

# The European Heart Rhythm Association symptom classification for atrial fibrillation: validation and improvement through a simple modification

Gareth J. Wynn<sup>1,2,3\*</sup>, Derick M. Todd<sup>1,2,3</sup>, Matthew Webber<sup>1,2</sup>, Laura Bonnett<sup>4</sup>, James McShane<sup>2</sup>, Paulus Kirchhof<sup>5,6</sup>, and Dhiraj Gupta<sup>1,2,3</sup>

<sup>1</sup>Institute of Cardiovascular Medicine and Science, Liverpool and London, UK; <sup>2</sup>Liverpool Heart and Chest Hospital, Liverpool L14 3PE, UK; <sup>3</sup>National Heart and Lung Institute, Imperial College London, London SW3 6LY, UK; <sup>4</sup>Department of Biostatistics, Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, Liverpool L69 3GA, UK; <sup>5</sup>University of Birmingham Centre for Cardiovascular Sciences and SWBH NHS trust, Birmingham B15 2TT, UK; and <sup>6</sup>Department of Cardiology and Angiology, Hospital of the University of Münster, Münster, Germany

Received 4 October 2013; accepted after revision 18 November 2013; online publish-ahead-of-print 16 February 2014

## Aims

To validate the European Heart Rhythm Association (EHRA) symptom classification in atrial fibrillation (AF) and test whether its discriminative ability could be improved by a simple modification.

## Methods and results

We compared the EHRA classification with three quality of life (QoL) measures: the AF-specific Atrial Fibrillation Effect on Quality-of-life (AFEQT) questionnaire; two components of the EQ-5D instrument, a health-related utility which can be used to calculate cost-effectiveness, and the visual analogue scale (VAS) which demonstrates patients' own assessment of health status. We then proposed a simple modification [modified EHRA (mEHRA)] to improve discrimination at the point where major treatment decisions are made. Quality of life data and clinician-allocated EHRA class were prospectively collected on 362 patients with AF. A step-wise, negative association was seen between the EHRA class and both the AFEQT and the VAS scores. Health-related utility was only significantly different between Classes 2 and 3 ( $P < 0.001$ ). We developed and validated the mEHRA score separating Class 2 (symptomatic AF not limiting daily activities), based on whether the patients were 'troubled by their AF' (Class 2b) or not (Class 2a). This produced two distinct groups with lower AFEQT and VAS scores and, importantly, both clinically and statistically significant lower health utility ( $\Delta$ utility 0.9,  $P = 0.01$ ) in Class 2b than Class 2a.

## Conclusion

Based on patients' own assessment of their health status and the disease-specific AFEQT, the EHRA score can be considered a useful semi-quantitative classification. The mEHRA score has a clearer separation in health utility to assess the cost efficacy of interventions such as ablation, where Class 2b symptoms appear to be the appropriate treatment threshold.

## Keywords

Atrial fibrillation • Symptoms • Quality of life • Symptom score • EHRA

## Introduction

Atrial fibrillation (AF) is an emerging medical epidemic affecting 1–2% of the general population and up to 15% of elderly patients.<sup>1–4</sup>

As well as doubling mortality rate, there is a significant effect on morbidity due to symptoms such as palpitation, fatigue, dyspnoea, and exercise intolerance.<sup>5,6</sup> The resultant impact on quality of life (QoL) is as dramatic as for patients with severe structural heart disease.<sup>7</sup> Up to

\* Corresponding author. Research Unit, Liverpool Heart and Chest Hospital, Thomas Drive, Liverpool L14 3PE, UK. Tel: +44 151 600 1016; fax: +44 151 600 1617, Email: gjw@doctors.org.uk

© The Author 2014. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

### What's new?

- The EHRA score has been validated as a semi-quantitative measure of AF related symptoms and patients' perception of their general state of health.
- A simple modification of the EHRA score (the mEHRA score) provides better discrimination for patients with mild to moderate symptoms.
- Patients with mild symptoms, which the patient does not find troublesome, have a Health-related Quality of Life comparable to asymptomatic patients suggesting they may not be appropriate candidates for intervention.

one-third of patients present with limiting symptoms, and inpatient care is a major consumer of healthcare resources.<sup>8</sup>

In 2007, an expert group of the German Atrial Fibrillation Competence NETwork (AFNET) and the European Heart Rhythm Association (EHRA) published recommendations for the conduct of clinical trials in AF.<sup>9</sup> Noting that no accepted and easily applicable measure for the AF-related symptoms exists, this group of experts proposed and described in their recommendations a new scoring system, the EHRA classification, to assess and quantify symptoms related to AF.<sup>9</sup> The EHRA classification is based on the impact of symptoms on daily activity during presumed episodes of AF. It is simple to use and has a format similar to the New York Heart Association (NYHA) symptom classification system for patients with heart failure, making it relatively intuitive for practicing clinicians while hopefully being more appropriate to AF-related symptoms.<sup>10</sup> When the EHRA score was published, the authors made a specific mention of the need for validation.<sup>9</sup> A similar score was thereafter validated by a Canadian group.<sup>6</sup> Since its initial proposal, the EHRA score has entered widespread use and has even been used in the European Society of Cardiology guidelines for the management of AF in the recommendations for rate and rhythm control.<sup>11–16</sup>

