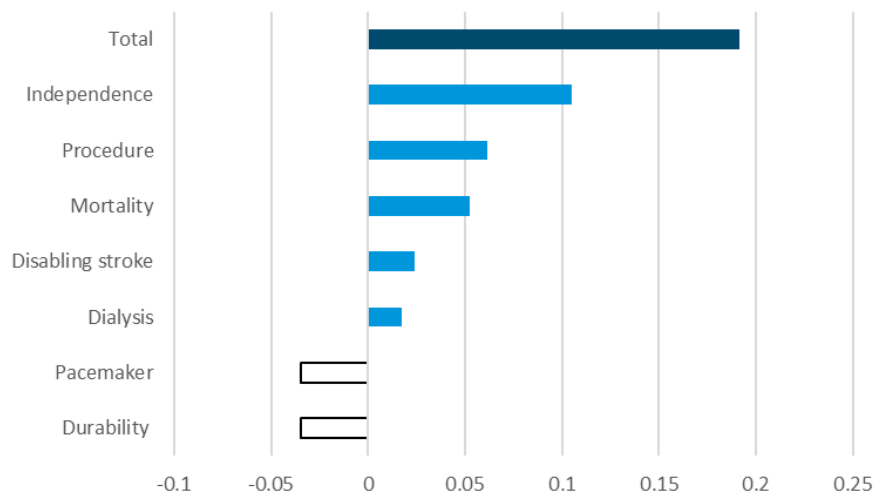


Figure 3. Incremental value transcatheter aortic valve replacement vs. surgical aortic valve replacement in patients <60 years old.

Table 6. Threshold Analysis.

	Minimum Acceptable Benefit/ Maximum Acceptable Risk			Outcome with TAVR
	Whole Population	<60 years old	≥ 60 years old	
Mortality (1 Month)	12.6%	11.6%	13.4%	1.1%
Disabling Non-Fatal Stroke	20.7%	21.0%	19.9%	0.8%
Independence	6.5%	6.0%	8%	47.9%
New Permanent Pacemaker	33.0%	30.6%	35.2%	12.3%
Requirement for Dialysis	21.6%	19.9%	23.7%	3.2%
Proven to Work	0yrs	0yrs	0yrs	10yrs

mortality, stroke and need for dialysis associated with TAVR than they did on longer period over which the procedure has been proven to work and reduced risk of needing a pace maker associated with SAVR.

Current guidelines from the American Heart Association for the procedural treatment of AS do not take into account recent clinical data supporting the use of TAVR²⁵. Based on the recent clinical results of TAVR and the findings of this study, regulators may reach different conclusions about the need to protect patients from risks historically associated with TAVR. For instance, TAVR may not result in the increased risk of stroke that regulators might expect it to, and patients may be willing to tolerate the greater need for a permanent pacemaker in order to experience the benefits of TAVR.

The BRA revealed substantial heterogeneity in patient preferences for AS treatment. Some preference heterogeneity is explained by patient age, with older patients being less willing to tolerate the invasiveness of SAVR, instead preferring to accept greater potential risks associated with other procedure attributes in order to reduce the invasiveness. However, preference heterogeneity raises concerns about the attendance of participants to the preference elicitation tasks. A small proportion (8.68%) of participants straight-lined on all questions. While this might indicate a lack of attendance, it may also capture strong preferences for/against the invasiveness of SAVR. Furthermore, all participants interpreted their response to the practice questions correctly, and only a small proportion of respondents demonstrated low health literacy or numeracy. This provides some reassurance that the preference heterogeneity observed in this study reflects a genuine difference in preference, rather than being the result of patients failing to complete the survey in a meaningful way.

Two other studies that used ASW to elicit patient preferences have been published^{26,27}. Both studies also observed substantial preference heterogeneity. One of these studies²⁷ provided evidence supporting the validity and reliability of the preference outputs, both by replicating the results of the ASW with a thresholding exercise, and by comparing participants' responses with their qualitative statements on the basis for their answers. This provides some reassurance about the validity of responses to the ASW exercise used for the current study. This may suggest that methods such as ASW, which elicit individual-level patient preferences, capture more preference heterogeneity than population-level methods, such as discrete choice experiments. Further work could usefully continue to validate the results of ASW exercises and test the hypothesis that individual-level preferences method captures greater heterogeneity in patient preferences.

Only one other study of AS patients' treatment preferences has been published to date²⁸. The study design was sufficiently different to the current study—focusing on patients' willingness to accept the mortality risk associated with interventions—that it is not possible to directly compare the results. However, the study by Hussain *et al.*²⁸ did reveal a higher risk tolerance among

patients with greater disease burden (defined as weekly incidence of restricting symptoms, perceived change in health compared with 1 year earlier, EQ-VAS scores, and the New York Heart Association (NYHA) classification). Our study failed to identify an association between patient preferences and NYHA classification, though this might be due to the limited sample size and the small proportion of the sample in the more severe stages of the NYHA classification. Further research could usefully gather data from a larger sample of AS patients to determine the association of preferences and patient characteristics, such as NYHA classification or whether patients have previously undergone treatment for AS.

While a majority of patients in the current study preferred TAVR, a number of patients (around 20%) preferred SAVR. This, and the underlying heterogeneity in patient preferences, support the need for a shared decision-making (SDM) tool that will help patients and surgeons choose procedures based on both clinical indications and patient risk tolerance. The Patient-Centered Outcomes Research Institute (PCORI) has developed a SDM tool to support patients choose between SAVR and TAVR¹⁵. However, this tool includes a narrower range of treatment attributes—stroke risk, mortality risk and discharge home—than those included in the analysis reported in this study. Furthermore, the tool does not include a component to elicit a patient's preferences.

As always, the conclusions of the study should be drawn in light of its limitations. First, it was not possible to engage patients in the design of the elicitation exercise. Instead, the patient voice was reflected in the design through a review of the limited preference research undertaken with AS patients to date and through the engagement that the experts who were consulted had themselves had with patients. Further work would usefully confirm with patients that no attributes were excluded from the study. Second, in the first survey round we relied on patient self-report of their AS diagnoses and severity, and in the second wave providing confirmation of AS diagnosis was voluntary. Third, the sample is healthier and younger than the population currently eligible for TAVR and SAVR²⁹. AS patients were recruited from the membership of Heart Valve Voice and Mended Hearts as well as through M3 Global Research, and it is possible that patients who are motivated to join these organizations may have different preferences than the broader population. Further AS patient preference research should replicate this study, including older patients with more severe disease burden. Fourth, it was necessary to assume that partial value functions were linear. The sample size and the number of attributes for which preferences were being collected meant collecting data on the shape of the partial value functions would have overburdened respondents.

Finally, the application of ASW in BRA is relatively novel and raises a number of questions. First, it relies on the idea that patients who are indifferent between adding one or another feature to a good or service would be willing to accept either improvement. The concept of indifference is commonly evoked

in methods to assess the value of changes in health outcome. Further, the baseline against which improvements are assessed is the worse level of performance of the treatments being evaluated and over 80% of patients had previously been treated for AS. Thus, in most cases it might be reasonable to assume that patients would accept the changes, as they represent outcomes better than those they've already accepted. However, to infer that they would be willing to accept improvements in outcomes assumes that they were well informed when they made these treatment choices. Further research should test participants willingness to accept changes in attributes.

