ORIGINAL RESEARCH

Participant Engagement and Preference Study for Clinical Outcomes Associated With Atrial Fibrillation

The PEARL-AF Study

Shelby D. Reed, PhD,[a] Josephine L. Harrington, MD,[a] Daniel P. Morin, MD, MPH,[b] Samir F. Saba, MD,[c]
Jay Alan Montgomery, MD,[d] Robert W. Harrison, MD,[e] Daniel R. Frisch, MD,[f] Thomas Viethen, MD,[g]
Miriam Tamm, PhD,[g] Jessica Xiao, PharmD,[h] Hardi Mundl, MD,[g] Rosa Coppolecchia, DO, MPH,[h]
Jui-Chen Yang, MEM,[a] Matthew J. Wallace, MA,[a] Juan Marcos Gonzalez, PhD,[a] Manesh R. Patel, MD[a,e]

ABSTRACT

BACKGROUND Quantifying patients' preferences for health outcomes associated with atrial fibrillation (AF) and its treatments offers a replicable approach to considering the patient perspective in regulatory decision-making.

OBJECTIVE The authors conducted a preference survey to estimate the relative importance of AF-related events for use in clinical trial analyses to estimate net health benefits with anticoagulants.

METHODS The survey included nontechnical descriptions of three severities of stroke, systemic embolism, myocardial infarction (MI) with or without subsequent heart failure (HF), major bleeding, clinically relevant nonmajor bleeding, and death. A best-worst scaling question format was used in which patients were shown 10 sets of four events and asked to select what they considered to be most and least serious.

RESULTS One thousand twenty-eight patients, mean age 69.2 years, 40.4% female, completed the survey. Best-worst scaling importance weights were significantly different across all events except between major bleeding and MI with HF. Death was considered the most serious (reweighted to 1), followed by severe disabling stroke (0.83), then major bleeding (0.53) or MI with HF (0.50), moderate-severity stroke (0.28) and systemic embolism (0.13). Clinically relevant nonmajor bleeding, MI without HF, and minor stroke (0.10, 0.06, and 0.04, respectively) were considered least serious. Events ordered by importance were consistent across age, sex, and race, but relative weights across events varied by sex and race.

CONCLUSIONS Patients expressed relatively high levels of concern about major bleeding compared to moderate-severity stroke or systemic embolism, endpoints frequently used in AF trials. Estimated weights could be used in patient-centered net-benefit determinations for AF therapies. (JACC Adv. 2024;3:101370) © 2024 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

ABBREVIATIONS AND ACRONYMS

AF = atrial fibrillation

BWS = best-worst scaling

CR-NMB = clinically relevant nonmajor bleeding

HF = heart failure

MI = myocardial infarction

mRS = modified Rankin Scale

PRO = patient-reported outcome

The U.S. Food and Drug Administration has issued several guidance documents encouraging greater involvement of patients throughout drug development and submission of patient experience data, which includes patient-reported outcomes (PROs) and patient preference information.[1] Food and Drug Administration's 2023 guidance on benefit-risk assessment encourages investigators and sponsors to submit patient preference information to provide patients' views on the relative importance of benefits and risks.[2] Increasingly, cardiovascular trials have included PROs and other endpoints that are important to patients. In the setting of atrial fibrillation (AF), patient-reported measures pertaining to symptoms, functioning, and treatment burden are important to consider given their link with medication nonadherence and discontinuation. However, in the evaluation of benefits and risks of new anticoagulation therapies, while necessary, PROs are insufficient because they do not incorporate patients' concerns about treatment-related adverse events or their preferences for health outcomes that treatments are designed to prevent.

Most clinical trials evaluating treatments for AF are designed to evaluate whether the treatments reduce the incidence of a composite endpoint, frequently inclusive of death and nonfatal thromboembolic events like ischemic stroke and systemic embolism. Although one would not argue that these events are important and undesirable to patients, their relative importance is not considered in the analysis. Thus, a treatment that reduces the incidence of more frequent, less severe events could be found superior to an alternative treatment that reduces the incidence of less frequent, more severe events. A further complicating factor in evaluating AF treatments is the inverse relationship between risks of thromboembolic and bleeding events. Given these issues, a comprehensive evaluation of benefits and risks is paramount. Regulators and experts tasked with developing practice guidelines are accustomed to jointly considering benefits and risks among treatment alternatives.

Thus, to ensure that new treatments improve patients' overall wellbeing, it is necessary to incorporate their concerns about the relative importance of AF-related outcomes in benefit-risk assessments. Using quantitative benefit-risk assessment approaches paired with stated-preference methods, it is possible to obtain transparent relative-importance weights of endpoint events from patients or other stakeholders and apply those weights to event-specific incidence rates for alternative treatments to derive a measure of net benefit.[3]

In anticipation of evaluating the net benefit of new anticoagulants, specifically factor XI/XIa inhibitors, being studied in large, randomized clinical trials, we designed a multi-site, cross-sectional patient-preference study to quantify the relative importance of endpoints commonly included in clinical trials testing anticoagulation therapies in AF.

