Structured Review of Patient-Reported Outcome Instruments for Assessing Atrial Fibrillation

Amy Barrett,[1] Dana DiBenedetti,[1] Hemant Phatak,[2] Uchenna Iloeje[2]

[1] RTI Health Solutions, Research Triangle Park, NC, United States; [2] Bristol-Myers Squibb, Princeton, NJ, United States

BACKGROUND

Atrial Fibrillation

- Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia, with an overall prevalence of 1% and a prevalence of approximately 10% in patients aged 80 years and older.[1]
- Significant increased morbidity and mortality are observed in patients with AF. Symptoms associated with AF are primarily caused by rapid and irregular heartbeat and include palpitations, shortness of breath, dizziness, anxiety, and reduced physical capacity, which result in impaired health-related quality of life (HRQOL).[2] Sequelae of AF include thromboembolic events and precipitation or worsening of heart failure.
- Multiple treatment options, including pharmacotherapy and ablative strategies, must be carefully evaluated on a patient-by-patient basis, because treatment options entail risk and may have limited efficacy depending on the subtype of AF.[1]
- Because few interventions have been shown to reduce mortality and serious morbidity, the assessment of patient-reported outcomes (PROs), including HRQOL and symptom severity or frequency, is crucial in the treatment of AF and in the study of new therapies for AF. Furthermore, in AF clinical trials of antiarrhythmic therapies, endpoints focused on the maintenance of heart rhythm may not accurately reflect the degree to which patients' AF symptoms improve.[3] Moreover, limited information is available to determine which instruments (if any) are available for assessing PROs in AF patients treated with antiarrhythmic therapy.

FDA Guidance for PRO Measures

- The United States (US) Food and Drug Administration (FDA) has issued detailed guidance on the requirements for PRO measures that are to be used to support regulatory approval or promotional claims.[4]
- The guidance, developed with input from the FDA's Study Endpoints and Label Development (SEALD) group, describes both the recommended development of a PRO measure and the psychometric properties for which evidence must be presented.
- This guidance clearly stipulates that any PRO measure used to support labeling or promotional claims must be developed with extensive input from patients to establish content validity and be thoroughly validated in the target population.

OBJECTIVE

- To identify and evaluate the key characteristics, strengths, and weaknesses of existing AF-specific PRO measures, focusing on how well the measures meet current regulatory guidance requirements set out by the FDA in its PRO Guidance for Industry.[4]

METHODS

This review was conducted in two phases.

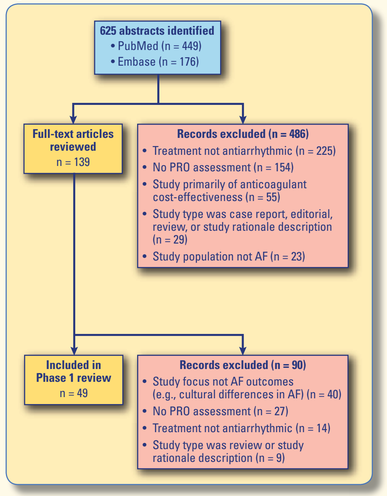
- In phase 1, a comprehensive review of multiple sources (PubMed, Embase, Patient-Reported Outcome and Quality of Life Instruments Database [PROQOLID], ClinicalTrials.org) was conducted to identify potential PRO measures, including measures assessing AF-specific HRQOL, symptoms, functional status, treatment satisfaction, or other patient-reported domains. Literature database searches were limited to studies published in English and describing research in humans. Clinical trials were limited to those including an antiarrhythmic treatment of AF.
- In phase 2, data related to the development and measurement properties of the instruments meeting the prespecified instrument criteria were gathered and compared. These data were sought through additional searches of PubMed and information provided by the instrument developers, either online or upon request.

RESULTS

The results of the initial phase 1 review of multiple sources included the following:

- PubMed and Embase: 625 abstracts identified; 139 full-text articles reviewed and 49 included (Figure 1)
- ClinicalTrials.gov: 55 antiarrhythmic AF trials identified; 13 appeared to include AF-specific PROs
- PROQOLID: 4 PRO measures related to AF or arrhythmia identified
- In total, 25 PRO measures with use in AF were identified, including 7 AF-specific measures, 5 generic HRQOL measures, 2 measures of functional status in cardiology conditions, and 11 measures developed for use in other conditions (e.g., measures assessing mood or illness intrusiveness)

Figure 1. Reference Source Flowchart for Phase 1 Literature Review



Note: The initial searches were conducted June 2012.