To put the EHRA score into context with existing tools to assess disease-related QoL in AF patients, we compared the EHRA score with accepted and validated measures of health-related QoL. Specifically, we used both a disease specific tool, the Atrial Fibrillation Effect on QualiTY-of-life (AFEQT) questionnaire, and a general tool, the very well-established EQ-5D questionnaire.<sup>17–19</sup> The AFEQT is very sensitive to changes in symptom burden but cannot be compared with other conditions. EQ-5D, however, is applicable to a wide range of health conditions and provides a single index value for health status, called the health utility, which can be used to calculate 'quality-adjusted life years' (QALYs) for health economic evaluation.<sup>20</sup> Previous studies have suggested that the minimal meaningful difference in the EQ-5D-derived health utility is 0.07.<sup>20</sup> In line with this, a previous study of the cost-effectiveness of catheter ablation in the UK setting found catheter ablation to be cost-effective with an estimated utility difference of 0.09 between symptomatic AF on drugs compared with sinus rhythm following ablation.<sup>21</sup>

We aimed to validate the EHRA score by using these general and disease-specific QoL measures. We hypothesized that the discriminative power of the EHRA score could be improved and attempted to achieve this through a simple modification.

**Table 1** Modified EHRA (mEHRA) classification

mEHRA score	Symptoms	Description
1	None	
2a	Mild	Normal daily activity not affected, <u>symptoms not troublesome to patient</u>
2b	Moderate	Normal daily activity not affected <u>but patient troubled by symptoms</u>
3	Severe	Normal daily activity affected
4	Disabling	Normal daily activity discontinued

Underlined text represents the modification to the original descriptions of EHRA classes.

## Methods

### Phase 1

Consecutive patients with a diagnosis of AF attending designated heart rhythm/electrophysiology clinics at a single specialist cardiac hospital in England (Liverpool Heart and Chest Hospital) were invited to complete the AFEQT questionnaire and the EQ-5D instrument, including the visual analogue scale (VAS).<sup>18,19</sup> At the same clinical visit, the reviewing clinician was asked to independently score the patient according to the EHRA symptom classification. In keeping with real-world practice, these assessments were completed by a range of clinicians experienced in the management of patients with AF, including consultant cardiologists, trainee physicians, and arrhythmia nurses. Clinicians were provided with, and requested to complete, a classification form that listed the published definitions for each class. To replicate how the classification is likely to be used in routine practice, no specific training was provided beyond that described above and access to the original publication in which the EHRA score was proposed. This mirrored the methods used to validate an alternative classification system.<sup>6</sup>

Quality of life was assessed by the AFEQT score (global), the health-related utility based on the EQ-5D instrument, and the VAS. For each measure, a higher score represents a higher QoL. AFEQT and VAS are scored from 0 to 100. Health utility (EQ-5D) ranges between 1 (perfect health) through 0 (death) to -0.59 (worse than death). Mean QoL scores were compared between neighbouring EHRA classes to assess the classification system's accuracy in semi-quantifying QoL.

### Phase 2

We proposed a modified EHRA (mEHRA) classification by subdividing Class 2 into Classes 2a (mild) and 2b (moderate) according to the degree to which the patient was 'troubled by their symptoms' (Table 1). All the patients categorized as Class 2 during Phase 1 of the study were independently re-categorized as either Class 2a or 2b by two clinicians (electrophysiology fellows) who were both blinded to the corresponding QoL scores. Clinical letters were reviewed with specific attention made to the extent to which the patients appeared 'troubled by their symptoms', given that their daily activities were not affected (which would indicate Class 3 symptoms). We sought to assess the patient's own perception of the impact of AF on their well-being. Those suffering from anxiety, loss of confidence, or symptoms that they found unpleasant were graded as Class 2b. There was agreement between the reviewers for all the cases reclassified into mEHRA

Classes 2a and 2b. Where the clinical letters were lacking in detail, the hospital records were re-reviewed and a consensus agreed upon. The mEHRA score was assessed and validated using the same methods used to validate the EHRA score in the initial phase of the study.

### Phase 3

Having completed Phases 1 and 2, we wished to verify and validate the findings from our retrospective scoring of the mEHRA score, by comparing our findings with an independent cohort in whom the mEHRA score had been used prospectively. The method of data collection for this phase matched exactly with that in Phase 1, except that the clinicians were asked to classify symptoms according to the new, expanded mEHRA score rather than the original EHRA score. We compared the prospective and the retrospective scores, for the two new mEHRA classes, for each of the three QoL measures used in the previous two phases.

### Atrial Fibrillation Effect on Quality-of-Life questionnaire

The AFEQT questionnaire is a validated, disease-specific, self-administered QoL instrument.<sup>17</sup> It has 20 questions with four conceptual domains: symptoms (four questions specifically targeted to assess AF-related symptoms), Treatment Concerns (six questions that assess AF treatment concerns in patients), Daily Activities (eight questions that evaluate daily function in AF patients), and Treatment Satisfaction (two questions asking about how well the current treatment controls their AF and relieves symptoms). Each of the 20 questions is marked on a 7-point Likert scale. A published algorithm exists to allow the calculation of a score, between 0 and 100 (where higher is better), for each domain and a global score based on the first three domains. The questionnaire specifies a recall period of the preceding 4 weeks.

### EQ-5D questionnaire (Three-level version)

EQ-5D is a standardized measure of health status developed by the EuroQol Group to provide a simple, generic measure of health for clinical and economic appraisal.<sup>18</sup> Two hundred and forty-three possible health states can be defined from the five questions and these can be converted into a single summary index called the 'health utility' based on country-

specific value sets.<sup>22</sup> The health utility measure is of particular interest as it is generalizable to other diseases and to the general population. Therefore, it allows the calculation of QALYs and thus the cost-effectiveness of interventions such as catheter ablation.

