openheart

Internationally validated score to predict the outcome of non-paroxysmal atrial fibrillation ablation: the 'FLAME score'

Vennela Boyalla [ID],[1,2] Julian W E Jarman,[1,2] Vias Markides,[1,2] Wajid Hussain,[1,2] Tom Wong,[1,2] R Hardwin Mead,[3] Gregory Engel,[3] Melissa H Kong,[3] Rob A Patrawala,[3] Roger A Winkle[3]

[1]Heart Division, Royal Brompton and Harefield Hospitals, London, UK
[2]National Heart & Lung Institute, Imperial College London, London, London, UK
[3]Silicon Valley Cardiology, East Palo Alto, California, USA

Correspondence to
Dr Vias Markides; v.markides@rbht.nhs.uk

ABSTRACT

Background The clinical effectiveness of ablating non-paroxysmal atrial fibrillation (non-PAF) relies on proper patient selection. We developed and validated a scoring system to predict non-PAF ablation outcomes.

Methods Data on 416 non-PAF ablations were analysed using binary logistic regression at a London centre. Identified preprocedural variables, which independently predicted freedom from atrial tachyarrhythmia. Twenty-one possible predictive variables and a model with c-statistic 0.751—explained outcome variation in London at mean follow-up 12±3 months. An additive point score (range 0–9) was developed—the FLAME score: female=1; long-lasting persistent atrial fibrillation=1; left atrial diameter in mm: 40 to <45 = 1, 45 to <50 = 2, 50 to <55=3, ≥55 =4; mitral regurgitation (MR) mild to moderate=1; extreme comorbidity=2. Extreme comorbidities include severe MR, moderate mitral stenosis, mitral replacement, hypertrophic cardiomyopathy or congenital heart disease.

Results The FLAME score was applied to data (882 non-PAF ablations) at a Californian centre, and predicted the outcome of both single (p<0.0001) and multiple (p<0.0001) procedures. For first ablation (follow-up 2.1 years (median, IQR 1.0–4.1)), FLAME score: 0–1 predicts 62% success, 2–4 44% and ≥5 29% (P_{trend} <0.0001). After the final ablation (mean procedures: 1.4±0.6, follow-up 1.8 years (median, IQR 0.8–3.6)), FLAME score: 0–1 predicts 81% success, 2–4 65% and ≥5 44% (P_{trend} <0.0001).

Conclusions FLAME score is easily calculated, derived in London, and predicted single and multiple procedural outcomes for non-PAF ablations in California. In patients with a high score, even multiple procedures are usually ineffective.

Key questions

What is already known about this subject?
► Numerous predictive scores have been developed over the last decade to determine outcomes for patients undergoing catheter ablation for atrial fibrillation (AF). However, these scores were developed for all types of AF, from paroxysmal AF (PAF) to non-PAF. Success rates for ablating patients with PAF are high, and therefore, prediction scores are of limited use in this subtype. Conversely, the non-PAF cohort is difficult to treat with lower success rates following ablation.

What does this study add?
► The Female, Long-lasting, Atrial diameter, Mitral, Extreme score is the first outcome prediction tool focusing only on the non-PAF population undergoing catheter ablation. It is easy to calculate based on baseline clinical variables (history, sex and echocardiography). The majority of patients with non-PAF require more than one procedure to achieve sinus rhythm. The score can stratify patients both for first and multiprocedural outcomes after catheter ablation.

How might this impact on clinical practice?
► Primarily, the impact of this score will be precision medicine, giving clinicians a tool to analyse individual patient risk for ablation outcomes. Secondary and more importantly is to educate patients by setting expectations for their ablation journey.

INTRODUCTION

Catheter ablation is an effective treatment for most patients with symptomatic paroxysmal atrial fibrillation (AF), establishing its superiority over antiarrhythmic drugs (AADs) in restoring sinus rhythm in multiple randomised controlled trials.[1][2] However, its efficacy at treating non-paroxysmal AF (PAF) is lower,[3] with long-term maintenance of sinus rhythm following a single procedure below 30% in some studies,[4] although improvement beyond 70% with multiple procedures has been reported.[4][5]

Internationally practice has gravitated towards offering catheter ablation as first-line treatment for symptomatic PAF, given a clear trajectory evidenced by high success rates.[2] However, the success rate in ablating non-PAF is highly variable,[4][5] therefore, appropriate patient selection is paramount. A scoring system for non-PAF, predicting successful restoration of sinus rhythm after single or multiple ablations, would be invaluable for guiding patients before

embarking on an invasive treatment journey. Such a score has not been recommended in international guidelines, despite evidence correlating catheter ablation outcomes with several clinical variables,[3 6] such as left atrial (LA) diameter and duration of non-PAF.