METHODS

An "object-case" best-worst scaling (BWS) exercise, a method based in random-utility theory, was chosen as the primary preference-elicitation method.[4] Asking respondents to rank a long list of options is known to fail tests of validity and reliability.[5,6] The BWS method obtains valid and reliable quantitative rank scores by showing participants multiple subsets of objects drawn from a larger list, such as medical events or health states, and asking them to choose the two that are most extreme in each list, a far easier task than ranking the whole set. With multiple iterations, it is possible to rank order the objects for each participant and to compute a relative-importance weight for each event for a sample of participants. Object-case BWS was chosen as the preference-elicitation method because it is a more efficient way to obtain weights for specific trial endpoints of regulatory interest than a discrete-choice experiment, which typically obtains preference information over a larger set of tradeoffs involving probabilistic outcomes. Furthermore, the intent was not to evaluate the impact of potential benefits and risks on patients' well-being but to generate patient-centered weights that could be applied to realized events in a clinical trial.

A seven-member patient advisory committee and a five-member steering committee of cardiologists were formed to guide development of a patient-centric survey that also accurately portrayed clinical endpoint events. To inform the selection of events for the BWS exercise, the study team identified clinical endpoints used in previous and ongoing trials of oral anticoagulation therapies in AF. These primarily included thrombotic events (ischemic stroke, systemic embolism, and myocardial infarction) and bleeding events (including major and clinically relevant nonmajor). Study investigators and patient advisors recognized that the relative importance of these events would vary depending upon their severity, medical intervention required, and a patient's recovery and health status following the

event. **Table 1** lists the nine clinical events selected for the study and the corresponding clinical trial endpoints. Three severity levels for stroke events correspond to modified Rankin scale (mRS) global disability outcomes. The least severe stroke corresponds to a 0 or 1 on the mRS. Strokes leading to moderate disability and severe disability correspond to scores of 3 and 5 on the mRS, respectively. Severity levels for bleeding events align with definitions from the International Society on Thrombosis and Haemostasis Scientific and Standardization Committee (Supplemental Table 1).[7,8]

The study team drafted the survey instrument, which was then reviewed by the patient advisory committee and clinical steering committee with particular attention to descriptions of clinical events. Descriptions and graphics used for stroke events were informed by consensus recommendations from the Stroke Therapy Academic Industry Roundtable XI.[9] The study team considered descriptions of events used in published patient preference studies on benefits and risks associated with AF treatments.[10-12] The survey instrument included the following sections: screening questions; introduction to the survey with a brief overview about AF; questions on personal health history; event descriptions; questions to assess participants' understanding and retention of survey content; BWS questions; and token-allocation and ranking exercises.

The study team conducted ten virtual pretest interviews with individual patients with AF between March 8, 2022, and April 10, 2022. The interviews tested the appropriateness and understandability of the survey content. A programmed version of the survey was used in the interviews, during which participants were asked to read the survey text aloud and to articulate their thoughts related to its content and their decision processes. An early observation illuminated through the interviews was that participants were making assumptions about the level of recovery following medical events based on personal experience or that of friends or family. To standardize this information for all participants, the team reframed the events as "people who had each event" to deterministically describe realized functional health status following the event. For example, descriptions for three stroke severity levels are provided in **Figure 1**.

In the prompt for the BWS questions, lists of four people with different events were shown, and participants were asked to select the person with the "most serious" and the person with the "least serious" medical problem (example in **Figure 2**). The survey instrument was iteratively revised between