### Visual analogue scale

The VAS forms an integral but distinct part of the EQ-5D instrument. Respondents are asked to mark a single point on a linear scale that represents their health status on the day of completion. The scale extends from 0 (worst imaginable health state) to 100 (best imaginable health state). The VAS provides a quantitative measure of state of health, as judged by the individual patients.

### Statistical analysis

Analyses were performed using StatsDirect software version 2.7.8. Analysis of QoL scores (AFEQT score, EQ-5D-derived utility, and VAS) using the Shapiro–Wilk *W* test showed non-Gaussian distribution that was not corrected by logarithmic transformation. However, the central limit theorem is generally taken to imply that an assumption of normality is not necessary for parametric testing to be valid if the group sizes are greater than about 30. This is the case for all but one group in our analysis. In addition, presentation of the data as means with standard deviation was adjudged to be of greater clinical relevance than presentation of medians with the inter-quartile ranges. Continuous data are therefore presented as mean  $\pm$  standard deviation and compared using the *t*-test. A similar approach has been used in previous work using QoL measures.<sup>23</sup> A two-tailed test was used where there was no *a priori* expectation of the direction of the difference between the groups. However, where a comparison was made between two adjacent groups, resulting in assessment for a difference in a single direction only, one-tailed testing was considered more appropriate. Analysis of variance was compared by using one-way analysis of variance (ANOVA). Trend across groups was assessed by using Cuzick's test. Proportions were compared by using the  $\chi^2$  statistic. Intra- and inter-observer variability were calculated, and assessed by using the  $\kappa$  statistic. A *P* value of 0.05 or lower was considered significant for all tests.

**Table 2** Baseline characteristics by EHRA class

	EHRA 1	EHRA 2	EHRA 3	EHRA 4	Total	<i>P</i> value for trend
<i>N</i> , number	149	99	90	24	362	n/a
Age (mean $\pm$ SD)	61.1 ( $\pm$ 11.4)	59.7 ( $\pm$ 12.2)	57.9 ( $\pm$ 13.0)	62.2 ( $\pm$ 9.9)	60.0 ( $\pm$ 12.0)	NC
Male gender	70.3%	58.6%	73.0%	62.5%	66.9%	NC
% PAF <sup>a</sup>	49.0%	65.6%	62.2%	41.7%	56.4%	0.36
Diabetes	6.8%	7.1%	8.8%	20.0%	8.3%	NC
Heart failure	5.4%	3.0%	7.7%	12.0%	5.8%	NC
Hypertension	54.4%	36.4%	46.2%	72.0%	48.6%	0.003
Pacemaker	5.4%	7.1%	9.9%	12.0%	7.5%	NC
COPD	4.8%	6.1%	8.8%	4.0%	6.1%	NC
Previous stroke/TIA	8.2%	6.1%	13.2%	12.0%	9.1%	NC
Prior ablation	77.6%	52.5%	25.3%	24.0%	53.9%	<0.0001
Subsequent ablation	6.8%	34.3%	61.5%	76.0%	32.9%	<0.0001

NC, not calculated (where ANOVA revealed no significant variance between groups, a test for trend was not performed).

<sup>a</sup>Predominant pattern at time of assessment.

COPD, chronic obstructive pulmonary disease; TIA, transient ischaemic attack; PAF, paroxysmal atrial fibrillation. The term 'ablation' is used to denote catheter ablation of atrial fibrillation (e.g. pulmonary vein isolation).

The study was performed as part of a wider institutionally approved patient reported outcome measures service improvement programme at the recruiting centre.

## Results

Quality of life and symptom data were collected on 362 patients attending the heart rhythm clinics during 2012. All patients received physician-allocated EHRA classification at the same clinical visit. The baseline characteristics are given in Table 2. There were no clear differences between the groups in terms of age, sex, or proportion classified as having paroxysmal, as opposed to persistent or permanent, AF. Previous studies have found QoL to be lower in females.<sup>24</sup> However, in our cohort, patient gender did not have a significant effect on any of the three QoL measures (EQ-5D  $P = 0.56$ , VAS  $P = 0.70$ , and AFEQT  $P = 0.14$ ). Hypertension was by far the most common comorbid condition (48.6% of all patients) and showed a

significant trend towards increasing prevalence in higher EHRA classes ( $P = 0.003$ ). Other comorbidities were infrequently present and were similar across classes. However, there was a clear trend seen whereby those in less severe symptom classes were the most likely to have previously undergone ablation and those in the most severe classes were considerably more likely to subsequently go on to have an ablation for atrial fibrillation. ( $P < 0.0001$  for both). A small proportion of asymptomatic patients went on to have an ablation in keeping with the published data from the EUR Observation study where 13% of patients were asymptomatic, citing a desire for a drug free lifestyle, improved QoL, and/or the maintenance of sinus rhythm.<sup>25</sup>