We developed an internationally relevant scoring system by first identifying baseline clinical variables independently predictive of a successful outcome of catheter ablation of non-PAF at a hospital in London (UK) and derived a scoring system to predict procedural outcomes. We then externally validated this scoring system among an independent cohort of patients at a hospital in California (USA).

METHODS

Score development: London

The Royal Brompton & Harefield Hospitals (London, UK) are a specialist heart and lung centre in which seven cardiologists were performing catheter ablation (procedural details are in online supplemental material) for non-PAF. As a national referral centre, the casemix is relatively complex, including patients with adult congenital heart disease. All patients provided informed written consent. Patients are reviewed regularly during the year after discharge, with ambulatory 24-hour ECG recordings to document arrhythmias standardly performed at 6 and 12 months postoperatively, and additional 7-day ECG recordings performed as required in response to symptoms. AADs were discontinued at the discretion of the clinician.

We retrospectively reviewed the records of all patients who underwent initial or redo non-PAF catheter ablations during this period. Redo procedures, following the first procedure for non-PAF, were included where the recurrent arrhythmia was PAF, non-PAF or LA tachycardia (AT). The individual procedure outcome (IPO), the outcome following an initial or redo procedure. This was defined as successful if, following the procedure examined, there was an absence of greater than 30s of atrial arrhythmia (fibrillation, flutter or tachycardia) based on symptoms and ambulatory ECG recordings, at all follow-up visits following a 'blanking period' of 3 months.

We examined 17 preoperative clinical variables (table 1) as possible predictors of IPO, as well as whether a redo procedure, the number of previous ablations, presence of AT in redo procedures and follow-up duration. Longstanding persistent AF was defined as continuous AF of greater than 12 months' duration. One clinical variable, 'extreme comorbidity', was a composite of rarely occurring comorbidities thought likely to reduce the chance of success but difficult to examine individually. Extreme comorbidity was defined as the presence of any one of severe mitral regurgitation (MR), moderate or severe mitral stenosis, mitral valve replacement, hypertrophic cardiomyopathy or structural congenital heart disease with significant ongoing haemodynamic impact on the atria. It is also important to note that rate control, usually with beta-blockers, was offered to all patients with

appropriate dose titration. In those with new left ventricular dysfunction, electrical cardioversion, often aided by concomitant administration of AADs such as amiodarone, was also frequently offered and performed.

We performed univariate analysis to identify variables associated with IPO and then entered possible explanatory variables into a multivariate binary logistic regression analysis sequentially to construct a model to explain the variation in IPO as the dependent variable. Using baseline clinical variables independently predictive of IPO, we then constructed an additive point score—the 'Female, Long-lasting, Atrial diameter, Mitral, Extreme (FLAME) score'- to predict non-PAF ablation outcomes.

Score testing: California

Silicon Valley Cardiology (California, USA) is a specialist cardiac centre where four cardiologists were performing catheter ablation for patients with non-PAF. All patients provided informed written consent. The catheter ablation techniques employed were similar to those in London (further details in online supplemental material). There was no significant change in procedural outcomes for non-PAF during the period studied.[7] Follow-up arrangements were similar to those in London, and AADs were discontinued in all patients.

Data on all patients who underwent initial or redo non-PAF catheter ablations during this period were collected prospectively. Inclusion criteria were the same as those employed in London. We examined the outcome following the patient's first procedure and their final procedure. For both procedures, success was defined in the same way as in London. Kaplan-Meier analysis and calculation of the c-statistic were used to determine the predictive accuracy of the FLAME score, and Ptrend (Mantel-Haenszel test) was used as a test for trend in success rate as the FLAME score increases.

Statistical analysis

Population differences were evaluated with Pearson's χ^2 test (categorical data), Kruskal-Wallis test (>2 categories), and Student's t-test (continuous data). Univariate relationships to IPO used Pearson's χ^2 or Fisher's exact test (categorical data), Mantel-Haenszel test of trend (>2 categories) and Student's t-test (continuous data). Tests were performed two tailed, and values of $p<0.05$ were considered statistically significant. Variables were entered into multivariate stepwise binary logistic regression in order of univariate significance, with IPO the dependent variable. Variables with $p<0.05$ in multivariate analysis, or borderline significance but clinically relevant, were retained in the final model and the c-statistic and Hosmer-Lemeshow test calculated. An adjustment was not made for multiple testing as reported results are explorative. Kaplan-Meier log-rank testing and c-statistic calculation quantified the predictive accuracy of the score. Mantel-Haenszel test of trend assessed stepwise change in success rate by score. Time at risk was accounted for in London data by including follow-up duration in multivariate analysis and in Californian data with Kaplan-Meier log-rank testing.