The results of the phase 2 review of measures included the following:

- Among the 7 AF-specific measures appearing to meet the inclusion criteria (patient-reported assessments of AF-specific HRQOL, symptoms, functional status, treatment satisfaction, or other patient-reported domains), 1 measure, the Canadian Cardiovascular Society Severity of Atrial Fibrillation (CCS-SAF) Scale,[5] was determined to be a clinician-reported measure and was excluded.
- Six PRO measures (Table 1) were the basis of more in-depth searches. During phase 2, 15 additional studies were selected for full-text review, but no additional studies were included.

Table 1 presents an overview of the six included measures, all of which were primarily assessments of AF-related HRQOL or AF symptoms.

Table 1. Overview of AF-Specific PRO Measures of Interest

Objective	Recall Period	Domains/Items
Primarily HRQOL		
AF6[6]		
Evaluates patient-reported symptoms and symptom impact of persistent AF before and after direct-current cardioversion or during evaluation for pharmacological versus nonpharmacological treatment	7 days	6 items • Breathing difficulties at rest • Breathing difficulties on exertion • Limitations in day-to-day life due to AF • Feeling of discomfort due to AF • Tiredness due to AF • Worry or anxiety due to AF
Atrial Fibrillation Effect on Quality-of-Life (AFEQT)[7]		
AF-specific HRQOL questionnaire assessing the impact of AF and its treatment on patient symptoms, functioning, and daily activities	4 weeks	20 items in 4 domains • Symptoms (4) • Daily activities (8) • Treatment concern (6) • Treatment satisfaction (2)
Atrial Fibrillation Quality of Life (AF-QoL)[8]		
AF-specific HRQOL questionnaire for patients with any type of AF	4 weeks	18 items in 3 domains[g] • Psychological (7) • Physical (8) • Sexual activity (3)
Questionnaire for Quality of Life in Atrial Fibrillation (QLAF)[10]		
AF-specific questionnaire for HRQOL and symptom assessment Interviewer administered (patient reported but not patient completed)	Not specified	28 items in 7 domains • Palpitation (6) • Breathlessness (4) • Chest pain (5) • Dizziness (5) • Drugs (4) • Direct-current cardioversion (2) • Ablation (2)
Primarily Symptom Assessment		
University of Toronto Atrial Fibrillation Symptom Severity (AFSS)[11]		
AF-specific scale developed to capture subjective and objective ratings of AF disease burden, including frequency, duration, and severity of episodes, and health care utilization	Varies by item	19 items • Total AF burden = AF frequency + AF duration + AF severity (4) • Global well-being (1) • AF symptoms (bothersomeness) (7) • Health care utilization (4) • Demographic data (2) • Current AF status (1)
Symptom Checklist—Frequency and Severity (SCL)[12]		
Arrhythmia-related symptom assessment developed in the late 1990s for evaluating the impact of early catheter ablation and pacing technologies on a variety of arrhythmias[3]	4 weeks	16 items (symptoms associated with AF) Respondents rate frequency (from 0 to 4) and severity (from 1 to 3) of each symptom Frequency and severity scores are not combined

Table 2 provides a summary of the type of patient involvement documented in the development of the AF-specific instruments.

Table 2. Summary of Content Validity Characteristics Based on Instrument Development

Involvement of Target Population	AF6	AFEQT	AF-QoL	QLAF	AFSS	SCL
Item generation/ modification[a]	✓	✓	✓	—	—	—
Evaluation of item completeness and acceptability[b]	—	✓	—	—	—	—
Item-reduction process[c]	—	✓	—	—	—	—

✓ = Yes; — = No (not reported or not adequately documented).

[a] Individual interviews or focus groups were conducted with target population.

[b] Through pilot testing, feasibility testing, or cognitive debriefing with individual interviews, the target population evaluated the completeness of item coverage and performed an initial assessment of clarity and readability.

[c] Item reduction was based on content analysis of feedback from members of the target population.

Table 3 presents a summary of the evaluated measurement properties of each of the AF-specific instruments of interest.

Table 3. Summary of Psychometric Properties Reported in the Literature for AF-Specific PRO Instruments of Interest

Psychometric Property	Instrument (Year Introduced)					
	Published After the Draft FDA PRO Guidance[13]				Published Before the Draft FDA PRO Guidance[13]	
	AF6 (2009)	AFEQT (2011)	AF-QoL (2007)	QLAF (2010)	AFSS (1998)	SCL (1996)
Internal consistency[a]	✓	✓	✓	✓	✓	NR
Test-retest reliability[b]	—	✓	✓	✓	✓	NR
Content validity[c]	✓	✓	✓	—	—	—
Construct validity, convergent[d]	✓	✓	✓	NR	✓	NR
Construct validity, divergent[d]	NR	✓	NR	NR	NR	NR
Discriminant validity[e]	NR	✓	✓	NR	NR	NR
Responsiveness, longitudinal validation study[f]	✓	✓	✓	✓	NR	NR
Responsiveness, RCT[g]	—	—	—	—	✓	✓

NR = not reported; RCT = randomized clinical trial.