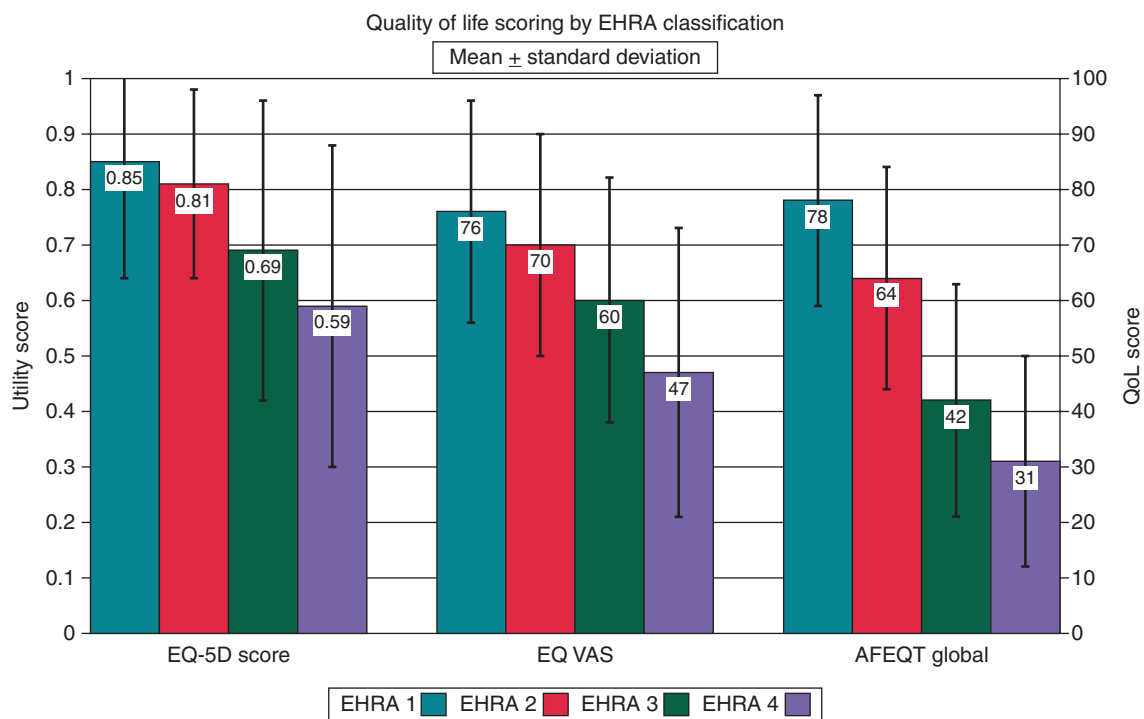
## Phase 1

Results for the three QoL measures are shown in Table 3 and Figure 1. Analysis of variance and regression analysis confirmed significant

**Table 3** Mean (and standard deviation) shown for each EHRA class

EHRA class	Utility (by EQ-5D)	P value	VAS	P value	AFEQT	P value
1	0.85 (± 0.21)	n/a	76.2 (± 19.9)	n/a	78.4 (± 19.0)	n/a
2	0.81 (± 0.17)	0.08	70.3 (± 20.3)	0.02	63.6 (± 20.0)	<0.0001
3	0.69 (± 0.27)	<0.001	59.6 (± 21.9)	<0.001	42.1 (± 21.1)	<0.0001
4	0.59 (± 0.29)	0.08	46.9 (± 25.9)	0.03	31.3 (± 18.6)	0.01

P values compare each class with the next lowest class in terms of symptom severity. VAS, visual analogue scale; AFEQT, AFEQT global score.



**Figure 1** Quality of life (QoL) scores (mean ± standard deviation) by EHRA class.

negative correlation between the EHRA class and QoL as assessed by all three measures. To determine the ability of the EHRA classification as a semi-quantitative tool, each EHRA class was compared with the class immediately below in terms of QoL (i.e. EHRA 2 with EHRA 1, EHRA 3 with EHRA 2, and EHRA 4 with EHRA 3). On using the disease-specific AFEQT score, significant differences were seen at each grade boundary suggesting that the EHRA score was an effective means of categorizing patients' symptoms. Similarly, on using the patient-based VAS, there was a significant difference between each and its immediate neighbour suggesting that the EHRA score effectively categorized patients in terms of their own assessment of their health state. However, when comparing the health-related utility, derived from the EQ-5D questionnaire, although there was a significant difference of 0.12 ( $P < 0.001$ )

between Classes 2 and 3, the difference between Classes 1 and 2 was only 0.04 ( $P = 0.08$ ). This observation prompted us to develop the mEHRA classification.