Second, unlike some other elicitation methods, such as discrete choice experiments, the study assumes that respondents don't make mistakes when they are giving their answers, which is unlikely if respondents are near to their indifference points. It is thus necessary to assume that any such mistakes average out when assessing preferences for the study sample.

Third, as all respondents start their evaluation of attributes at the same point, ASW may be subject to anchoring. To mitigate this risk, 'worst' levels in the choice tasks were defined as the 'worst' clinically relevant level. However, in some instances this may have introduced a ceiling effect, with insufficient improvement in the attribute available to identify the indifference point. In these instances, it was necessary to increase the 'worst' level beyond that experienced by patients. Further work could usefully explore this apparent tension between anchoring and ceiling effects inherent within ASW.

Conclusions

Most AS patients are willing to tolerate sizable increases in clinical risk in exchange for the benefits associated with TAVR. A BRA incorporating data from patients' preferences for the attributes of AS treatments revealed a strong preference for TAVR compared to SAVR. The analysis also revealed substantial heterogeneity in individual patient preferences, partly associated with patient age. Further work is required to understand this heterogeneity, and whether additional patient

characteristics such as NYHA class are associated with different preferences. In the meantime, SDM tools should incorporate the factors identified in this model to assist patients and clinicians in achieving a more patient-centered treatment decision.

Data availability

Underlying data

Open Science Framework: AS patient preference data. DOI [10.17605/OSF.IO/UGD8X](https://doi.org/10.17605/OSF.IO/UGD8X)²⁴.

This project contains the following underlying data:

- TAVR Manuscript_Full Dataset_Updated.xlsx (full dataset of patient responses including codebook).

Extended data

Open Science Framework: AS patient preference data. DOI [10.17605/OSF.IO/UGD8X](https://doi.org/10.17605/OSF.IO/UGD8X)²⁴.

This project contains the following extended data:

- TAVR Manuscript Appendices S1_S2_Updated.docx (contains S1 Appendix, including S1 Table 1 and legends for S1 Figures 1-6; and S2 Appendix, including legends for S2 Figures 1-6).
- TAVR Manuscript Figures_Updated.docx (S1 Figures 1-6 and S2 Figures 1-6).
- TAVR Survey Contents_Updated.docx (a copy of the questionnaire given to each participant).

Data are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](https://creativecommons.org/licenses/by/4.0/) (CC0 1.0 Public domain dedication).

Acknowledgements

The authors would like to thank Heart Valve Voice and Mended Hearts for their support with the design of the survey, and their support recruiting patients.

References

1. Perera S, Wijesinghe N, Ly E, *et al.*: **Outcomes of patients with untreated severe aortic stenosis in real-world practice.** *N Z Med J.* 2011; **124**(1345): 40-8.
[PubMed Abstract](#)
2. Bach DS, Radeva JI, Birnbaum HG, *et al.*: **Prevalence, referral patterns, testing, and surgery in aortic valve disease: leaving women and elderly patients behind?** *J Heart Valve Dis.* 2007; **16**(4): 362-9.
[PubMed Abstract](#)
3. US Census Bureau: **Updated statistic from 2015- (Population Division, June 2015).** 2015.
4. Nkomo VT, Gardin JM, Skelton TN, *et al.*: **Burden of valvular heart diseases: a population-based study.** *Lancet.* 2006; **368**(9540): 1005-11.
[PubMed Abstract](#) | [Publisher Full Text](#)
5. Thaden JJ, Nkomo VT, Enriquez-Sarano M: **The global burden of aortic stenosis.** *Prog Cardiovasc Dis.* 2014; **56**(6): 565-71.
[PubMed Abstract](#) | [Publisher Full Text](#)
6. Bach DS: **Prevalence and characteristics of unoperated patients with severe aortic stenosis.** *J Heart Valve Dis.* 2011; **20**(3): 284-91.
[PubMed Abstract](#)
7. Ross J Jr, Braunwald E: **Aortic stenosis.** *Circulation.* 1968; **38**(1 Suppl): 61-7.
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Baron SJ: **Valve Disease.** 2017.
[Reference Source](#)
9. Leon MB, Smith CR, Mack M, *et al.*: **Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery.** *N Engl J Med.* 2010; **363**(17): 1597-607.
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Baron SJ, Arnold SV, Wang K, *et al.*: **Health Status Benefits of Transcatheter vs Surgical Aortic Valve Replacement in Patients With Severe Aortic Stenosis at Intermediate Surgical Risk: Results From the PARTNER 2 Randomized Clinical Trial.** *JAMA Cardiol.* 2017; **2**(8): 837-45.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