✓ = Instrument achieved or exceeded the established psychometric standard or the standard set by the authors of this review (see notes for the specific standard for each property).

— = Instrument did not meet the established psychometric standard or the standard set by the authors of this review (see notes for the specific standard for each property).

[a] Range for acceptable Cronbach's alpha: above 0.70 but not higher than 0.95.[14]

[b] Threshold for acceptable test-retest reliability: interclass correlation coefficient of 0.70 or greater.[15]

[c] Target population (patients with AF) provided documented input in the development of the instrument in one or more of the following areas: generation of item concept and wording, evaluation of completeness of item coverage, or assessment of item clarity and readability.

[d] At least one Pearson's correlation coefficient (r) value was categorized as moderate (0.10-0.50) or strong (> 0.50).[16]

[e] Discriminant validity demonstrated by statistically significant ($P < 0.05$) difference in at least one comparison of patient subgroups with differing clinical features.

[f] Responsiveness demonstrated by statistically significant ($P < 0.05$) results in at least one longitudinal validation study.

[g] Responsiveness demonstrated by statistically significant ($P < 0.05$) results in at least one RCT.

DISCUSSION

- The two symptom measures, the SCL and AFSS, were developed in the late 1990s, before the introduction of the current FDA PRO guidelines. These measures have the least available evidence in support of their psychometric properties.
- The HRQOL measures were all developed since the draft FDA PRO guidance was published in 2006.
- The level of patient involvement for the AFEQT appears to meet the standards set forth to establish content validity in the FDA PRO guidance. The AF6 and AF-QoL involved patients in the initial generation of items, which is an important step and provides a basis of evidence for content validity.
- The AFEQT has demonstrated achievement in the greatest number of the evaluated psychometric properties, although to date it has not been included in a clinical trial in AF, so no evaluation of responsiveness in that setting was possible.

AF-Specific HRQOL

- It is unlikely that the HRQOL instruments reviewed, in their current form, would be acceptable to the FDA to support a PRO promotional label claim. All of the measures likely would face challenges related to relatively long recall periods and multidimensional assessment of a complex measurement concept (HRQOL).
- The AF6, AF-QoL, and QLAF were developed in Sweden, Spain, and Brazil, respectively, and have been validated only in studies conducted in these countries.
- The AFEQT appears to be the strongest available instrument for measuring HRQOL in AF, with the most rigorous and well-documented development and most successful demonstration of measurement properties, including reliability, content and construct validity, and responsiveness.
- Additional studies confirming the AFEQT's measurement properties are needed, given that currently only one study[7] presents all of the relevant development and validation data, and an instrument's properties ideally are demonstrated during repeated use and evaluation.

AF Symptom Assessment

- Neither the SCL or AFSS symptom assessments would be acceptable to the FDA to support a PRO promotional label claim without evidence of adequate psychometric properties and content validity.
- The SCL has no development or validation information in AF but has been widely used, particularly in trials of AF surgical interventions.
- The AFSS has limited published psychometric validation related to test-retest and internal consistency reliability and has been frequently used in clinical and observational studies.
- The AFSS and SCL have been used with similar frequency in trials of AF antiarrhythmic therapy, and both displayed some responsiveness, though generally they did not detect between-group differences related to treatment. The evidence of responsiveness for these measures is limited, and the results are mixed.

CONCLUSIONS

- Use of a PRO measure that meets the standards of the FDA PRO guidance in a clinical trial may result in the potential for a PRO label claim. If included in a drug product label, AF-specific PRO results may be used in promotional materials. Data appearing in an FDA label can be used (without risk) to support promotional activities.
- There does not appear to be any AF-specific HRQOL or symptom measure that would be likely to support an FDA PRO label claim in its current form.

REFERENCES

Please see handout.

CONTACT INFORMATION

Amy Barrett, MSPH, MA
Director, Patient-Reported Outcomes

RTI Health Solutions
200 Park Offices Drive
Research Triangle Park, NC 27709

Phone: +44 7850774732
E-mail: abarrett@rti.org

Presented at: ISPOR 18th Annual International Meeting
May 18-22, 2013
New Orleans, LA, United States

FUNDING

Funding for this study and development of this poster provided by Bristol-Myers Squibb, Princeton, NJ, United States.