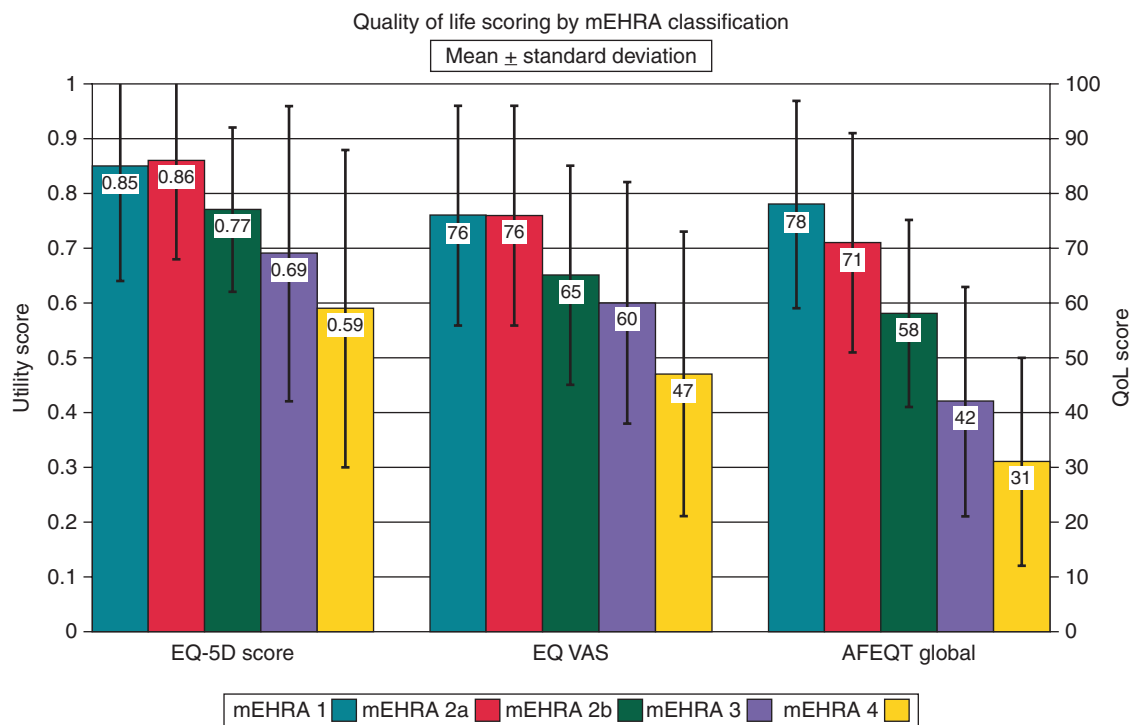
### Phase 2

Of the 99 patients originally classified as EHRA Class 2, 90 had sufficient detail in their archived clinical letters to be reclassified in accordance with the proposed mEHRA classification into either Class 2a or 2b. Of these, 43 were classified as Class 2a and 47 were classified as Class 2b. The subdivision of Class 2 into Classes 2a and 2b, resulted in clearly separate groups, with the Class 2a patients having AFEQT scores and health utilities much closer to the Class 1 patients and the Class 2b patients having significantly more symptoms as judged by the AFEQT and a significant reduction in the

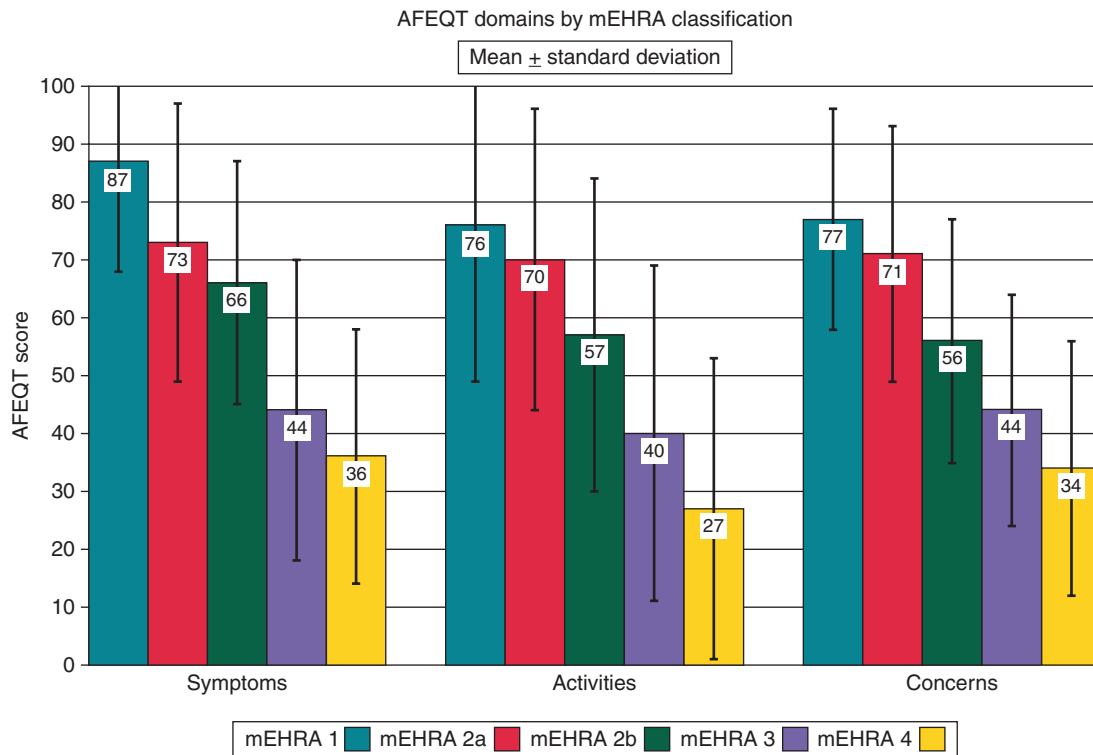
**Table 4** Mean (and standard deviation) shown for each mEHRA class

mEHRA class	Utility (by EQ-5D)	P value	VAS	P value	AFEQT	P value
1	0.85 (± 0.21)	n/a	76.2 (± 19.9)	n/a	78.4 (± 19.0)	n/a
2a	0.86 (± 0.18)	0.41	75.6 (± 19.9)	0.43	70.9 (± 19.8)	0.01
2b	0.77 (± 0.15)	0.01	65.2 (± 20.1)	0.01	58.3 (± 17.3)	<0.001
3	0.69 (± 0.27)	0.02	59.6 (± 21.9)	0.09	42.1 (± 21.1)	<0.0001
4	0.59 (± 0.29)	0.08	46.9 (± 25.9)	0.03	31.3 (± 18.6)	0.01

P values compare each class with the next lowest class in terms of symptom severity. VAS, visual analogue scale; AFEQT, AFEQT global score.



**Figure 2** Quality of life (QoL) scores (mean ± standard deviation) by mEHRA class. Classes 1, 3, and 4 are as for Figure 1. Class 2 has been split into Classes 2a (n = 43) and 2b (n = 47).