TABLE 1 Medical Events in the Survey and Corresponding Clinical Trial Endpoints

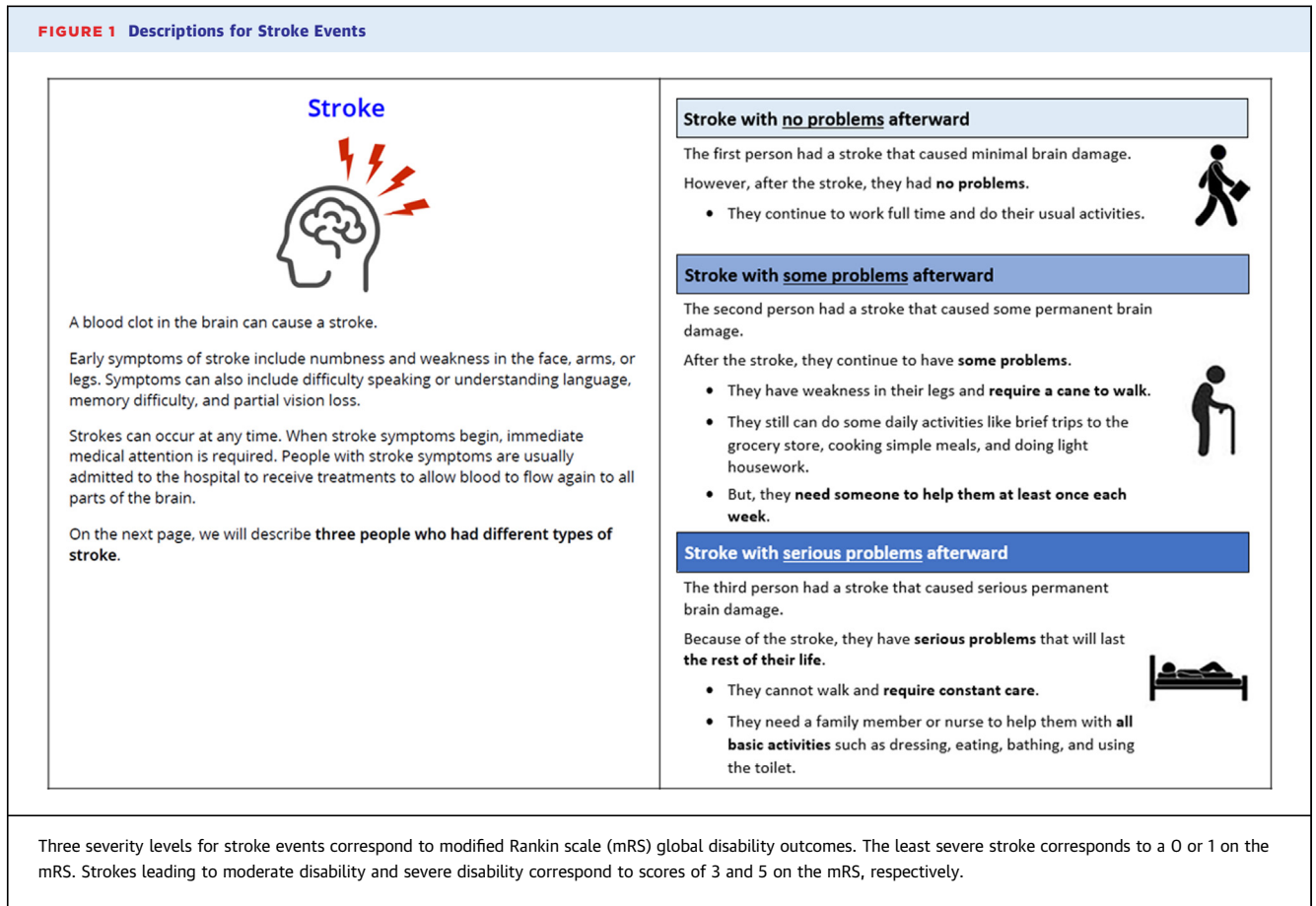| Medical Events	Clinical Trial Endpoints
Stroke with no problems afterward	Minor stroke (modified Rankin scale 0, 1)
Stroke with some problems afterward	Moderate-severity stroke (modified Rankin scale 2, 3)
Stroke with serious problems afterward	Disabling stroke (modified Rankin scale 4, 5)
Clot in the leg	Systemic embolism
Heart attack with no problems afterward	Nonfatal myocardial infarction; no functional deficit
Heart attack with weakened heart afterward	Nonfatal myocardial infarction resulting in heart failure
Death	All-cause death or cardiovascular death
Bleeding requiring medical care and/or change in atrial fibrillation medication	Clinically relevant nonmajor bleed (CR-NMB)
Bleeding requiring emergency care and blood transfusion	Major bleeding event

interviews. For the last three participants, interviewers withheld verbal questions and observed participants as they completed the online survey, similar to field conditions, to examine how they completed survey questions, to gauge completion time, and to identify potential problems with self-administration. The survey instrument is provided in online supplemental materials.

EXPERIMENTAL DESIGN. The four events (ie, people) shown in each BWS question were determined using a balanced incomplete block design.[13] The design comprised 300 choice sets, each of which included 10 lists of four events. Each survey participant received a set of ten BWS questions. Overall, each of the nine clinical events appeared 1,333 or 1,334 times; each pair of events appeared 499 or 500 times; and each event appeared in the first, second, third, or fourth position 333 or 334 times. There were no constraints in the experimental design, meaning that all combinations of events were possible.

In addition to our primary preference-elicitation approach with BWS, the survey included a token-allocation exercise. For the token-allocation exercise, participants were to allocate 100 tokens across the events, assigning more (fewer) tokens to events they considered more (less) serious. The ranking exercise asked participants to rank the clinical events from most to least serious. To avoid potential ordering effects in the ranking and token allocation exercises, the order of the events shown was randomized across participants.

The survey was programmed and hosted using Lighthouse Studio version 9.13.1 (Sawtooth Software Provo). The survey was also translated to Spanish and pretested with 6 Spanish-speaking patients. A Spanish version of the survey was available to recruiting partners.

FIGURE 1 Descriptions for Stroke Events

Three severity levels for stroke events correspond to modified Rankin scale (mRS) global disability outcomes. The least severe stroke corresponds to a 0 or 1 on the mRS. Strokes leading to moderate disability and severe disability correspond to scores of 3 and 5 on the mRS, respectively.

RECRUITMENT. Survey participants were recruited from two sources. One was an online platform (Medidata; Dassault Systèmes) that included individuals with a confirmed diagnosis of AF. Patients participating in the platform received email invitations to complete the online survey. The second source was partnering clinical sites at the University of Pittsburgh Medical Center, Vanderbilt University Medical Center, Duke University Health System, Jefferson Health/Pier Consortium, and Ochsner Heart & Vascular. Various