11. Gargiulo G, Sannino A, Capodanno D, *et al.*: **Transcatheter Aortic Valve Implantation Versus Surgical Aortic Valve Replacement: A Systematic Review and Meta-analysis.** *Ann Intern Med.* 2016; **165**(5): 334–44.
[PubMed Abstract](#) | [Publisher Full Text](#)
12. Siontis GC, Praz F, Pilgrim T, *et al.*: **Transcatheter aortic valve implantation vs. surgical aortic valve replacement for treatment of severe aortic stenosis: a meta-analysis of randomized trials.** *Eur Heart J.* 2016; **37**(47): 3503–12.
[PubMed Abstract](#) | [Publisher Full Text](#)
13. Yu PJ, Mattia A, Cassiere HA, *et al.*: **Should high risk patients with concomitant severe aortic stenosis and mitral valve disease undergo double valve surgery in the TAVR era?** *J Cardiothorac Surg.* 2017; **12**(1): 123.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Siemieniuk RA, Agoritsas T, Manja V, *et al.*: **Transcatheter versus surgical aortic valve replacement in patients with severe aortic stenosis at low and intermediate risk: systematic review and meta-analysis.** *BMJ.* 2016; **354**: i5130.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Patient-Centered Outcomes Research Institute (PCORI): **Optimizing Health Outcomes in Patients with Symptomatic Aortic Valve Disease.** [Reference Source](#)
16. Marsh K, Ijzerman M, Thokala P, *et al.*: **Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force.** *Value Health.* 2016; **19**(2): 125–37.
[PubMed Abstract](#) | [Publisher Full Text](#)
17. Mishra RK, Yang W, Roy J, *et al.*: **Kansas City Cardiomyopathy Questionnaire Score Is Associated With Incident Heart Failure Hospitalization in Patients With Chronic Kidney Disease Without Previously Diagnosed Heart Failure: Chronic Renal Insufficiency Cohort Study.** *Circ Heart Fail.* 2015; **8**(4): 702–8.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Kapadia S, Agarwal S, Miller DC, *et al.*: **Insights Into Timing, Risk Factors, and Outcomes of Stroke and Transient Ischemic Attack After Transcatheter Aortic Valve Replacement in the PARTNER Trial (Placement of Aortic Transcatheter Valves).** *Circ Cardiovasc Interv.* 2016; **9**(9): pii: e002981.
[PubMed Abstract](#) | [Publisher Full Text](#)
19. Thourani VH, Kodali S, Makkar RR, *et al.*: **Transcatheter aortic valve replacement versus surgical valve replacement in intermediate-risk patients: a propensity score analysis.** *Lancet.* 2016; **387**(10034): 2218–25.
[PubMed Abstract](#) | [Publisher Full Text](#)
20. Leon MB, Smith CR, Mack MJ, *et al.*: **Transcatheter or Surgical Aortic-Valve Replacement in Intermediate-Risk Patients.** *N Engl J Med.* 2016; **374**(17): 1609–20.
[PubMed Abstract](#) | [Publisher Full Text](#)
21. Green CP, Porter CB, Bresnahan DR, *et al.*: **Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure.** *J Am Coll Cardiol.* 2000; **35**(5): 1245–55.
[PubMed Abstract](#) | [Publisher Full Text](#)
22. Marchini J: **Lecture 6: The Normal Distribution: Department of Statistics.** University of Oxford.
23. Postmus D, Mavris M, Hillege HL, *et al.*: **Incorporating patient preferences into drug development and regulatory decision making: Results from a quantitative pilot study with cancer patients, carers, and regulators.** *Clin Pharmacol Ther.* 2016; **99**(5): 548–54.
[PubMed Abstract](#) | [Publisher Full Text](#)
24. Marsh K: **AS patient preference data.** 2020.
<http://www.doi.org/10.17605/OSF.IO/UGD8X>
25. Nishimura RA, Otto CM, Bonow RO, *et al.*: **2017 AHA/ACC Focused Update of the 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 Clinical Practice Guidelines.** *Circulation.* 2017; **135**(25): e1159–e95.
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Postmus D, Richard S, Bere N, *et al.*: **Individual Trade-Offs Between Possible Benefits and Risks of Cancer Treatments: Results from a Stated Preference Study with Patients with Multiple Myeloma.** *Oncologist.* 2018; **23**(1): 44–51.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Sri Bhashyam S, Quartel A, Gershman A, *et al.*: editors. **PSY202 - The occupational hazards of measuring risk tolerance: Convergent validity in preference elicitation.** *ISPOR European Conference; Barcelona, Spain. Value Health.* 2018; **21**(Supplement 3): S470.
[Publisher Full Text](#)
28. Hussain AI, Garratt AM, Beines JO, *et al.*: **Validity of standard gamble utilities in patients referred for aortic valve replacement.** *Qual Life Res.* 2016; **25**(7): 1703–12.
[PubMed Abstract](#) | [Publisher Full Text](#)
29. Otto CM, Kumbhani DJ, Alexander KP, *et al.*: **2017 ACC Expert Consensus Decision Pathway for Transcatheter Aortic Valve Replacement in the Management of Adults With Aortic Stenosis: A Report of the American College of Cardiology Task Force on Clinical Expert Consensus Documents.** *J Am Coll Cardiol.* 2017; **69**(10): 1313–46.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:   

Version 5

Reviewer Report 16 February 2021

<https://doi.org/10.5256/f1000research.54334.r79093>

© 2021 Gonzalez J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Juan Marcos Gonzalez 

Department of Medicine, Duke University School of Medicine, Durham, NC, USA

I thank the authors for addressing my previous concerns. The changes and clarifications offered have addressed my comments satisfactorily.

My only remaining comment is in response to the authors' question about the potential insufficiency of equivalences to measure unbiased marginal rates of substitution. There is certainly some literature in DCEs showing that inappropriate use of a forced-choice format can lead to variation in parameter estimates (Determann et al., 2019¹; and Campbell and Erdem, 2019²). While the need for an opt-out or status-quo alternative depends on the study objectives and the relevant decision context, there is also evidence that marginal rates of substitution change with the inclusion of cost in a forced-choice format (Pedersen et al., 2011³).

References

1. Determann D, Gyrd-Hansen D, de Wit GA, de Bekker-Grob EW, et al.: Designing Unforced Choice Experiments to Inform Health Care Decision Making: Implications of Using Opt-Out, Neither, or Status Quo Alternatives in Discrete Choice Experiments. *Med Decis Making*. **39** (6): 681-692 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Campbell D, Erdem S: Including Opt-Out Options in Discrete Choice Experiments: Issues to Consider. *Patient*. **12** (1): 1-14 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Pedersen L, Kjær T, Kragstrup J, Gyrd-Hansen D: Does the Inclusion of a Cost Attribute in Forced and Unforced Choices Matter?. *Journal of Choice Modelling*. 2011; **4** (3): 88-109 [Publisher Full Text](#)

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 4

Reviewer Report 02 November 2020

<https://doi.org/10.5256/f1000research.25257.r63117>

© 2020 Gonzalez J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Juan Marcos Gonzalez 

Department of Medicine, Duke University School of Medicine, Durham, NC, USA

Thank you for the opportunity to review this interesting manuscript. I commend the authors for looking at such a novel method for the elicitation of risk tolerance with stated-preference methods. The manuscript is well-written and the explanation of the method is clear. I do have some major concerns that I think should be addressed by the authors. I summarized my concerns below.

My biggest concern is the number of assumptions implicit in the derivation of the MAR and MAB measures with these data. In particular, the operating assumption that value equivalence in the context of the decisions elicited implies acceptability. I would agree with that premise if the question presented the right context, but here things seem a bit murky. Does the fact that I am indifferent between adding any one of two features to a good or service imply that I'm willing to accept either improvement? Not necessarily. While this assumption is not necessarily wrong, I think it needs to be better justified either with evidence from their qualitative/pilot work or through previous literature.

Almost equally problematic is the omission of measurement error in the analysis of these data. Respondents answered three questions with each pairwise comparison. Depending on the step procedure to change the specifics of the comparisons, there could be a systematic error structure for the measures. For example, a 1-unit improvement across questions could systematically fall short of a threshold among respondents who truly require low probabilities of adverse events to achieve equivalence. Conversely, an initial full swing in the relevant range followed by a gradual increase in the risk levels could end up overestimating the equivalences. I could not find any details on the algorithm followed by the authors as they tried to identify an indifference threshold with only three questions per comparison. That information needs to be included or highlighted.

Even if the respondent reached an equivalence threshold with the questions in the survey, there is still a measurement error for at least three reasons:

1. The current analysis assumes that responses are collected without errors and that respondents don't make mistakes as they give their answers. This is particularly dubious when respondents are near indifference and the attributes are relatively important (i.e., the equivalence ratio is small).
2. The study design ignores anchoring problems as all respondents started their evaluation of attribute equivalence at the same point. It is well known that anchoring problems can bias measures of preferences. Potentially, this would not be problematic if the baseline scenario was clinically relevant—like the current standard of care—but the principle for the baseline

levels seems to be unrelated to clinical considerations.

3. The range established a priori for the swings may not be sufficient to achieve indifference for some attributes. This would be considered a ceiling effect and could overestimate the importance weight of some attributes.