Table 1 Population characteristics in London and California

Variable	London Per patient (n=361)	California Per patient (n=619)	P value
Age (years, mean±SD)	60±11	63±10	0.001*
Female sex	98 (27)	153 (25)	0.401
Body mass index (kg/m^2, mean±SD)	29±6	30±5	0.133
Long-standing persistent AF (>1 year)	105 (29)	157 (25)	0.204
Left atrial diameter (mm, mean±SD)†	47±8	45±7	<0.0001*
Left ventricular function			<0.0001*
Normal	232 (64)	448 (72)	
Mildly impaired	50 (14)	122 (20)	
Moderately impaired	38 (11)	33 (5)	
Severely impaired	41 (11)	16 (3)	
Valvular heart disease‡	38 (11)	49 (8)	0.166
Extreme comorbidity§	39 (11)	27 (4)	0.0001*
Hypertension	105 (29)	318 (51)	<0.0001*
Diabetes mellitus	29 (8)	67 (11)	0.156
Coronary artery disease	53 (15)	90 (15)	1.000
Thyrotoxicosis within the last year¶	11 (3)	13 (2)	0.355
Pulmonary disease	10 (3)	38 (6)	0.002*
Stroke or TIA	21 (6)	48 (6)	0.253
Renal dysfunction	5 (1)	N/A	N/A
Previous antiarrhythmic drugs (median(IQR))**	1 (0–1)	1 (1–2)	<0.0001*
Previous use of amiodarone	115 (32)	209 (34)	0.540
FLAME score (mean±SD)	2.9±1.7	2.5±1.6	<0.0001*
	Per procedure (n=416)	Per patient (n=619)	
Follow-up (months, mean±SD/median (IQR))	12±3	25 (12–50)	<0.0001*

Values shown are number (%) unless otherwise indicated. Per patient values are calculated at the time of a patient's first procedure.
*Statistically significant p<0.05.
†Transthoracic echocardiographic left parasternal long axis anteroposterior diameter at end-systole.
‡Moderate or greater stenosis or regurgitation, replacement or repair of any left-sided valve.
§See figure 1 for definition
¶Thyrotoxic within the last year.
**Class I or III agents.
AF, atrial fibrillation; FLAME, Female, Long-lasting, Atrial diameter, Mitral, Extreme; N/A, not available; TIA, transient ischaemic attack.

Patients were censored at the time of failure or last follow-up. Analyses were performed using IBM SPSS Statistics software (V.20, IBM).

RESULTS

Score development: London

A total of 416 procedures were examined in London among 361 patients. Patients' baseline characteristics are summarised in table 1. The mean duration of follow-up was 12±3 months. AADs were discontinued following 368 (88%) procedures. Univariate and multivariate analyses of variables possibly predictive of the IPO are shown in table 2. By univariate analysis, significantly lower chance of therapy success was predicted by long-standing persistent AF >1 year (p<0.0001), increasing duration of follow-up (p<0.0001), increasing LA diameter (p=0.001), extreme comorbidity (p=0.001), lower age (p=0.002), recent thyrotoxicosis (p=0.006), previous use of amiodarone (p=0.008), increasing number of previous AADs (p=0.013) and female sex (borderline significance, p=0.051).

By multivariate analysis, a model with c-statistic 0.751 (95% CI, 0.701 to 0.801 (Hosmer-Lemeshow test=0.632) was constructed from variables independently predictive of significantly lower chance of therapy success: long-standing persistent AF >1 year (p=0.001), female

Table 2 Relationship between clinical variables and procedural success in London by univariate and multivariate analysis

Variable	Univariate analyses	Multivariate analyses	
	P value	RR (95% CI)*	P value
Long-standing persistent AF (>1 year)	<0.0001	0.40 (0.23 to 0.70)	0.001
Follow-up (months)	<0.0001	0.87 (0.79 to 0.95)†	0.003
Female sex	0.051	0.44 (0.25 to 0.78)	0.005
Age (years)	0.002		0.006
<50		0.26 (0.11 to 0.58)	0.001
50–59		0.65 (0.32 to 1.32)	0.234
60–69		0.81 (0.43 to 1.54)	0.517
≥70		1.000	1.000
Thyrotoxicosis within the last year	0.006	0.21 (0.06 to 0.83)	0.025
AT rather than AF (at redo procedure)	0.122	2.04 (1.10 to 3.82)	0.025
Left atrial diameter (mm)	0.001	0.96 (0.93 to 0.996)‡	0.028
Extreme comorbidity	0.001	0.46 (0.21 to 1.03)	0.059‡
Previous use of amiodarone	0.008		
Previous antiarrhythmic drugs§	0.013		
Valvular heart disease	0.062		
Left ventricular function	0.106		
Diabetes mellitus	0.177		
Renal dysfunction	0.178		
Hypertension	0.191		
Stroke or TIA	0.198		
Any previous AF ablation	0.327		
Coronary artery disease	0.374		
Pulmonary disease	0.429		
Number previous AF ablations	0.783		
Body mass index	0.987		