FIGURE 2 Example BWS Question

Participants could click or hover over medical problem to view its description. BWS = best-worst scaling.
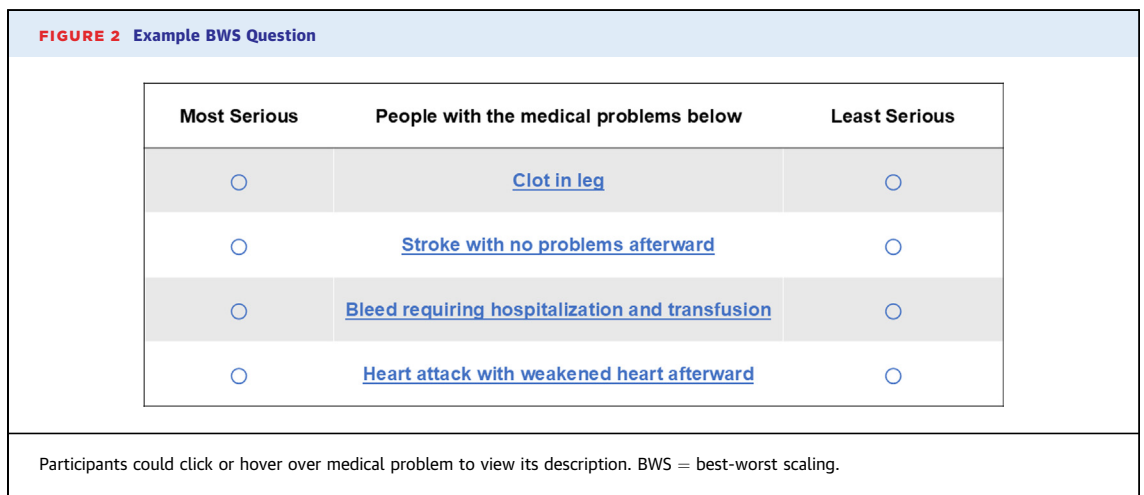
patient identification and recruitment approaches were used, including in-person flyers, posters, and electronic invitations sent via direct email or online health portals. Patients who completed the survey were offered a $40 electronic gift card. All patients participating in pretest interviews and the main survey provided informed consent (Duke University Health System Institutional Review Board protocols numbers: Pro00110367 and Pro00110208).

STATISTICAL ANALYSIS. Descriptive statistics were used to summarize participants' sociodemographic and clinical characteristics. BWS data were analyzed using conditional-logit regression models with the dependent variable coded 1 for the chosen most serious event, −1 for the chosen least serious event, and 0 for events not chosen. Independent variables were effect-coded such that the omitted event, in this case death, was the negative sum of the coefficients of all other events, making 0 equal to the mean effect. To aid interpretation as BWS relative-importance weights, resulting model coefficients were adjusted to sum to 100.[14] To facilitate computation of net-benefit measures, BWS weights were also rescaled relative to death, where death had a weight of 1, and all the other clinical events were rescaled accordingly. CIs were estimated using the Krinsky and Robb procedure with 10,000 simulations.[15] Data from the token allocation and ranking exercises were used to compute mean estimates to generate rankings for comparison with BWS results. All statistical analysis was performed in StataSE version 18 (StataCorp LLC).

SUBGROUP COMPARISONS. Separate conditional-logit models were fit to BWS data for subgroups defined by sex, age, and race. Likelihood-ratio tests were used to compare overall results between subgroups, and z-scores were computed to compare BWS weights for individual events between groups. To facilitate the use of our findings in quantitative net-benefit evaluations and comparisons across subgroups, BWS weights were also rescaled relative to death to facilitate comparisons of the relative importance of events vs death.

RESULTS

A total of 1,240 patients with AF initiated the survey, of which 1,028 (83%) completed the survey between September 26, 2022, and June 12, 2023. Among the 1,028 participants, the mean age was 69.2 ± 9.4 years, 59.1% were male (sex), 89.5% identified as White, and 6.7% as Black or African American (**Table 2**). There was a broad distribution in terms of highest level of formal education with 12.6% having a high school education or less, 30.3% having some college,

TABLE 2 Participant Characteristics (N = 1,028)	
Age, y	69.2 ± 9.4
Sex[a]	
Female	40.4% (415)
Male	59.1% (608)
Race[bc]	
American Indian or Alaska Native	0.7% (7)
Asian	1.4% (14)
Black or African American	6.7% (69)
Native Hawaiian or Other Pacific Islander	0.0% (0)
White	89.5% (920)
Other	1.0% (10)
Hispanic, Latino, or Spanish origin[d]	2.1% (22)
Highest level of education[e]	
High school diploma or equivalent or less	12.6% (130)
Some college but no degree	18.0% (185)
Technical school or associate's degree	12.3% (126)
4-y college degree ± some graduate school	26.4% (271)
Graduate or professional degree	29.8% (306)
Time since AF diagnosis[f]	
<1 y	266 (25.9%)
1-3 y	341 (33.2%)
4-10 y	250 (24.3%)
>10 y	158 (15.4%)
Not sure	12 (1.2%)
Treatments received for AF[g]	
Anticoagulant medicine	888 (86.4%)
Heart rhythm medicine	431 (41.9%)
Heart rate medicine	760 (73.9%)
Cardiac/catheter ablation	350 (34.0%)
Electrical cardioversion	399 (38.8%)
CHA$_2$DS$_2$-VASc score[g]	2.9 ± 1.5

Values are mean ± SD, % (n), or n (%). [a]5 "prefer not to answer". [b]Percentages may not add to 100%; [c]"prefer not to answer". [c]10. [d]37. [e]10. [f]1 missing. [g]5 missing. AF = atrial fibrillation.

technical school, or associate's degree, 26.4% having a 4-year degree, and 29.8% having a graduate degree. Median time spent completing the survey was 32.2 minutes (25th and 75th percentiles: 24.5 and 43.3 minutes).