**Figure 3** AFEQT sub-domain scores (mean  $\pm$  standard deviation) by mEHRA class.

health utility as judged by the EQ-5D. The results are shown in *Table 4* and *Figure 2*.

We also analysed the mEHRA class according to the three subdomains of the AFEQT score: Symptoms, Activities, and Concerns. As can be seen in *Figure 3*, there is a clear step-wise trend to lower scores as the mEHRA class increases. As these components of the AFEQT global score are not themselves individually validated we did not seek to analyse these on a class-by-class basis. However, the ANOVA confirmed a highly significant difference between the groups for all three subdomains ( $P < 0.0001$  for each) and Cuzick's trend test showed a strong, and highly significant, trend towards lower scores with increasing mEHRA class ( $P < 0.0001$  for each of the three subdomains).

### Reproducibility of the modified European Heart Rhythm Association class

To assess the reproducibility of the mEHRA score we measured both intra-observer and inter-observer variability. Agreement between the two assessors for the ratings of Class 2a or 2b was very good at 83.2% ( $\kappa$  0.70, 95% confidence interval (CI) = 0.53–0.87). Inter-observer variability was assessed by asking each assessor to reclassify a random sample of 20 of the original clinical letters used for Phase 2 of the study. This was performed after an interval of several months to avoid bias due to recall of the previous classification. This demonstrated excellent repeatability, with an agreement between the original and the repeat classification of 90% for one assessor and 95% for the other (combined  $\kappa$  0.85, 95% CI for  $\kappa$ , 0.54–1.16).

### Phase 3: prospective validation of the modified European Heart Rhythm Association

The mEHRA score was thereafter prospectively applied to a second cohort of patients attending the Heart Rhythm Clinics at Liverpool Heart and Chest Hospital. By using the modified scoring system, 165 patients were classified as either Class 2a ( $n = 85$ ) or 2b ( $n = 80$ ). These data are shown in *Table 5*. Prospective scoring showed the same pattern as retrospective scoring with significantly lower AFEQT, EQ-5D, and VAS scores for mEHRA Class 2b than Class 2a. There were no significant differences between any of the three QoL measures between this validation cohort and the initial cohort in either of the two new mEHRA classes (2a and 2b).

### Discussion

This comparison of the EHRA symptoms classification with one disease-specific QoL instrument (AFEQT) and another general measure for the health-related QoL (EQ-5D, incorporating the VAS) provides good evidence that the EHRA score can be used to assess AF-related symptoms without prior training. It is easily applied in clinical practice and has the potential to be useful in clinical trials. However, it does not discriminate sufficiently in patients with low-level symptoms in terms of the health-related utility. We found that the EHRA classification can be improved by subdividing Class 2 into two separate classes (2a and 2b).

Phase 1 of this study provides evidence that the EHRA classification, in its originally proposed format correlates well with the



**Table 5** Comparison of retrospective and prospective QoL scores, for each of the two proposed additional mEHRA classes

QoL measure		Retrospective n = 90	Prospective n = 165	P value Retrospective vs. prospective
Utility (EQ-5D)	2a	0.86 (± 0.18)	0.81 (± 0.22)	0.23
	2b	0.77 (± 0.15)	0.72 (± 0.22)	0.25
P value 2a vs. 2b		P = 0.01	P < 0.0001	
VAS	2a	75.6 (± 19.9)	77.9 (± 15.9)	0.51
	2b	65.2 (± 20.1)	67.0 (± 16.4)	0.60
P value 2a vs. 2b		P = 0.01	P = 0.009	
AFEQT	2a	70.9 (± 19.8)	67.7 (± 22.2)	0.42
	2b	58.3 (± 17.3)	54.1 (± 20.2)	0.25
P value 2a vs. 2b		P < 0.001	P < 0.0001	

disease-related QoL, as judged by the AFEQT questionnaire and with the patients' own perception of their health state (VAS). There is a step-wise, negative association between the EHRA class and both of these measures. The EHRA score can therefore be considered as a validated tool for symptom classification. In a similar manner to the NYHA functional class for heart failure and the Canadian Cardiovascular Society angina scale, the EHRA score allows clinicians to broadly categorize the severity of patients' symptoms. Where the specifics and the complexities of symptoms can be considerable, this sort of categorization provides a simple means of communicating and quantifying symptom severity. It allows a cross-sectional comparison between patients and a longitudinal comparison for individual patients or groups of patients.

However, using health-related utility as a measure of QoL, the EHRA score only showed significant discriminatory power at the boundary between mild (Class 2) and severe (Class 3) symptoms. To try to improve the discriminative ability for patients with mild-moderate symptoms, we subdivided the EHRA Class 2 patients into Classes 2a and 2b, based on the degree to which the patients were 'troubled by their symptoms', and found that the two subdivisions were significantly different from each other on all three QoL measures. Indeed, the health-related utility showed no significant difference between Class 1 (asymptomatic patients) and Class 2a (patients with symptoms but which are not troublesome and do not affect daily activity). Class 2b patients showed a significant reduction (0.09) in health utility compared with Class 2a patients (Table 4). As such it may be more appropriate to base the treatment decisions not simply on the presence or the absence of symptoms, but also on whether the symptoms cause trouble to the patient or not.