Without some sense of the measurement error of the instrument, either through theory (i.e., random utility theory) or empirical data (i.e., as done with PROMs), we would seem to be blind about the validity of the study results. These measurement issues need to be better addressed or included as limitations of the study.

Another important implicit assumption of the method is that of linearity or proportionality of weights within the attribute range. There is a lot of evidence that this assumption often does not hold, particularly with low risks of events. The assumption would seem to be particularly crucial here because the evaluation of the importance is not done in the direction (and potentially the range) of the relevant risk changes from a regulatory standpoint. That is, the authors evaluate the impact of reducing the risks from a high absolute level and then move the approximated equivalences to a different point within the studied range to predict the importance of increases in risk levels.

If data comes from the same clinically relevant range the authors could at least argue that the average effects within the range are recoverable, but if evaluating a completely different risk range, there is no guarantee that the linearity assumption is adequate even for the average effect.

Consider the following *extreme* example, suppose we were interested in the importance of increasing the risk of death within the range of 0% to 10%. Collecting preference information over that range, even if we imposed linearity or proportionality of preferences, would at least capture the correct average effect within that increase—assuming we use the right analysis tools. However, if instead, we approximate that information by looking at the importance of reducing the risk of death from 100% to 90%, there is no guarantee that this importance would be an accurate approximation of preferences for 0% to 10%. The only way the two measures would be identical is if the linearity assumption imposed by the authors holds for the full range of risk levels, but the authors don't test this to justify such an important—and potentially flawed—assumption.

Overall, I think this study results could offer a reasonable approximation of risk tolerance with adaptive swing weighting, but the validity of such approximation is unclear without further evaluation of the implicit assumptions made to derive MAR and MAB.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Health preference research and the use of these preference measures to infer risk tolerance.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 29 Jan 2021

Kevin Marsh, Evidera Inc, London, UK

Reviewer Feedback 1

My biggest concern is the number of assumptions implicit in the derivation of the MAR and MAB measures with these data. In particular, the operating assumption that value equivalence in the context of the decisions elicited implies acceptability. I would agree with that premise if the question presented the right context, but here things seem a bit murky. Does the fact that I am indifferent between adding any one of two features to a good or service imply that I'm willing to accept either improvement? Not necessarily. While this assumption is not necessarily wrong, I think it needs to be better justified either with evidence from their qualitative/pilot work or through previous literature.

Authors' Response 1

Thanks for this observation. This is an interesting question. Given the centrality of the concept of indifference in defining utility in economic theory, the associated notion of the marginal rate of substitution, and employment of the concept of indifference in various other preference elicitation methods – not least TTO, we had not questioned the relationship between indifference and willingness to accept. We have reference the above in justifying the approach. Though we'd be interested in the reviewers thoughts on whether there is a literature which challenges this notion that we are not familiar with.

Reviewer Feedback 2

Almost equally problematic is the omission of measurement error in the analysis of these data. Respondents answered three questions with each pairwise comparison. Depending on the step procedure to change the specifics of the comparisons, there could be a systematic

error structure for the measures. For example, a 1-unit improvement across questions could systematically fall short of a threshold among respondents who truly require low probabilities of adverse events to achieve equivalence. Conversely, an initial full swing in the relevant range followed by a gradual increase in the risk levels could end up overestimating the equivalences. I could not find any details on the algorithm followed by the authors as they tried to identify an indifference threshold with only three questions per comparison. That information needs to be included or highlighted.

Authors' Response 2

The algorithm used is reported in Extended Data, Appendix 1. We did not adopt either of the approaches that cause the reviewer some concern. Rather, each iteration of the level of performance cut the remaining space where the indifferent point was located in half.

Reviewer Feedback 3

Potential measurement error because: The current analysis assumes that responses are collected without errors and that respondents don't make mistakes as they give their answers. This is particularly dubious when respondents are near indifference and the attributes are relatively important (i.e., the equivalence ratio is small).

Authors' Response 3

We agree this is a key difference between swing weighting and methods like discrete choice experiments. We have added a note to this effect in the paragraph on limitations in the discussion section.

Reviewer Feedback 4

Potential measurement error because: The study design ignores anchoring problems as all respondents started their evaluation of attribute equivalence at the same point. It is well known that anchoring problems can bias measures of preferences. Potentially, this would not be problematic if the baseline scenario was clinically relevant—like the current standard of care—but the principle for the baseline levels seems to be unrelated to clinical considerations.

Authors' Response 4

We agree with the reviewers observations about the importance of anchoring. Where possible the 'worst' level in the choice tasks were defined as the worst level in the procedure choice to which the preference estimates were applied. However, in some instance the pilot suggested that this would not give us enough 'runway' to identify indifference – per the following comments, we risked ceiling effects if we did this. In those cases, the 'worst' level needed to be increased compared with that experienced by patients to create the 'runway'. We agree this might be a source of bias and have added a note to this effect in the discussion section.

Reviewer Feedback 5

Potential measurement error because: The range established a priori for the swings may not be sufficient to achieve indifference for some attributes. This would be considered a ceiling effect and could overestimate the importance weight of some attributes

Authors' Response 5

As noted in response to the previous comment, the range over which swings were explored was sometimes widened, precisely to try to avoid such ceiling effects. Nevertheless, the bi-modal distribution observed points to the potential that such effects may not have been entirely eradicated. Also noted above, in such designs it is possible that there is a tension between efforts to avoid anchoring and efforts to avoid ceiling effects. We have added a note on this in the discussion section.

Reviewer Feedback 6

Another important implicit assumption of the method is that of linearity or proportionality of weights within the attribute range. There is a lot of evidence that this assumption often does not hold, particularly with low risks of events. The assumption would seem to be particularly crucial here because the evaluation of the importance is not done in the direction (and potentially the range) of the relevant risk changes from a regulatory standpoint. That is, the authors evaluate the impact of reducing the risks from a high absolute level and then move the approximated equivalences to a different point within the studied range to predict the importance of increases in risk levels. If data comes from the same clinically relevant range the authors could at least argue that the average effects within the range are recoverable, but if evaluating a completely different risk range, there is no guarantee that the linearity assumption is adequate even for the average effect.

Authors' Response 6

We agree with this limitation. We had included this in the discussion of the limitations of the approach. We've re-emphasized this challenge in the discussion section, and call it out as a potential limit of the ASW method when working with more than a few attributes. The method was adopted as the goal of the study was to collect of data on preferences for many attributes with small sample sizes. However, in this context, it was not possible to also test the non-linearity of value functions without overburdening participants.

Competing Interests: No competing interests were disclosed.

Version 2

Reviewer Report 29 May 2019

<https://doi.org/10.5256/f1000research.21135.r48437>

© 2019 Jaffe R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Ross Jaffe**

Versant Ventures, San Francisco, CA, USA

The authors have addressed my major concerns, and I support indexing of this article.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Internal medicine physician with prior experience in outcomes research and shared decision making. Not actively involved in research in my present role as a venture capitalist, but I led the MDIC Patient Centered Benefit-Risk Project that developed a framework for incorporating patient preferences into FDA medical device regulation and developed the first catalog of patient preference methodologies included as an appendix to that report.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 21 May 2019

<https://doi.org/10.5256/f1000research.21135.r48436>

© 2019 Coylewright M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Megan Coylewright

Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA

The changes provided needed updates. No major changes to methods or conclusions.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: I have received honoraria from Edwards LifeSciences.