*Relative Risk (95% CI) of success, derived from binary logistic regression.
†Continuous variable; negative correlation with success.
‡Borderline statistical significance
§Class I or III agents
AF, indicates atrial fibrillation; AT, atrial tachycardia; TIA, transient ischaemic attack.

sex (p=0.005), increasing LA diameter (p=0.028), presence of extreme comorbidity (borderline significance, p=0.059), thyrotoxicosis within the last year (p=0.025), age <50 years (p=0.006), AF rather than AT (relevant to redo procedures) (p=0.025) and increasing duration of follow-up (p=0.003) (table 2).

'FLAME score'

An additive point score (range 0–9) relevant to the selection of patients to enter a programme of catheter ablation for non-PAF was developed. Baseline clinical variables which were not modifiable and were independently predictive of therapy failure were included. As extreme comorbidity (including severe MR) had relatively few procedures to allow detailed statistical assessment of its impact and was associated with poor outcomes (22% success), it was included despite borderline statistical significance. It was given a relatively high weighting, and the presence of mild to moderate MR was included with a lower weighting. Recent thyrotoxicosis was not included being considered modifiable. While there was an increased risk of therapy failure in patients <50 years, there was no significant difference in outcome among other age strata. A very high burden of comorbidity in the younger London patients (including extreme comorbidity in 14/78 (18%)<50 years vs 36/338 (11%)≥50 years; p=0.032), particularly congenital heart disease (8/78 (10%)<50 years vs 15/338 (4%)≥50 years; p=0.031), not fully reflected by the variables present in multivariate analysis, did not appear relevant to most populations. Age was, therefore, not included in the score. The score is shown in figure 1. A model containing the score's variables alone had a c-statistic of 0.675 (95% CI, 0.620 to 0.731, Hosmer-Lemeshow test=0.348).

Risk factor	Score
Female sex	1
Long-lasting persistent AF (>1 year)	1
Atrial diameter (left, mm)	
40 - <45	1
45 - <50	2
50 - <55	3
≥55	4
Mitral regurgitation (mild to moderate)	1
Extreme comorbidity*	2
Maximum score	**9**

*Extreme comorbidity is any of:
-severe mitral regurgitation
-moderate or greater mitral stenosis
-mitral valve replacement
-hypertrophic cardiomyopathy
-structural congenital heart disease

Figure 1 The FLAME score. Risk factors and their weightings comprising the FLAME Score. AF, atrial fibrillation.

Figure 2 Outcome of ablation in California by FLAME Score group. Kaplan-Meier curves describing freedomf rom recurrence of sustained atrial arrhythmia (by ECG or symptoms) following an initial (A) or final (B) ablation. Curves show FLAME score groups. FLAME, Female, Long-lasting, Atrial diameter, Mitral, Extreme.

Score performance: California

A total of 882 procedures were examined in California among 619 patients. Baseline characteristics and differences compared with London patients are summarised in table 1. The Californian patients had smaller LA (p<0.0001), better left ventricular function (p<0.0001) and less extreme comorbidity (p=0.0001), leading to a significantly lower FLAME score (p<0.0001). They were, however significantly older (p=0.001), with significantly more hypertension (p<0.0001) and pulmonary disease (p=0.002), and had tried more AADs (p<0.0001). The duration of follow-up was also longer (p<0.0001).

Median duration of follow-up after an initial procedure was 2.1 years (IQR 1.0–4.1) and did not differ by FLAME score (p=0.696). The score predicted the outcome following a patient's initial procedure in Kaplan-Meier analysis (p<0.0001) (figures 2 and 3A). Score 0–1 predicted 62% success, 2–4 44%, and ≥5 29% (P_{trend} <0.0001) (table 3). The c-statistic was 0.643 (95% CI 0.600 to 0.687, Hosmer-Lemeshow test=0.922).