Five questions were designed to evaluate retention of information presented in the event descriptions; percentages of participants providing correct responses ranged between 58.9% and 94.2%. The question with the lowest correct-response rate asked about the frequency with which a person with moderate-severity stroke needed help from another person. This question appeared several screens after the descriptions of the three types of stroke were presented.

Estimated BWS weights for all events relative to death (rescaled to 1) are presented in **Figure 3, Central Illustration**. Weights were significantly different ($P < 0.01$) between all clinical events except scores for

FIGURE 3   BWS Weights by Event Rescaled Relative to Death for Full Sample and by Age Group, Sex, and Race
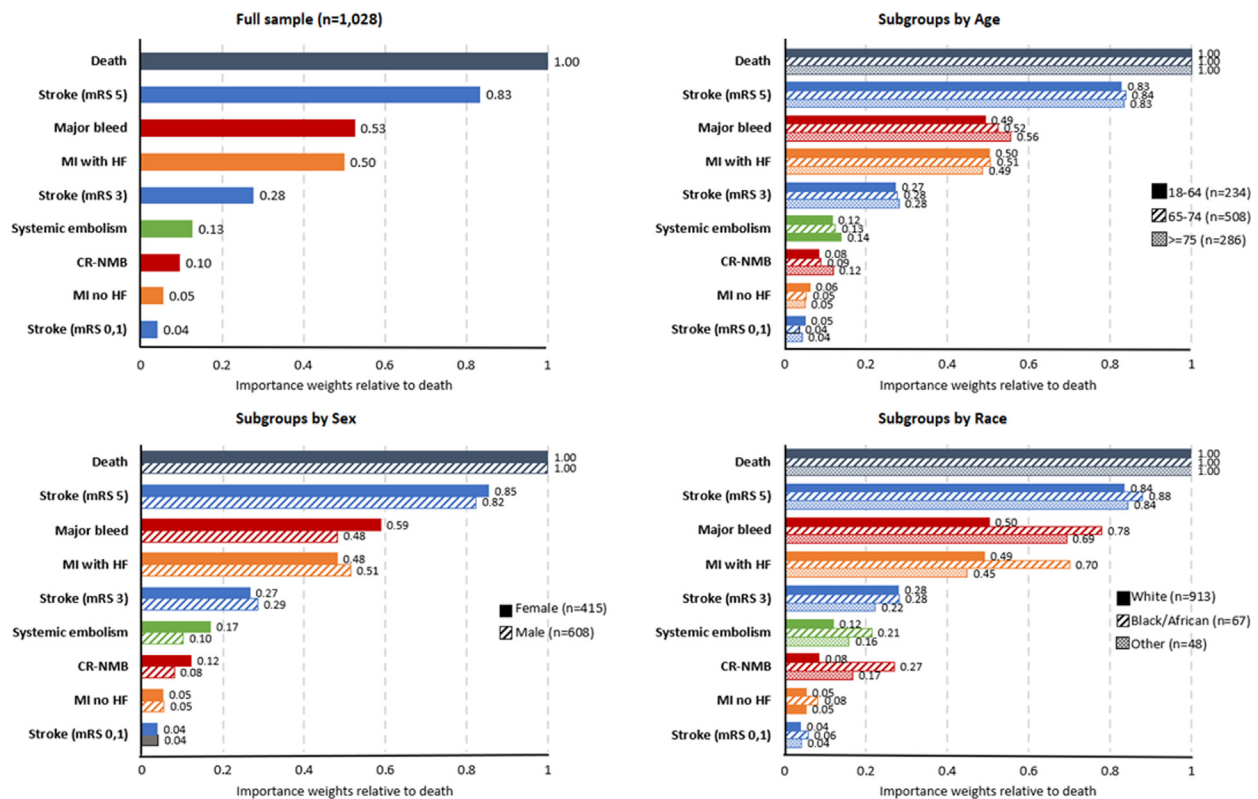
Brackets represent 95% CIs. BWS = best-worst scaling; CR-NMB = clinically relevant nonmajor bleed; HF = heart failure; MI = myocardial infarction; mRS = modified Rankin scale.

major bleed and MI with heart failure (HF) ($P = 0.104$). Death was considered the most important event, followed by disabling stroke (0.83, 95% CIs in Supplemental Table 2). Major bleeding and MI with subsequent HF were the next most important (0.53 and 0.50). Compared to these two events, moderate-severity stroke was half as important (0.28). Events with the smallest weights were systemic embolism (0.13), clinically relevant nonmajor bleeding (CR-NMB) (0.10), MI, and stroke with no residual limitations (0.06 and 0.04). The rank orders of events were consistent across the BWS, token-allocation, and ranking exercises when accounting for the lack of statistically significant difference between the BWS weights for major bleed and MI with HF (Supplemental Table 3).

SUBGROUP COMPARISONS. Male and female participants had statistically significant differences in BWS weights for six of nine events (Table 3). Males considered death, MI with HF, and moderate-severity stroke as significantly more serious than females. Females considered major bleeding, systemic embolism, and CR-NMB to be significantly more serious than males.

BWS weights generally were consistent across groups of participants aged 18 to 64 years, 65 to 74 years, and 75 years or older. There were statistically significant but small differences in BWS scores for some of the lower weighted events: CR-NMB, MI without subsequent HF, and minor stroke.