Reviewer Expertise: shared decision making

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 07 May 2019

<https://doi.org/10.5256/f1000research.20591.r46834>

© 2019 Coylewright M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Megan Coylewright

Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA

The co-authors have an ambitious agenda to identify patient preferences regarding the differing attributes for the treatment of aortic stenosis. They are to be commended for tackling an

important topic, and employing known methodologic standards for addressing the question.

They deploy a methodology of adapted swing weighting (ASW), identifying which treatment attributes patients would tolerate to move from an invasive (surgical aortic valve replacement, or SAVR) option to a minimally invasive (transcatheter aortic valve replacement, or TAVR) option. A set of pairwise comparisons were made between "less invasive" and other features of AVR, with 5 different iterations presented.

The study is written by employees of one of the TAVR companies (Edwards LifeSciences), as well as employees of a company that specializes in patient preference work, paid for by Edwards. It is not apparent that there was engagement of other stakeholders in the study design, such as patients, families, clinicians, administrators, policy makers, etc. There were 5 patients with the disease condition that trialed the exercises before the start of the study to ensure understandability. There is no information how these patients were recruited, or what changes were made in response to their feedback.

The authors' conclusions are that patients are willing to tolerate big increases in risk and exchange for benefits of a minimally invasive procedure.

Importance of the research question: Given the recent reconsideration of the TAVR national coverage determination, data as needed to inform interventions designed to increase the degree to which patient preferences are included in final decision making. This study suggests that there is heterogeneity in patient preferences and that patients may tolerate larger differences in risk than physicians and regulators traditionally have deemed acceptable when faced with a minimally invasive procedure.

Originality of the work: There is not a large amount of data on the preferences of patients for TAVR.

Strengths:

Identifying patient preferences and how they influence decision-making is an important topic. Pair-wise comparisons can be helpful to delineate tolerance of risk in one domain versus another. In the decision aid that was presented to the patient's demonstrating differences between the 2 therapies, the same picture was used for both surgery and transcatheter procedure which did potentially reduce bias.

Weaknesses:

The largest weakness is the fact that the attributes were not selected by patients known to be affected by the disease condition; a survey was asked of participants in advocacy groups online. The eligible patients were created from the membership of heart valve voice and mended hearts. They were directed to an online screening tool where patient reported disease conditions were recorded. The study sample was largely white and highly educated. No patients had severe limitations to physical activity, which has been shown to significantly impact decision-making. This has been a consistent weakness in study design for clinical trials as well. The patients selected for the survey, in addition, were not necessarily patients that would be considering this decision.

Data analysis:

Responses from the survey were used to compute the maximum acceptable increase in risk or maximum acceptable reduction in benefits that participants would be willing to tolerate in exchange for being involved in a less invasive procedure.

Interpretation of data:

93 patients completed the survey. Patients were not representative of the TAVR population and described above in the weaknesses. The authors importantly describe substantial heterogeneity and patient's responses to the exercise. Importantly, this correlated with patient age, which was lower in the sample than the TAVR population. Patients were more likely to prefer TAVR when considering benefits of independence or risks of mortality or disabling stroke. They were more likely to favor surgical AVR when considering risks of pacemaker and uncertainty regarding durability.

Presentation of data and story:

Abstract: The abstract describes that the purpose of the research is to determine which outcomes associated with TAVR and S AVR patients consider most important. Patients were not included in the selection of potential outcomes associated with the therapies. The abstract is stand-alone, describing the findings of trade-offs for risks and benefits. However, there should be a further statement about whether or not the patients had aortic stenosis and whether they were symptomatic as this is known to impact decision-making.

Introduction: The introduction would need to be revised regarding the literature around pacemaker and stroke. Otherwise, the introduction lays out the rationale and explains the goals for the study.

Methods: The methods provide sufficient detail that experiments could be repeated however, much of the methods could be relegated to a supplement online.

Results: The results are hypothesis generating but are limited in the reliability given uncertainty regarding patient diagnosis, severity of AS and symptoms status. In addition, the attributes that the patients are commenting on are defined by the research team, with leadership from the company, and therefore have potential bias.

The findings are clinically relevant as they describe differences and patient preferences around risk and benefit. They highlight the need for shared decision making approach. They also raise the issue of which outcomes are most important to patients. Additional research is needed to identify that from patients with symptoms themselves.

Discussion: The discussion around the relationship between the guidelines and regulators needs editing. It is not clear on what the authors mean when they say that the current guidelines "do not take into account recent clinical data supporting the use of TAVR". There is an important point about how patients willingness to tolerate certain risks may differ from policy makers and physicians and an argument can be made to include patient preferences into final decision making.

Limitations described include how well the participants responding to the survey paid attention to the questions. This is described as a reason for some of the heterogeneity. Limitation would be the small size of the patient's sample and those patients not having the disease that is being studied nor being symptomatic. The limitations focus primarily on the selection of ASW exercises as a way to assess patient preferences rather than the patient population.

The tables and figures do stand-alone. The references appear appropriate. I appreciate the opportunity to review this interesting manuscript.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: I have received honoraria from Edwards LifeSciences for speaking on shared decision making.

Reviewer Expertise: shared decision making

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 08 May 2019

Kevin Marsh, Evidera Inc, London, UK

Review 2: Meghan Coylewright

The co-authors have an ambitious agenda to identify patient preferences regarding the differing attributes for the treatment of aortic stenosis. They are to be commended for tackling an important topic, and employing known methodologic standards for addressing the question.

They deploy a methodology of adapted swing weighting (ASW), identifying which treatment attributes patients would tolerate to move from an invasive (surgical aortic valve replacement, or SAVR) option to a minimally invasive (transcatheter aortic valve replacement, or TAVR) option. A set of pairwise comparisons were made between "less invasive" and other features of AVR, with 5 different iterations presented.

The study is written by employees of one of the TAVR companies (Edwards LifeSciences), as well as employees of a company that specializes in patient preference work, paid for by Edwards. It is not apparent that there was engagement of other stakeholders in the study design, such as patients, families, clinicians, administrators, policy makers, etc. There were 5 patients with the disease condition that trialed the exercises before the start of the study to ensure understandability. There is no information how these patients were recruited, or what changes were made in response to their feedback.

The authors' conclusions are that patients are willing to tolerate big increases in risk and exchange for benefits of a minimally invasive procedure.

Importance of the research question: Given the recent reconsideration of the TAVR national coverage determination, data as needed to inform interventions designed to increase the degree to which patient preferences are included in final decision making. This study suggests that there is heterogeneity in patient preferences and that patients may tolerate larger differences in risk than physicians and regulators traditionally have deemed acceptable when faced with a minimally invasive procedure.

Originality of the work: There is not a large amount of data on the preferences of patients for TAVR.