The score also predicted the outcome following a patient's final procedure in Kaplan-Meier analysis
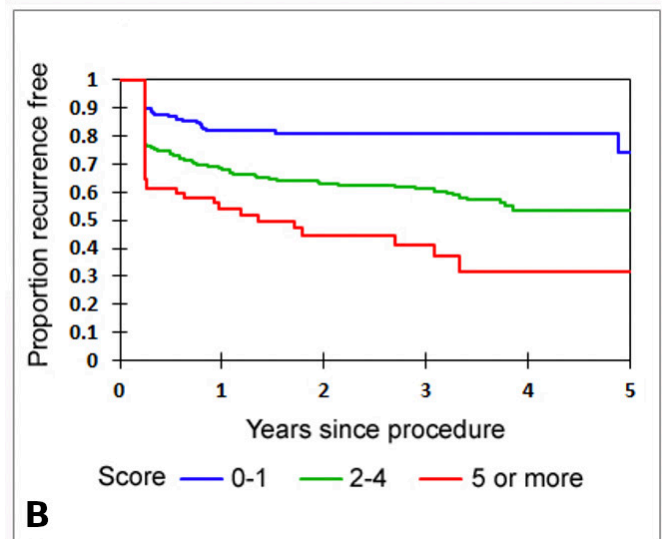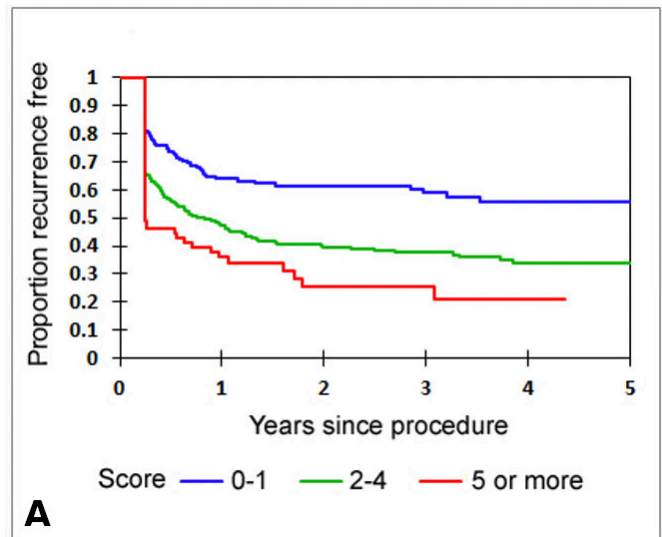
(p<0.0001). Patients underwent mean 1.4±0.6 procedures. With increasing score, the number of repeat procedures increased significantly, being performed in 28% of patients with scores 0–1, 40% with scores 2–4, and 46% with scores ≥5 (p=0.001) (table 3). The median follow-up duration after the final procedure was 1.8 years (IQR 0.8–3.6) and did not differ by *FLAME Score* (p=0.958). Score 0–1 predicted 81% success, 2%–4 65%, and ≥5 44% success (P_{trend} <0.0001) (table 3, figures 2 and 3B). The c-statistic was 0.690 (95% CI, 0.645 to 0.734, Hosmer-Lemeshow test=0.309).

No individual elements of the score dominated the score's predictive power excessively. The highest c-statistic value for any individual element of the score with respect to outcome following a patient's final procedure was 0.598 (95% CI 0.550 to 0.646) for atrial diameter.
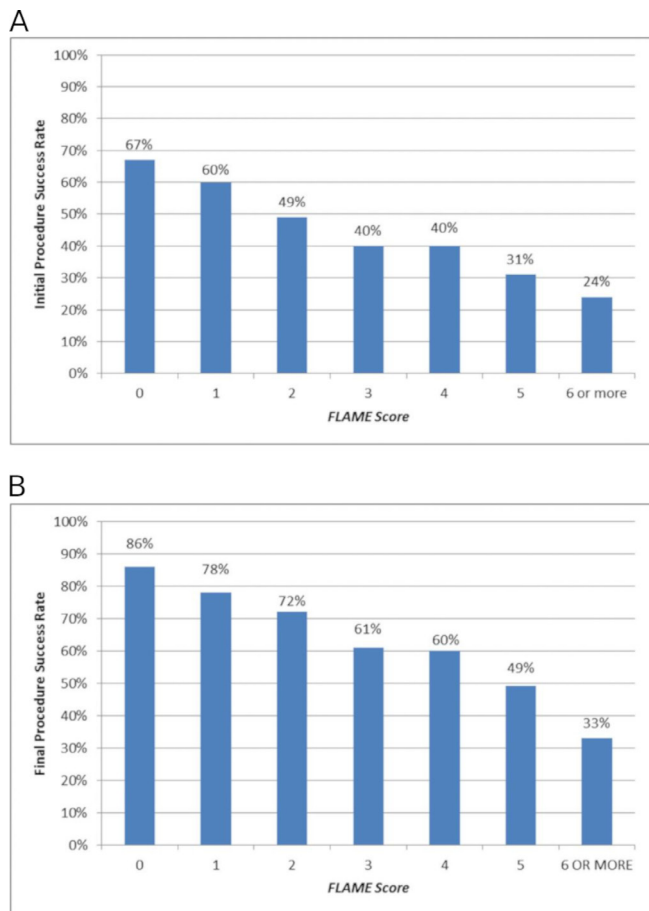
Figure 3 Success rate following the initial (A) and final (B) ablation in California by FLAME score. FLAME, Female, Long-lasting, Atrial diameter, Mitral, Extreme.