Although only 11% of participants identified as a race other than White, there were significant differences in BWS weights across race groups for six of nine events, and the differences were of greater magnitude than observed between sexes or across age groups (Table 3). Participants identifying as Black or a race other than Black or White both considered major bleeding and CR-NMB as significantly more serious than White participants. Conversely, Black participants considered death, disabling stroke, and

CENTRAL ILLUSTRATION The PEARL-AF Study

Objective: To quantify the relative importance of AF-related events for use in clinical-trial analyses of anticoagulants

Methods: Web-based preference survey included:
- Patient-friendly descriptions of each event, medical care required, and recovery period or continuing health or functional deficits.

- Object-Case Best-Worst Scaling exercise

- Medical history and sociodemographic characteristics

Relative importance of clinical outcomes relative to death
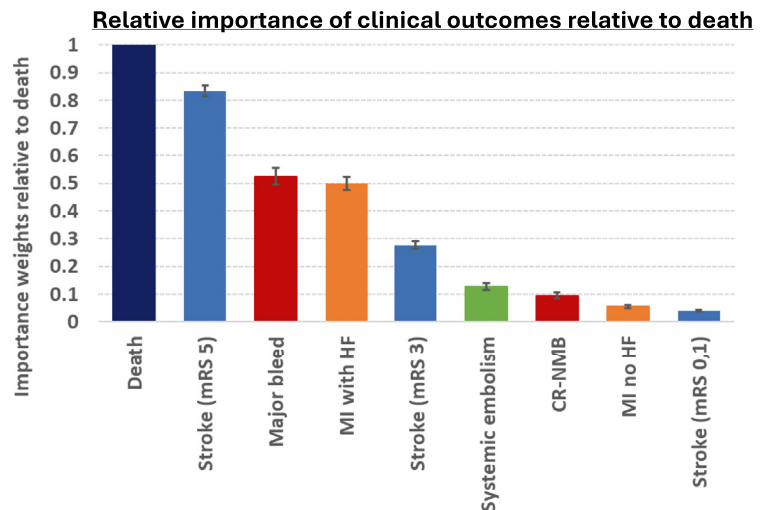
Event-specific weights could be applied to clinical trial data to compute net benefit.

Reed SD, et al. JACC Adv. 2024;3(12):101370.

moderate-severity stroke as relatively less serious than White participants.

DISCUSSION

Health preference researchers design their survey instruments to control experimental stimuli to the extent possible. This requires developing survey content that is clearly understood and relevant to patients while being clinically accurate. It allows researchers, clinicians, and policymakers greater insight and certainty about what patients were considering when completing the survey. This level of transparency and consistency is key to generating reliable evidence on patients' preferences.

Our best-practice preference study found that patients prioritized death and severe disabling stroke as the most serious events, but we also found that patients were nearly as concerned about major bleeding associated with anticoagulation therapy as they were about the thromboembolic events it is intended to prevent. This is important because many clinical trials evaluating anticoagulation therapy in AF include only thromboembolic events in composite endpoints. Bleeding events are generally considered adverse events and are not typically considered in analyses of treatment efficacy.

Consistent with a 2001 valuation study using a standard-gamble exercise,[16] our findings show that the importance of a stroke to patients is highly dependent on its severity and subsequent disability. A 2017 systematic review of published papers on patient preferences in anticoagulation in AF reported that prevention of severe stroke was the single most important factor for patients, but most of the studies were limited to stroke or stroke and bleeding as the only health outcomes evaluated. The authors noted that quality of the studies was moderate and limited by small sample sizes.[17] More recent studies reveal that the most significant features driving patients' choices of anticoagulation options are reduction of stroke, bleeding risk, and individual recommendations of their provider.[10-12] Other studies have attempted to quantify preferences by evaluating patients' willingness to pay.[18,19] None of the studies included a broader range of medical events and their associated severity, both of which could be important in patients' choices about anticoagulation options and in evaluating the net benefit afforded by alternative anticoagulation therapies. Furthermore, most of these studies noted significant heterogeneity in patient-level preferences, but they lacked adequate sample sizes to evaluate whether preferences varied across patient characteristics. In contrast, our study included a