At first sight, the EQ-5D questionnaire does not intuitively represent a sensitive QoL assessment for a patient with AF. The value in its use, however, is the assessment of a health utility score at baseline, which can then be reassessed following treatment. This allows calculation of the cost-efficacy of the intervention, and thereby a comparison of medical interventions between different disciplines. There is a growing demand in all countries to understand the cost-effectiveness of treatments used for common conditions to ensure the efficient and appropriate use of health resources. This is particularly the case when considering potentially costly treatments such as catheter ablation. In addition, there is an increasing focus on the patient reported health status as the most relevant outcome of medical

interventions, and as clinicians we should embrace this method of assessing outcomes.<sup>26</sup>

The EQ-5D QoL questionnaire has previously been studied in AF, although not in association with the EHRA score. Berg *et al.*<sup>27</sup> reported the findings from the EQ-5D in the Euro Heart Survey. The EQ-5D was completed by 5050 patients attending specialist hospital departments in 35 European countries. The population studied was somewhat older than ours with a mean age of 66 years. The mean utility was 0.75, which would suggest that most patients had a symptom level around mEHRA Class 2b. A repeat survey after 1 year completed by 3045 patients showed only a minor (0.013) improvement in the health utility. A lower health utility was associated with AF-specific symptoms, but also other variables including increasing age, history of stroke, and the inability to take regular exercise. Only 2.5% of the patients enrolled into the Euro Heart Survey received treatment with catheter ablation.<sup>28</sup> A sub-analysis of the Birmingham Atrial Fibrillation Treatment of the Aged (BAFTA) study, looking at an older population (mean age 82, range 75–99) showed a comparable baseline utility for males (0.77) but a lower utility for females (0.68).<sup>24</sup>

In assessing interventions, doctors often focus on observable events, such as the frequency and the duration of AF episodes, which may not tally with the patients' perspective of their QoL.<sup>7,29</sup> The motivation of a patient to seek treatment, however, is based on the hope that symptoms will improve. We propose that the reclassification of EHRA Class 2 into Classes 2a and 2b can be of clinical use in selected patients with moderate AF-related symptoms. The reclassification of EHRA Class 2 is focused on the impact that AF has on the patient, by stating either that AF symptoms are 'not troublesome to patient' (Class 2a) or that the 'patient is troubled by symptoms' (Class 2b). This distinction will be intuitive for many clinicians, and appears as easily applicable as the original score. From our results, we can see that this clearly differentiates the two groups with not only a statistically significant difference in the health utility, but with an absolute difference of 0.09, a clinically meaningful difference.

## Limitations

The EHRA score, and by extension the mEHRA score, is intended to be used for patients with either paroxysmal or non-paroxysmal AF and we did not differentiate between the two types in our analyses. It is conceivable that, within an individual mEHRA class, there is a difference in the QoL between the two AF types. The subdivision of

EHRA Class 2 in Phase 2 was performed retrospectively by using clinical letters and could be considered subjective as it was based on a judgement of whether the patient was 'troubled by their symptoms'. This was however performed by two independent physicians, who were both blinded to the QoL scores of the patients in Phase 1 of the study. In a small minority of cases ( $n = 9$ ), subdivision was not possible because of inadequate details in the clinical letter and therefore these patients could not be included in Phase 2. By looking at and comparing a prospective cohort, we have shown that retrospective scoring is unlikely to have a significant effect on the conclusions drawn. In addition, although our assessments were made by a range of healthcare providers, all were experienced in the management of patients with AF and the population studied is from a specialist tertiary centre where a large proportion of patients are managed invasively. The classification may be less appropriately used in the hands of others groups of physicians or nurses or in other patient populations. Within these limitations, the mEHRA appears a practicable and useful addition to the EHRA score.

## Conclusions

The EHRA classification, as originally proposed, is a valid means of quantifying AF symptom severity and correlates well with AFEQT and with the generic EQ-5D. Subdividing EHRA Class 2 into Classes 2a and 2b by a single additional question has the potential to discriminate two clearly separate groups, with Class 2b patients having significantly impaired QoL due to AF, whereas those with Class 2a symptoms do not differ significantly from asymptomatic patients. This simple modification may further improve the clinical usefulness of the EHRA score, particularly when considering interventions such as ablation, where Class 2b symptoms appear to be the appropriate treatment threshold.

## Acknowledgement

The authors wish to acknowledge the invaluable contribution of Damian Cullen in managing the data.

**Conflict of interest:** none declared.

## Funding

Funding to pay the Open Access publication charges for this article was provided by a non-commercial Departmental/Hospital fund.