DISCUSSION

The clinically challenging and thereby pertinent cohort of non-PAF patients would benefit from a validated outcome prediction score used before embarking on a journey of ablative therapy. Analysis of IPOs of non-PAF catheter ablations at a London hospital identified independent predictors of outcome from which the easily calculated FLAME score was derived, relevant to the

selection of patients for entry into a catheter ablation programme. When validated among patients undergoing non-PAF catheter ablations in California, it effectively stratified the outcomes of both first and multiple procedures. Patients with high scores had reduced success rates following their first procedure, and their final procedure, with fewer maintaining sinus rhythm despite undergoing more repeat procedures than patients with lower scores. Conversely, among patients with a low score (0–1), a success rate of approximately 80% could be achieved over 5 years of follow-up, and with over 70% of patients requiring just a single procedure.

The FLAME score is the first cohort-specific (non-PAF) outcome prediction score for patients undergoing radiofrequency ablation. MB-LATER and ALARMEc scores have a singular aim focus on predicting outcomes for repeat ablations,[8 9] therefore unable to equip physicians to guide a patient at the beginning of their rhythm management journey, which is crucial. The APPLE and CAAP-AF scores have been developed using both derivation and validation cohorts but are not focused on non-PAF, unlike the FLAME score.[10 11] Furthermore, the APPLE score allocates a point for impaired LV function (<50% ejection fraction), although current evidence recommending ablation as a treatment strategy for improving systolic function.[10 12] Despite the proposal of several scores as predictors of AF recurrence following catheter ablation, the FLAME score is unique—the only externally validated score that can predict both initial and multiprocedural success among non-PAF patients.

The variables we found independently associated with the outcome are intuitively understandable and consistent with existing literature.[3–6 13–19] LA diameter is the variable most frequently identified as an independent predictor of non-PAF ablation outcome.[3 5 6 13–17] Additional predictors that have been identified previously in the multivariable analysis include female sex,[3 4 14–16] duration of AF,[3 4 6 13 16 19] and valvular disease.[5] Hypertrophic cardiomyopathy has rarely been examined but was a powerful risk factor for failure when tested.[16] Structural

Table 3 Success rates and other variables by FLAME score in California

FLAME score	Patientsn (%)	Initial procedure				Final procedure			
		Follow-up months mean±SD	Initial procedure: successful	Success rate		Procedures per patient mean±SD	Follow-up months mean±SD	Final procedure: successful	Success rate
0	57 (9)	26±20	38	67%		1.3±0.7	24±19	49	86%
1	123 (20)	32±23	74	60%		1.3±0.6	28±20	96	78%
2	149 (24)	32±23	73	49%		1.4±0.6	28±20	107	72%
3	133 (22)	31±23	53	40%		1.5±0.6	26±20	81	61%
4	85 (14)	32±22	34	40%		1.4±0.6	26±20	51	60%
5	51 (8)	34±20	16	31%		1.5±0.7	28±18	25	49%
6 or more	21 (3)	32±19	5	24%		1.5±0.5	29±16	7	33%
P value		0.696		<0.0001		0.001	0.958		<0.0001

FLAME, Female, Long-lasting, Atrial diameter, Mitral, Extreme.

congenital heart disease is relevant to few, often younger, patients and has not previously been examined, but seems intuitive, and in our experience, an important risk factor for failure. Our finding that left ventricular dysfunction does not predict ablation failure independently is also consistent with most previous studies. With evidence for improvement in functional status with ablation,[20] one could argue that for a given FLAME score, ablation is relatively indicated in heart failure patients. In patients with structural congenital heart disease or mitral valve disease, the choices are harder as, while the benefits of sinus rhythm may be greater, the success rate of catheter ablation is lower.