Strengths:

Identifying patient preferences and how they influence decision-making is an important topic. Pair-wise comparisons can be helpful to delineate tolerance of risk in one domain versus another.

In the decision aid that was presented to the patient's demonstrating differences between the 2 therapies, the same picture was used for both surgery and transcatheter procedure which did potentially reduce bias.

Weaknesses:

The largest weakness is the fact that the attributes were not selected by patients known to be affected by the disease condition; a survey was asked of participants in advocacy groups online. The eligible patients were created from the membership of heart valve voice and mended hearts. They were directed to an online screening tool where patient reported disease conditions were recorded. The study sample was largely white and highly educated. No patients had severe limitations to physical activity, which has been shown to significantly impact decision-making. This has been a consistent weakness in study design for clinical trials as well. The patients selected for the survey, in addition, were not necessarily patients that would be considering this decision.

Response: We thank the reviewer for their interest in the study and their comments. We agree with concerns about the representative of the sample. The observation of the limitation of the study has been extended in the discussion section to read:

"Finally, the current study is based on a relatively small sample of patients, and the sample is healthier and younger than the population currently eligible for TAVR and SAVR. AS patients were recruited from the membership of Heart Value Voice and Mended Hearts, and it is possible that patients who are motivated to join these organizations may have different preferences than the broader population. Further AS patient preference research should replicate this study in a larger sample of patients, including more patients with more severe disease burden".

Data analysis:

Responses from the survey were used to compute the maximum acceptable increase in risk or maximum acceptable reduction in benefits that participants would be willing to tolerate in exchange for being involved in a less invasive procedure.

Interpretation of data:

93 patients completed the survey. Patients were not representative of the TAVR population and described above in the weaknesses. The authors importantly describe substantial heterogeneity and patient's responses to the exercise. Importantly, this correlated with patient age, which was lower in the sample than the TAVR population. Patients were more likely to prefer TAVR when considering benefits of independence or risks of mortality or disabling stroke. They were more likely to favor surgical AVR when considering risks of pacemaker and uncertainty regarding durability.

Presentation of data and story:

Abstract: The abstract describes that the purpose of the research is to determine which outcomes associated with TAVR and S AVR patients consider most important. Patients were not included in the selection of potential outcomes associated with the therapies. The abstract is stand-alone, describing the findings of trade-offs for risks and benefits. However, there should be a further statement about whether or not the patients had aortic stenosis and whether they were symptomatic as this is known to impact decision-making.

Response: *The following statement has been added to the abstract: All patients self-reported as being diagnosed with AS, and as either having received treatment for their AS in the previous 10 years or as experiencing limitations in their physical activity as a result of their AS.*

Introduction: The introduction would need to be revised regarding the literature around pacemaker and stroke. Otherwise, the introduction lays out the rationale and explains the goals for the study.

Response: *We have not updated the introduction. Perhaps the reviewer could provide more detail on the literature that should be summarized in the introduction.*

Methods: The methods provide sufficient detail that experiments could be repeated however, much of the methods could be relegated to a supplement online.

Response: *We have not moved any content to a supplementary appendix. We would be happy to consider doing so if the editors thought that was more appropriate given the audience for the journal.*

Results: The results are hypothesis generating but are limited in the reliability given uncertainty regarding patient diagnosis, severity of AS and symptoms status. In addition, the attributes that the patients are commenting on are defined by the research team, with leadership from the company, and therefore have potential bias.

Response: *We agree that it would have been preferable to have been able to engage patients more in the design of the study. The following statement has been added to the discussion section:*

"it was not possible to engage patients in the design of the elicitation exercise. Instead,

the patient voice was reflected in the design through a review of the limited preference research undertaken with AS patients to date and through the engagement that the experts who were consulted had themselves had with patients. Further work would usefully confirm with patients that no attributes were excluded from the study”.

We have also added to the paragraph in the discussion on limitations the fact that we relied on patient self report of their diagnoses and AS severity.

The findings are clinically relevant as they describe differences and patient preferences around risk and benefit. They highlight the need for shared decision making approach. They also raise the issue of which outcomes are most important to patients. Additional research is needed to identify that from patients with symptoms themselves.

Response: We agree this would improve the research. As above, a note to this effect has been added to the discussion.

Discussion: The discussion around the relationship between the guidelines and regulators needs editing. It is not clear on what the authors mean when they say that the current guidelines "do not take into account recent clinical data supporting the use of TAVR". There is an important point about how patients willingness to tolerate certain risks may differ from policy makers and physicians and an argument can be made to include patient preferences into final decision making. Limitations described include how well the participants responding to the survey paid attention to the questions. This is described as a reason for some of the heterogeneity. Limitation would be the small size of the patient's sample and those patients not having the disease that is being studied nor being symptomatic. The limitations focus primarily on the selection of ASW exercises as a way to assess patient preferences rather than the patient population.

Response: See earlier responses. The discussion of the limitations of the study has been extended to cover these concerns.

Competing Interests: No competing interests were disclosed.

Reviewer Report 25 April 2019

<https://doi.org/10.5256/f1000research.20591.r46833>

© 2019 Jaffe R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

 **Ross Jaffe**

Versant Ventures, San Francisco, CA, USA

This article by Marsh et al assesses patient preferences regarding the important medical decision about whether to undergo traditional surgical aortic valve replacement (SAVR) versus the newer transcatheter aortic valve replacement (TAVR) for aortic stenosis. This is an important effort to understand what attributes of benefit and risk from aortic valve replacement are most important to patients and understand how patients tradeoff those attributes in deciding which therapy to choose. With shared decision-making becoming increasingly used as our US healthcare system evolves to be more patient-centric, studies such as this one are important to expanding our understanding of how best to provide the information needed by patients to make informed decisions about their care. This study confirms a belief that many clinicians have that most patients prefer TAVR to SAVR because it is less invasive, providing evidence that about 75% patients prefer the less invasive approach. These patients are willing to tolerate a somewhat higher risk of stroke, pacemaker placement, and dialysis, and less evidence of long-term duration for the benefits of less invasiveness. It also identifies that younger patients (<60 yo) perspectives may differ from those of older patients (>60 yo) somewhat, although both groups generally prefer the less invasive procedure.

I support the publication of this study. While this study is not perfect from my point of view, it is an important contribution to the literature, both about how to treat aortic stenosis and about patient preference assessment more generally. Understanding what treatment attributes are most important to patients and how patients trade off benefits and risks is important to understanding how best to inform patients about their treatment options and help them make decisions that best reflect their individual preferences. As the paper notes on p. 11, there has only been one other, more limited study of patient preferences in aortic stenosis, so this study is an important expansion of our understanding of patient preferences that clinical area.

Additionally, from a methodological perspective and as also noted in the paper, there are only a few other studies that use adapted swing weights (ASW) to assess patient preferences. Swing weighting is one of only fourteen methods identified in the Medical Device Innovation Consortium (MDIC) review of preference assessment methodologies. (See: MDIC Patient Centered Benefit Risk Project Report, Appendix A: Catalog of Methods for Assessing Patient Preferences for Benefits and Harms of Medical Technologies, May 2015, available at: <https://mdic.org/resource/patient-centered-benefit-risk-pcbr-framework/>). This study is an important contribution to the literature about swing weighting methods, and should allow comparison to studies of patient preference using other methods to help researchers and clinicians better understand how best to assess patient preferences.