## References

- Davis RC, Hobbs FD, Kenkre JE, Roalfe AK, Iles R, Lip GY et al. Prevalence of atrial fibrillation in the general population and in high-risk groups: the ECHOES study. *Europace* 2012;**14**:1553–9.
- Bonhorst D, Mendes M, Adragao P, De SJ, Primo J, Leiria E et al. Prevalence of atrial fibrillation in the Portuguese population aged 40 and over: the FAMA study. *Rev Port Cardiol* 2010;**29**:331–50.
- Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA* 2001;**285**:2370–5.
- Fitzmaurice DA, Hobbs FD, Jowett S, Mant J, Murray ET, Holder R et al. Screening versus routine practice in detection of atrial fibrillation in patients aged 65 or over: cluster randomised controlled trial. *BMJ* 2007;**335**:383.
- Singh SN, Tang XC, Singh BN, Dorian P, Reda DJ, Harris CL et al. Quality of life and exercise performance in patients in sinus rhythm versus persistent atrial fibrillation: a Veterans Affairs Cooperative Studies Program Substudy. *J Am Coll Cardiol* 2006;**48**:721–30.
- Dorian P, Guerra PG, Kerr CR, O'Donnell SS, Crystal E, Gillis AM et al. Validation of a new simple scale to measure symptoms in atrial fibrillation: the Canadian Cardiovascular Society Severity in Atrial Fibrillation scale. *Circ Arrhythm Electrophysiol* 2009;**2**:218–24.
- Dorian P, Jung W, Newman D, Paquette M, Wood K, Ayers GM et al. The impairment of health-related quality of life in patients with intermittent atrial fibrillation: implications for the assessment of investigational therapy. *J Am Coll Cardiol* 2000;**36**:1303–9.
- Hawe E, Baillie L, Praet C. Estimating the direct costs of atrial fibrillation to the NHS in the constituent countries of the UK and at SHA level in England, an OHE Consulting Report. London: The Office of Health Economics; 2009.
- Kirchhof P, Auricchio A, Bax J, Crijns H, Camm J, Diener HC et al. Outcome parameters for trials in atrial fibrillation: recommendations from a consensus conference organized by the German Atrial Fibrillation Competence NETwork and the European Heart Rhythm Association. *Europace* 2007;**9**:1006–23.
- Ganiats TG, Browner DK, Dittrich HC. Comparison of quality of well-being scale and NYHA functional status classification in patients with atrial fibrillation. New York Heart Association. *Am Heart J* 1998;**135**(Pt 1):819–24.
- Chiang CE, Naditch-Brule L, Murin J, Goethals M, Inoue H, O'Neill J et al. Distribution and risk profile of paroxysmal, persistent, and permanent atrial fibrillation in routine clinical practice: insight from the real-life global survey evaluating patients with atrial fibrillation international registry. *Circ Arrhythm Electrophysiol* 2012;**5**:632–9.
- Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I et al. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). *Europace* 2010;**12**:1360–420.
- Camm AJ, Lip GY, De CR, Savelieva I, Atar D, Hohnloser SH et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation—developed with the special contribution of the European Heart Rhythm Association. *Europace* 2012;**14**:1385–413.
- Steg PG, Alam S, Chiang CE, Gamra H, Goethals M, Inoue H et al. Symptoms, functional status and quality of life in patients with controlled and uncontrolled atrial fibrillation: data from the RealiseAF cross-sectional international registry. *Heart* 2012;**98**:195–201.
- Kirchhof P, Breithardt G, Camm AJ, Crijns HJ, Kuck KH, Vardas P et al. Improving outcomes in patients with atrial fibrillation: rationale and design of the Early treatment of Atrial fibrillation for Stroke prevention Trial. *Am Heart J* 2013;**166**:442–8.
- Kirchhof P, Ammentorp B, Darius H, de Caterina R, Le Heuzey J, Schilling RJ et al. Management of atrial fibrillation in seven European countries after the publication of the 2010 ESC guidelines on atrial fibrillation: primary results of the prevention of thromboembolic events—European Registry in Atrial Fibrillation Registry. *Europace* 2014;**16**:6–14.
- Spertus J, Dorian P, Bubien R, Lewis S, Godejohn D, Reynolds MR et al. Development and validation of the Atrial Fibrillation Effect on Quality-of-Life (AFEQT) Questionnaire in patients with atrial fibrillation. *Circ Arrhythm Electrophysiol* 2011;**4**:15–25.
- The EuroQol Group: EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208.
- Brooks R. EuroQol: the current state of play. *Health Policy* 1996;**37**:53–72.
- Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005;**14**:1523–32.
- McKenna C, Palmer S, Rodgers M, Chambers D, Hawkins N, Golder S et al. Cost-effectiveness of radiofrequency catheter ablation for the treatment of atrial fibrillation in the United Kingdom. *Heart* 2009;**95**:542–9.
- Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* 1998;**316**:736–41.
- Jansson KA, Granath F. Health-related quality of life (EQ-5D) before and after orthopedic surgery. *Acta Orthop* 2011;**82**:82–9.
- Roalfe AK, Bryant TL, Davies MH, Hackett TG, Saba S, Fletcher K et al. A cross-sectional study of quality of life in an elderly population (75 years and over) with atrial fibrillation: secondary analysis of data from the Birmingham Atrial Fibrillation Treatment of the Aged study. *Europace* 2012;**14**:1420–7.
- Arbelo E, Brugada J, Hindricks G, Maggioni A, Tavazzi L, Vardas P et al. ESC-EURObservational Research Programme: the Atrial Fibrillation Ablation Pilot Study, conducted by the European Heart Rhythm Association. *Europace* 2012;**14**:1094–103.
- Rumsfeld JS, Alexander KP, Goff DC Jr, Graham MM, Ho PM, Masoudi FA et al. Cardiovascular health: the importance of measuring patient-reported health status: a scientific statement from the American Heart Association. *Circulation* 2013;**127**:2233–49.
- Berg J, Lindgren P, Nieuwlaet R, Bouin O, Crijns H. Factors determining utility measured with the EQ-5D in patients with atrial fibrillation. *Qual Life Res* 2010;**19**:381–90.
- Nieuwlaet R, Capucci A, Camm AJ, Olsson SB, Andresen D, Davies DW et al. Atrial fibrillation management: a prospective survey in ESC member countries: the Euro Heart Survey on Atrial Fibrillation. *Eur Heart J* 2005;**26**:2422–34.
- Frost MH, Bonomi AE, Ferrans CE, Wong GY, Hays RD. Patient, clinician, and population perspectives on determining the clinical significance of quality-of-life scores. *Mayo Clin Proc* 2002;**77**:488–94.