Despite multivariate analysis, important predictive variables may not be identified where data are few or of poor quality, or if the variable is not examined, and the latter situation may lead to alternative confounded variables appearing predictive. Unfortunately, meta-analyses cannot overcome these problems without access to patient-level data. Another critical factor is the inclusion of intraprocedural or postprocedural variables in previous studies, such as AF cycle length, electro-anatomical mapping (ATLAS score),[21] termination during ablation or early recurrence of AF (MB-LATER).[9] These may impair the ability to detect important preoperative predictive variables due to confounding, for instance, patients with termination during ablation may also have smaller atria. We, therefore avoided this approach.

With the development of consensus around electrically confirmed pulmonary vein isolation with additional substrate modification as the cornerstone of non-PAF ablation, success rates have plateaued in recent years,[7] despite ongoing debate around the correct substrate modification techniques to employ.[22] Preoperative clinical variables, reflecting the severity of the structural and electrophysiological substrate, seem likely to influence success rates irrespective of the exact ablation technique employed,[16] and our findings confirm this concept.

Our results showed that with FLAME score's 0–4, patients could obtain good long-term results with a mean of <1.5 procedures. However, it is important to remember that such a scoring system can only aid in defining one side of a risk–benefit equation, and it does not necessarily follow that high scoring patients should be denied ablation. Weighed against the likelihood of success must be patient' need, which may be particularly pronounced in those who are most highly symptomatic, intolerant of AADs, or haemodynamically compromised. Additionally, the success rates described in this study reflect the 'final procedure' at the time of this study, but not the final clinical outcomes for all patients, many of whom may undergo further procedures in the future. Furthermore, the binary definition of treatment failure used in this analysis may fail to describe the full extent of clinical benefit derived by patients in whom their AF was rendered paroxysmal. However, in the presence of a particularly high FLAME score, the best treatment strategy should be carefully evaluated, and alternative techniques such as surgical or hybrid ablation might be considered.

Limitations

Several questions cannot be answered within the current study design. An exhaustive list of variables was not examined, such as the history and success of cardioversion procedures, and others may exist which reduce success in some patients. Additionally, although previous studies found LA diameter to perform similarly to LA volume,[5 23] and it is more commonly calculated, LA volume may improve prediction at extremes of size.[23] Like others,[5] we found that a binary LA diameter less or greater than 50 mm had predictive power similar to its use as a continuous variable—we chose the latter as, aside from similar statistical power, it seems likely to have a continuous effect. Conversely, duration of non-PAF was tested in a binary manner, in keeping with worldwide definitions and because of the difficulty in determining longer durations accurately; however, we do not know whether it would be more powerful as a continuous variable, nor whether subcategories of persistent AF defined by factors such as a requirement for, or resistance to, electrical cardioversion independently predict outcome. The London population included a minority of patients seen infrequently in other centres, including those with structural congenital heart disease, leading to the unusual finding of the lowest age band having an adverse predictive effect. However, higher age bands had no significant differences in outcome between them, consistent with most other studies,[3–6 13–19] and the score effectively stratified the Californian population where none of the patients had structural congenital heart disease. Intraoperative factors were purposely not tested, which may reduce the specificity of the results, and there was heterogeneity of ablation techniques employed and treatment decisions regarding postoperative AADs; however, our intention was to create a score internationally relevant to current and future practice. Additionally, there was a significant difference in follow-up duration between the populations, although interestingly, the separation of Kaplan-Meier curves observed in the Californian population was almost completed within the duration of the London population follow-up. Finally, in line with international recommendations, while we have used the recurrence of ≥30 s of atrial arrhythmia as our outcome measure, other outcomes such as reduction in the burden of AF or change in the quality of life, could be more clinically meaningful and thus the outcomes described may understate the benefit of this procedure to patients.

CONCLUSION

The FLAME score, easily calculated from baseline clinical variables, was derived from an analysis of individual procedural outcomes of catheter ablation for persistent AF in London. It effectively stratified the outcomes of first or multiple catheter ablations for persistent AF in

California. Such a score may help to better advise individual patients about the effectiveness of catheter ablation for non-PAF.