This study does have a few issues that should be highlighted to help put the results in context. I also note an area for further assessment of the data as well as areas for future research focus. Please note that I come at this study as someone with an interest in patient preferences from clinical and regulatory policy perspective, and do not have the expertise to comment on the specific methodology or statistical analysis, which I will leave to experts in those areas.

Major concerns:

1. The mixing of treated and untreated patients in the participant population: In Box 1 on p. 6, the inclusion criteria describe that patients in the study could have already had a procedure within the last 10 years or could be untreated. From a shared decision making point of view, patients

express their preferences prior to treatment – so it is most important to understand the benefit-risk attributes that are most important to patients not yet treated and how such patients trade off such risks. The mixing of both untreated and treated patients may make it hard to understand how the pre-procedure patients view these issues.

Prior treatment could significantly influence a patient's preference for one treatment option or another, but it is hard to know *a priori* how treatment would influence a patient's preferences.

Prior treatment may introduce a confirmation bias that patients tend to prefer the procedure that they chose to have, and therefore require much greater benefit or much less risk from the alternative procedure compared to treatment naïve patients. Alternatively, if patients had a negative experience with their prior procedure, they may find the benefit/risk profile of the alternative procedure much more attractive than that of the procedure that they had. Additionally, their experience of specific benefits or risks from their procedure may skew their weighting of those specific attributes compared to other patients.

For this study, it would be important in Table 4 to add a breakdown of the patient treatment history, specifically the number and percentage of patients that have had SAVR, TAVR, or were untreated.

Additionally, it would be helpful to add a comparison of the preferences of patients in each of these categories to show how prior treatment affects the MIR/MRB for each attribute, perhaps in a table similar to Table 5 except substituting treatment category for age. One concern is that with a sample size of 93 patients, sub-categorization by treatment status may result in too few patients in any one category to have confidence in the results. If this is a problem, it should be acknowledged that future studies may be needed to better understand the effect of prior treatment on preferences in aortic valve replacement.

2. Representativeness of patients involved in member organizations: The study recruited participants from two member organizations (Heart Value Voice and Mended Hearts). While it is understandable why the membership of these organizations facilitated identifying patients with aortic valve disease. However, patients who are motivated enough to join such organizations may have different preferences than the broader population of patients with the disease eligible for treatment of their aortic valve disease. There is no way to assess this potential difference in this study, but the authors could acknowledge this potential source of bias in the sample population in their discussion of the results and encourage future study in one or more different aortic valve disease populations.

3. Lack of clarity in the attribute of "Time over which the procedure has been proven to work": From the definition of this attribute in the paper, it is difficult to know whether patients interpreted this attribute as a measure of how long they could expect benefit, i.e., duration of effect, or whether patients also viewed this as how much clinical experience there was with a treatment, i.e., uncertainty in the knowledge about the effect. (See section 2 of the MDIC Patient Centered Benefit Risk Report for a nice discussion of uncertainty and how it relates to patient preferences.) This ambiguity raises the question about whether this attribute as described elicited preferences about expected duration, or elicited preferences about patient tolerance of uncertainty about the effect of TAVR, or some combination of the two. Showing an example of how this "proven to work" attribute was shown to patients akin to Figure 1 might clarify this ambiguity. Future studies might try to separate these attributes, particularly comparing an

established therapy like SAVR with a newer treatment like TAVR.

Minor issues and typos:

1. Introduction, last paragraph of left column, p. 3: the beginning of the second sentence is awkward: "However, little is known about the weights that patients assign to which attributes," I would suggest "However, little is known about the weights that patients assign to each attribute. . . ."
2. Methods, Attribute selection, last sentence (p.3): note "was were". I would remove the were.
3. Table 1, p. 4: Disabling non-fatal stroke: in the description, first sentence, note the "one 1" – should be "one month".
4. Table 2, p. 4, description of "independence" attribute: the description of the measure of this attribute does not mention that this is measured at one month, whereas all the other measures note the time frame of the measure.
5. Analysis, p. 6, top right column, first paragraph: there looks to be a calculation error in the next to the last sentence. $4.8\%/2\% = 2.4$, not 1.4 as written.
6. Comparisons of TAVR and SAVR, p. 7: right column, top paragraph – "In each case, patents. . ." should be "In each case, patients. . ."
7. Figures 2 and 3, p. 9: Each x-axis should be labeled with what the unit of measurement is. This will help the figures be more self-explanatory.

I hope that these comments are helpful. Given my interest in the use of patient preferences in the FDA regulatory process and increasing shared decision making in medicine broadly, I am pleased that the authors undertook this study to better understand patient preferences in aortic stenosis treatment, and I support its publication. It should be a nice addition to both the aortic stenosis treatment literature and the patient preference literature.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: I was asked to speak to the senior management of Edwards Lifesciences, which provided the grant for this study, in August 2017 about the use of patient preference information in FDA regulatory approval of medical devices. I was not paid for the talk, but was

reimbursed travel expenses. Additionally, one of the authors, Barry Liden, has served with me on the Steering Committee for the MDIC Patient Centered Benefit Risk (PCBR) Project, where I was the MDIC Board member leading that project. Barry now chairs the MDIC Science of Patient Input (SPI) Steering Committee, where I am also a member.

Reviewer Expertise: Internal medicine physician with prior experience in outcomes research and shared decision making. Not actively involved in research in my present role as a venture capitalist, but I led the MDIC Patient Centered Benefit-Risk Project that developed a framework for incorporating patient preferences into FDA medical device regulation and developed the first catalog of patient preference methodologies included as an appendix to that report.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 08 May 2019

Kevin Marsh, Evidera Inc, London, UK

Review 1: Ross Jaffe

This article by Marsh et al assesses patient preferences regarding the important medical decision about whether to undergo traditional surgical aortic valve replacement (SAVR) versus the newer transcatheter aortic valve replacement (TAVR) for aortic stenosis. This is an important effort to understand what attributes of benefit and risk from aortic valve replacement are most important to patients and understand how patients tradeoff those attributes in deciding which therapy to choose. With shared decision-making becoming increasingly used as our US healthcare system evolves to be more patient-centric, studies such as this one are important to expanding our understanding of how best to provide the information needed by patients to make informed decisions about their care. This study confirms a belief that many clinicians have that most patients prefer TAVR to SAVR because it is less invasive, providing evidence that about 75% patients prefer the less invasive approach. These patients are willing to tolerate a somewhat higher risk of stroke, pacemaker placement, and dialysis, and less evidence of long-term duration for the benefits of less invasiveness. It also identifies that younger patients (<60 yo) perspectives may differ from those of older patients (>60 yo) somewhat, although both groups generally prefer the less invasive procedure.