TABLE 3 Comparisons of BWS Weights by Sex, Age, and Race

	n	Death	Stroke (mRS 5)	MI With HF	Major Bleed	Stroke (mRS 3)	Systemic Embolism	CR-NMB	MI No HF	Stroke (mRS 0,1)	P Value[e]
Sex											
BWS weights											
Male	608	**29.58**[a]	24.34	**15.21**[a]	**14.24**[a]	**8.44**[a]	**3.02**[a]	**2.36**[a]	1.62	1.20	<0.0001
Female	415	**28.03**[a]	23.93	**13.47**[a]	**16.47**[a]	**7.44**[a]	**4.72**[a]	**3.39**[a]	1.46	1.08	
Relative weights vs death											
Male		1.00	0.82	0.51	0.48	0.29	0.10	0.08	0.05	0.04	
Female		1.00	0.85	0.48	0.59	0.27	0.17	0.12	0.05	0.04	
Age											
BWS weights											
18-64 y	234	29.35	24.29	14.77	14.45	8.01	3.41	**2.45**[b]	**1.80**[b]	**1.46**[a]	<0.0001
65-74 y	508	28.99	24.28	14.68	15.18	8.05	3.63	**2.60**[b,c]	1.57	**1.03**[a]	
75+ y	286	28.52	23.78	13.87	15.83	8.00	3.97	**3.39**[b,c]	**1.42**[b]	1.04	
Relative weights vs death											
18-64 y		1.00	0.83	0.50	0.49	0.27	0.12	0.08	0.06	0.05	
65-74 y		1.00	0.84	0.51	0.52	0.28	0.13	0.09	0.05	0.04	
75+ y		1.00	0.83	0.49	0.56	0.28	0.14	0.12	0.05	0.04	
Race[d]											
BWS weights											
White	913	**29.38**[a]	**24.54**[a]	**14.42**[b]	**14.76**[a,b]	**8.22**[a,b]	3.50	**2.49**[a,b]	1.54	1.14	<0.0001
Black	67	**23.46**[a]	**20.65**[a,c]	16.42	**18.28**[a]	**6.60**[a]	5.02	**6.31**[a]	1.94	1.34	
Other	48	27.57	**23.26**[c]	**12.38**[b]	19.09	**6.15**[b]	4.38	**4.60**[b]	1.43	1.13	
Relative weights vs death											
White		1.00	0.84	0.49	0.50	0.28	0.12	0.08	0.05	0.04	
Black		1.00	0.88	0.70	0.78	0.28	0.21	0.27	0.08	0.06	
Other		1.00	0.84	0.45	0.69	0.22	0.16	0.17	0.05	0.04	

Bold indicates statistically significant differences. [a]$P < 0.05$ for comparisons between male vs female, 18 to 64 vs 65 to 74 y, or White vs Black. [b]$P < 0.05$ for comparisons between 18 to 64 vs 75+ y or White vs other race. [c]$P < 0.05$ for comparisons between 65 to 74 vs 75+ y or Black vs other race. [d]Participants were categorized as: White if White was the only race option selected; Black if Black was selected with or without any other race options selected; and Other if any race other than Black was selected and White was not the only option selected. [e]P values based on log likelihood tests.

BWS = best-worst scaling; CR-NMB = clinically relevant nonmajor bleed; mRS = modified Rankin scale.

large sample size, clinical events stratified by severity, and deterministic descriptions about recovery and health status following each event.

Our study of 1,028 patients found significant differences in the relative importance of several events between sex and race groups. Female participants and Black and other-race participants indicated greater concern about major bleeding and CR-NMB relative to death, MI with HF, and moderate-severity stroke compared to their male and White counterparts, respectively. Although differences by sex and age groups were statistically significant, they were small in magnitude compared to differences observed across racial groups. Other preference studies related to benefits and risks of AF therapies did not include sufficient numbers of non-White participants to reliably test for potential differences between racial groups, in part because many have been conducted outside the United States.[17]

Quantitative approaches to combining importance weights from patients or other stakeholders with clinical evidence are increasingly prevalent in the

medical literature. These approaches include multi-criteria decision analysis and innovative analytic approaches.[20,21] Our findings from the BWS can be applied in these approaches to incorporate the patient perspective when evaluating the net-benefit associated with anticoagulation therapies for AF. When planning for these types of analyses, it is important to consider how trial endpoints align with event descriptions used in preference-elicitation surveys.

For weights used within a quantitative benefit-risk framework, an ISPOR Task Force recommends the use of preference-elicitation methods that require respondents to evaluate tradeoffs.[4] Methods recommended by the task force include the threshold technique, swing weighting, and discrete-choice experiments, but these methods cannot readily accommodate a large number of attributes. The threshold technique is limited to estimating a single threshold where benefit-risk tradeoffs are acceptable in cases where a treatment's benefits or risks can be defined. Furthermore, swing weighting and discrete-choice

experiments require cognitively challenging comparisons when considering multiple probabilistic attributes, and these exercises are more difficult to complete on mobile phones with limited screen sizes. Thus, these approaches were not tenable in achieving our goal to generate importance weights for nine events in a broad, representative group of older AF patients.

STUDY LIMITATIONS. A couple of limitations are useful to consider. First, study participants were primarily recruited from large academic health systems; thus, the findings may not be representative of the broader AF population. Second, although the sample is large, some groups of interest, including patients with a history of bleeding events or strokes, were too small to reliably test for potential differences. Third, even for groups where significant differences were noted, such as across racial groups, further qualitative and quantitative research would be helpful to identify potential explanatory factors. Fourth, our study findings are a function of the descriptions used for clinical events. The robustness of our results could be tested in the future using modified descriptions. Lastly, although our study findings are applicable to benefit-risk evaluations in AF, they may have more limited utility in clinical decision-making as current tools do not allow clinicians to predict an individual's risks of less or more severe bleeds or strokes.

CONCLUSIONS

Patients with AF at risk for stroke and systemic embolism expressed relatively higher than anticipated levels of concern about major bleeding compared to moderate-severity stroke or systemic embolism, which are clinical outcomes frequently used as primary efficacy endpoints in AF trials. Weights from this study that be applied to trial-based event frequencies in patient-centered net-benefit determinations for AF therapies. In the future, for both regulatory approval and shared decision-making aimed at personalized and patient-centered care for AF, net clinical benefit assessments for anticoagulants will be needed.

ADDRESS FOR CORRESPONDENCE: Dr Shelby D. Reed, Duke Clinical Research Institute, Duke University School of Medicine, 300 West Morgan Street, Durham, North Carolina 27701, USA. E-mail: Shelby.reed@duke.edu.