ORCID iD
Vennela Boyalla http://orcid.org/0000-0002-0285-171X

REFERENCES

1. Terasawa T, Balk EM, Chung M, *et al*. Systematic review: comparative effectiveness of radiofrequency catheter ablation for atrial fibrillation. *Ann Intern Med* 2009;151:191.
2. Hakalahti A, Biancari F, Nielsen JC, *et al*. Radiofrequency ablation vs. antiarrhythmic drug therapy as first line treatment of symptomatic atrial fibrillation: systematic review and meta-analysis. *Europace* 2015;17:370–8.
3. Bhargava M, Di Biase L, Mohanty P, *et al*. Impact of type of atrial fibrillation and repeat catheter ablation on long-term freedom from atrial fibrillation: results from a multicenter study. *Heart Rhythm* 2009;6:1403–12.
4. Rostock T, Salukhe TV, Steven D, *et al*. Long-Term single- and multiple-procedure outcome and predictors of success after catheter ablation for persistent atrial fibrillation. *Heart Rhythm* 2011;8:1391–7.
5. D'Ascenzo F, Corleto A, Biondi-Zoccai G, *et al*. Which are the most reliable predictors of recurrence of atrial fibrillation after transcatheter ablation?: a meta-analysis. *Int J Cardiol* 2013;167:1984–9.
6. Themistoclakis S, Schweikert RA, Saliba WI, *et al*. Clinical predictors and relationship between early and late atrial tachyarrhythmias after pulmonary vein antrum isolation. *Heart Rhythm* 2008;5:679–85.
7. Winkle RA, Mead RH, Engel G, *et al*. Trends in atrial fibrillation ablation: have we maximized the current paradigms? *J Interv Card Electrophysiol* 2012;34:115–23.
8. Mujović N, Marinković M, Marković N, *et al*. Prediction of very late arrhythmia recurrence after radiofrequency catheter ablation of atrial fibrillation: the MB-LATER clinical score. *Sci Rep* 2017;7:40828.
9. Wójcik M, Berkowitsch A, Greiss H, *et al*. Repeated catheter ablation of atrial fibrillation: how to predict outcome? *Circ J* 2013;77:2271–9.
10. Kornej J, Hindricks G, Shoemaker MB, *et al*. The apple score: a novel and simple score for the prediction of rhythm outcomes after catheter ablation of atrial fibrillation. *Clin Res Cardiol* 2015;104:871–6.
11. Winkle RA, Jarman JWE, Mead RH, *et al*. Predicting atrial fibrillation ablation outcome: the CAAP-AF score. *Heart Rhythm* 2016;13:2119–25.
12. Turagam MK, Garg J, Whang W, *et al*. Catheter ablation of atrial fibrillation in patients with heart failure: a meta-analysis of randomized controlled trials. *Ann Intern Med* 2019;170:41.
13. Nademanee K, Schwab MC, Kosar EM, *et al*. Clinical outcomes of catheter substrate ablation for high-risk patients with atrial fibrillation. *J Am Coll Cardiol* 2008;51:843–9.
14. Winkle RA, Mead RH, Engel G, *et al*. Relation of early termination of persistent atrial fibrillation by cardioversion or drugs to ablation outcomes. *Am J Cardiol* 2011;108:374–9.
15. Shim J, Joung B, Park JH, *et al*. Long duration of radiofrequency energy delivery is an independent predictor of clinical recurrence after catheter ablation of atrial fibrillation: over 500 cases experience. *Int J Cardiol* 2013;167:2667–72.
16. McCready JW, Smedley T, Lambiase PD, *et al*. Predictors of recurrence following radiofrequency ablation for persistent atrial fibrillation. *Europace* 2011;13:355–61.
17. Saad EB, d'Avila A, Costa IP, *et al*. Very Low Risk of Thromboembolic Events in Patients Undergoing Successful Catheter Ablation of Atrial Fibrillation With a CHADS$_2$ Score ≤3. *Circulation* 2011;4:615–21.
18. Mohanty S, Mohanty P, Di Biase L, *et al*. Impact of metabolic syndrome on procedural outcomes in patients with atrial fibrillation undergoing catheter ablation. *J Am Coll Cardiol* 2012;59:1295–301.
19. Komatsu Y, Uno K, Otomo K, *et al*. Atrial defibrillation threshold as a novel predictor of clinical outcome of catheter ablation for persistent atrial fibrillation. *Europace* 2011;13:213–20.
20. Jones DG, Haldar SK, Hussain W, *et al*. A randomized trial to assess catheter ablation versus rate control in the management of persistent atrial fibrillation in heart failure. *J Am Coll Cardiol* 2013;61:1894–903.
21. Mesquita J, Ferreira AM, Cavaco D, *et al*. Development and validation of a risk score for predicting atrial fibrillation recurrence after a first catheter ablation procedure – atlas score. *EP Eur* 2018;20:f428–35.
22. Verma A, Jiang C-yang, Betts TR, *et al*. Approaches to catheter ablation for persistent atrial fibrillation. *N Engl J Med Overseas Ed* 2015;372:1812–22.
23. Parikh SS, Jons C, McNitt S, *et al*. Predictive capability of left atrial size measured by CT, TEE, and TTE for recurrence of atrial fibrillation following radiofrequency catheter ablation. *Pacing Clin Electrophysiol* 2010;33:532–40.