I support the publication of this study. While this study is not perfect from my point of view, it is an important contribution to the literature, both about how to treat aortic stenosis and about patient preference assessment more generally. Understanding what treatment attributes are most important to patients and how patients trade off benefits and risks is important to understanding how best to inform patients about their treatment options and help them make decisions that best reflect their individual preferences. As the paper notes on p. 11, there has only been one other, more limited study of patient preferences in aortic stenosis, so this study is an important expansion of our understanding of patient preferences that clinical area.

Additionally, from a methodological perspective and as also noted in the paper, there are only a few other studies that use adapted swing weights (ASW) to assess patient preferences. Swing weighting is one of only fourteen methods identified in the Medical Device Innovation Consortium (MDIC) review of preference assessment methodologies. (See: MDIC Patient Centered Benefit Risk Project Report, Appendix A: Catalog of Methods for Assessing Patient Preferences for Benefits and Harms of Medical Technologies, May 2015, available at: <https://mdic.org/resource/patient-centered-benefit-risk-pcbr-framework/>). This study is an important contribution to the literature about swing weighting methods, and should allow comparison to studies of patient preference using other methods to help researchers and clinicians better understand how best to assess patient preferences.

Response: We thank the reviewer for their interest in our study

This study does have a few issues that should be highlighted to help put the results in context. I also note an area for further assessment of the data as well as areas for future research focus. Please note that I come at this study as someone with an interest in patient preferences from clinical and regulatory policy perspective, and do not have the expertise to comment on the specific methodology or statistical analysis, which I will leave to experts in those areas.

Major concerns:

1. The mixing of treated and untreated patients in the participant population: In Box 1 on p. 6, the inclusion criteria describe that patients in the study could have already had a procedure within the last 10 years or could be untreated. From a shared decision making point of view, patients express their preferences prior to treatment – so it is most important to understand the benefit-risk attributes that are most important to patients not yet treated and how such patients trade off such risks. The mixing of both untreated and treated patients may make it hard to understand how the pre-procedure patients view these issues.

Prior treatment could significantly influence a patient's preference for one treatment option or another, but it is hard to know *a priori* how treatment would influence a patient's preferences. Prior treatment may introduce a confirmation bias that patients tend to prefer the procedure that they chose to have, and therefore require much greater benefit or much less risk from the alternative procedure compared to treatment naïve patients. Alternatively, if patients had a negative experience with their prior procedure, they may find the benefit/risk profile of the alternative procedure much more attractive than that of the procedure that they had. Additionally, their experience of specific benefits or risks from their procedure may skew their weighting of those specific attributes compared to other patients.

For this study, it would be important in Table 4 to add a breakdown of the patient treatment history, specifically the number and percentage of patients that have had SAVR, TAVR, or were untreated.

Additionally, it would be helpful to add a comparison of the preferences of patients in each of these categories to show how prior treatment affects the MIR/MRB for each attribute, perhaps in a table similar to Table 5 except substituting treatment category for age. One concern is that with a sample size of 93 patients, sub-categorization by treatment status may result in too few patients in any one category to have confidence in the results. If this is a problem, it should be acknowledged that future studies may be needed to better understand the effect of prior treatment on preferences in aortic valve replacement.

Response: This is an important observation. We have updated Table 4 as suggested. We have not replicated Table 5, but have instead added the following note that no difference in preferences were observed between participants who had / hadn't been treated previously for their AS.

"No other correlation was found between participant characteristics and MIR/MRB. This includes whether a participant reported having previously undergone treatment for their AS. While this might be expected to influence preferences, the ability of the analysis to identify this influence is limited by the small sample size and the relatively small proportion of participants who reported not having previously received AS treatment (11.8%)."

The discussion section has also been updated to include the following statement:

"Further research could usefully gather data from a larger sample of AS patients to determine the association of preferences and patient characteristics, such as NYHA classification or whether patients have previously undergone treatment for AS."

2. Representativeness of patients involved in member organizations: The study recruited participants from two member organizations (Heart Valve Voice and Mended Hearts). While it is understandable why the membership of these organizations facilitated identifying patients with aortic valve disease. However, patients who are motivated enough to join such organizations may have different preferences than the broader population of patients with the disease eligible for treatment of their aortic valve disease. There is no way to assess this potential difference in this study, but the authors could acknowledge this potential source of bias in the sample population in their discussion of the results and encourage future study in one or more different aortic valve disease populations.

Response: We agree that the recruitment strategy may be a source of bias in the preference data. We have acknowledged this in the discussion, which now includes the statement:

"

Finally, the current study is based on a relatively small sample of patients, and the sample is healthier and younger than the population currently eligible for TAVR and SAVR. AS patients were recruited from the membership of Heart Valve Voice and Mended Hearts, and it is possible that patients who are motivated to join these organizations may have different preferences than the broader population. Further AS patient preference research should replicate this study in a larger sample of patients, including more patients with more severe disease burden"

3. Lack of clarity in the attribute of “Time over which the procedure has been proven to work”: From the definition of this attribute in the paper, it is difficult to know whether patients interpreted this attribute as a measure of how long they could expect benefit, i.e, duration of effect, or whether patients also viewed this as how much clinical experience there was with a treatment, i.e., uncertainty in the knowledge about the effect. (See section 2 of the MDIC Patient Centered Benefit Risk Report for a nice discussion of uncertainty and how it relates to patient preferences.). This ambiguity raises the question about whether this attribute as described elicited preferences about expected duration, or elicited preferences about patient tolerance of uncertainty about the effect of TAVR, or some combination of the two. Showing an example of how this “proven to work” attribute was shown to patients akin to Figure 1 might clarify this ambiguity. Future studies might try to separate these attributes, particularly comparing an established therapy like SAVR with a newer treatment like TAVR.

Response: We have updated Figure 1 to use an example of a survey question that includes the ‘time proven to work’ attribute.

Minor issues and typos:

1. Introduction, last paragraph of left column, p. 3: the beginning of the second sentence is awkward: “However, little is known about the weights that patients assign to which attributes,” I would suggest “However, little is known about the weights that patients assign to each attribute. . . .” **Response: Updated as suggested**
2. Methods, Attribute selection, last sentence (p.3): note “was were”. I would remove the were. **Updated as suggested**
3. Table 1, p. 4: Disabling non-fatal stroke: in the description, first sentence, note the “one 1” – should be “one month”. **Response: Updated as suggested**
4. Table 2, p. 4, description of “independence” attribute: the description of the measure of this attribute does not mention that this is measured at one month, whereas all the other measures note the time frame of the measure. **Response: The current description includes a timeframe. No update made**
5. Analysis, p. 6, top right column, first paragraph: there looks to be a calculation error in the next to the last sentence. $4.8\%/2\% = 2.4$, not 1.4 as written. **Response: Thank you for noticing this. Updated.**
6. Comparisons of TAVR and SAVR, p. 7: right column, top paragraph – “In each case, patents. . . .” should be “In each case, patients. . . .” **Response: Updated**
7. Figures 2 and 3, p. 9: Each x-axis should be labeled with what the unit of measurement is. This will help the figures be more self-explanatory. **Response: Labels have been added to the Figures.**

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research