PERSPECTIVES

COMPETENCY IN SYSTEMS-BASED PRACTICE: When evaluating whether treatment benefits are not outweighed by their associated risks, value judgments about their relative importance are necessary. To ensure that new treatments improve patients' overall wellbeing and to support the choice of treatment, quantitative benefit-risk assessments can be paired with stated-preference methods to incorporate patients' views on the relative importance of clinical outcomes, which might be different than physicians' views.

TRANSLATIONAL OUTLOOK: Our study reports on a transparent, replicable approach using object-case best-worst scaling to estimate relative importance weights from patients for nine clinical events relevant to benefits and risks associated with anticoagulants in AF. Our estimates of event-specific relative importance weights could be applied to trial-based event frequencies in patient-centered net-benefit determinations for AF therapies to aid in regulatory and clinical decision-making.

REFERENCES

1. United States Food and Drug Administration. FDA patient-Focused drug development guidance series for enhancing the incorporation of the patient's voice in medical product development and regulatory decision making. Accessed April 9, 2024. https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical

2. United States Food and Drug Administration. Benefit-risk assessment for new drug and biological products guidance for Industry. Accessed April 9, 2024. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/benefit-risk-assessment-new-drug-and-biological-products

3. Tervonen T, Veldwijk J, Payne K, et al. Quantitative benefit-risk assessment in medical product decision making: a good practices report of an ISPOR Task Force. *Value Health*. 2023;26:449–460.

4. Louviere JJ, Flynn TN. Using best-worst scaling choice experiments to measure public perceptions and preferences for healthcare reform in Australia. *Patient*. 2010;3:275–283.

5. Sudman S, Bradburn NM, Wansick B. *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco, CA: Jossey-Bass; 1982.

6. Feather NT. The measurement of values: effects of different assessment procedures. *Aust J Psychol*. 1973;25:221–231.

7. Schulman S, Kearon C. Subcommittee on control of anticoagulation of the scientific and standardization committee of the international society

on Thrombosis and Haemostasis. Definition of major bleeding in clinical investigations of anti-hemostatic medicinal products in non-surgical patients. *J Thromb Haemostasis*. 2005;3:692-694.

8. Kaatz S, Ahmad D, Spyropoulos AC, Schulman S, Subcommittee on Control of Anticoagulation. Definition of clinically relevant non-major bleeding in studies of anticoagulants in atrial fibrillation and venous thromboembolic disease in non-surgical patients: communication from the SSC of the ISTH. *J Thromb Haemostasis*. 2015;13: 2119-2126.

9. Saver JL, Chaisinanunkul N, Campbell BCV, et al. Standardized nomenclature for modified Rankin Scale global disability outcomes: consensus recommendations from stroke therapy academic Industry roundtable XI. *Stroke*. 2021;52: 3054-3062.

10. Okumura K, Inoue H, Yasaka M, et al. Comparing patient and physician risk tolerance for bleeding events associated with anticoagulants in atrial fibrillation-evidence from the United States and Japan. *Value Health Reg Issues*. 2015;6:65-72.

11. Edwards NT, Greanya ED, Fan Kuo I, Loewen PS, Culley CL. Patient preferences regarding atrial fibrillation stroke prophylaxis in patients at potential risk of atrial fibrillation. *Int J Clin Pharm*. 2017;39:468-472.

12. Lahaye S, Regpala S, Lacombe S, et al. Evaluation of patients' attitudes towards stroke prevention and bleeding risk in atrial fibrillation. *Thromb Haemostasis*. 2014;111:465-473.

13. Smith NF, Street DJ. The use of balanced incomplete block designs in designing randomized response surveys. *Aust N Z J Statistics*. 2003;45(2):181-194.

14. Gonzalez JM. A Guide to measuring and interpreting attribute importance. *Patient*. 2019;12:287-295.

15. Krinsky I, Robb AL. On approximating the statistical properties of elasticities. *Rev Econ Stat*. 1986;68:715-719.

16. Robinson A, Thomson R, Parkin D, Sudlow M, Eccles M. How patients with atrial fibrillation value different health outcomes: a standard gamble study. *J Health Serv Res Policy*. 2001;6:92-98.

17. Loewen PS, Ji AT, Kapanen A, McClean A. Patient values and preferences for antithrombotic therapy in atrial fibrillation. A Narrative Systematic Review. *Thromb Haemostasis*. 2017;117:1007-1022.

18. Wright JN, Vazquez SR, Kim K, Jones AE, Witt DM. Assessing patient preferences for switching from warfarin to direct oral anticoagulants. *J Thromb Thrombolysis*. 2019;48:596-602.

19. Shafrin J, Bruno A, MacEwan JP, et al. Physician and patient preferences for nonvalvular atrial fibrillation therapies. *Value Health*. 2016;19:451-459.

20. Ozga A, Rauch G. Introducing a new estimator and test for the weighted all-cause hazard ratio. *BMC Med Res Methodol*. 2019;19:1-16.

21. Bakal J, Westerhout C, Armstrong P. Impact of weighted composite compared to traditional composite endpoints for the design of randomized controlled trials. *Stat Methods Med Res*. 2015;24: 980-988.

KEY WORDS atrial fibrillation, net clinical benefit, patient-centered drug development, patient preferences

APPENDIX For supplemental tables and the PEARL-AF Survey Instrument, please see the online version of this paper.