Atrial fibrillation risk scores to evaluate left atrial substrate based on voltage analysis in long-standing persistent type of arrhythmia

Radosław Marek Kiedrowicz, Maciej Wielusiński, Andrzej Wojtarowicz, Jarosław Kaźmierczak
Cardiology Department, Pomeranian Medical University, Szczecin, Poland

ABSTRACT

Background: Pre-ablation identification of left atrial (LA) low voltage areas (LVA) among long-standing persistent atrial fibrillation (LSPAF) population remains challenging.

Aims: The aim of the study was to analyze the potential of selected scores originally developed to assess arrhythmia recurrences, thromboembolic complications, or progression from paroxysmal to persistent AF to predict the presence of LA-LVA in LSPAF patients.

Methods: One hundred and fifty-two patients underwent pulmonary vein isolation followed by high-density-high-resolution LA voltage mapping. AF risk scores, such as APPLE, ATLAS, CAAP-AF, DR-FLASH, CHA2DS2-VASc, and HATCH were retrospectively calculated. A receiver operating characteristic curve analysis was performed to evaluate the ability of the scores to predict LVA.

Results: Low voltage areas were detected in 52% of the patients. 28% of the patients with LVA presented severe global LVA burden, whereas 56% of the patients showed a disseminated pattern of remodeling. CAAP-AF ≥7, DR-FLASH ≥4, and CHA2DS2-VASc ≥3 predicted the presence of LVA, whereas ATLAS ≤7 indicated the absence of LVA. ATLAS ≤8, CAAP-AF ≤9, DR-FLASH ≤4, and CHA2DS2-VASc ≤3 predicted the absence of severe LVA. APPLE ≤3 and CHA2DS2-VASc ≤3 predicted the absence of a LVA disseminated pattern. Among predictive scores, ATLAS (AUC, 0.633, 95% CI, 0.543–0.723, P = 0.004), DR-FLASH (AUC, 0.696; 95% CI, 0.594–0.81; P <0.001), and CHA2DS2-VASc (AUC, 0.644; 95% CI, 0.518–0.77; P = 0.025) were the best predictors for the absence of LVA, severe LVA and a disseminated pattern of LVA, respectively.

Conclusions: Atrial fibrillation risk stratification with specific scoring systems can unmask the presence of LA-LVA in the LSPAF population.

Key words: atrial fibrillation, atrial fibrillation risk scores, long-standing persistent atrial fibrillation, low voltage areas, voltage mapping

INTRODUCTION

Important variables associated with atrial fibrillation (AF) pathogenesis and commonly considered for fibrotic remodeling are low voltage areas (LVA) in the left atrium (LA), detected by bipolar voltage mapping [1]. The pre-ablation identification of LVA could essentially contribute to a choice of particular individualized AF therapy, unmasking patients who are unlikely to remain in sinus rhythm or require extensive substrate modification. However, parameters predicting LVA burden are yet to be established, especially among the long-standing persistent AF (LSPAF) population [2]. This study aimed to analyze the potential of selected, easily applicable clinical scores originally developed to assess arrhythmia recurrences after AF ablation, thromboembolic complications, or progression from paroxysmal to persistent AF to predict the presence of LVA in LSPAF patients.

METHODS

Patient selection
One hundred and sixty-six consecutive LSPAF patients who underwent RF point-by-point pulmonary vein isolation (PVI) were enrolled prospectively. Direct current cardioversion was applied to restore sinus rhythm in all patients following PVI. Only individuals able to maintain sinus rhythm (n = 152, 92% of the total study popula-
Detection of LVA

All patients underwent high density-high resolution LA bipolar voltage mapping using the CARTO®3 system fitted with a CONFIDENSE™ module (Biosence-Webster). This mapping protocol is described in detail elsewhere [2]. Briefly, this was performed during coronary sinus (CS) pacing with a Pentaray catheter. The voltage map was created during CS pacing to reduce the occurrence of spontaneous atrial ectopy and to facilitate the identification of incorrectly annotated points. To ensure detailed mapping, the distance thresholding was set at 5 mm, the tissue proximity filter was always enabled and only mapping sites that were within a distance of 5 mm from the acquired shell contributed to the voltage map. Further discrete mapping using a Thermocool SmartTouch catheter (Biosence-Webster), which covered less than 10% of the total LA surface area (TSA) at sites presenting inadequate Pentaray-tissue contact, was performed if necessary. Electrograms (EGM) were only accepted if a contact force was ≥6 g, EGM amplitude ≥0.5 mV was defined as normal and <0.5 mV as diseased tissue. All points presenting low voltage were visually inspected and those incorrectly annotated were deleted from the map. An extension of all areas showing low voltage potentials at least 5 mm away from the ablation lesion set was measured with custom CARTO®3 system software. The global LVA burden was calculated as a sum of LVA and then expressed as a percentage of TSA. The part of the PV inside ablation encirclement, LA appendage and an area adjacent to the fossa ovalis were excluded from TSA calculations. The extent of global LVA burden >20% of TSA was arbitrarily considered as severe based on our observation that all detected LVA can be easily ablated if it occupies less than 20% of TSA [2]. The body of LA was segmented into 5 areas: septum, anterior, posterior, inferior, and lateral wall. If LVA were identified within 3 out of 5 LA segments it was considered a disseminated pattern of voltage-defined remodeling.

Statistical analysis

All continuous variables are expressed as a median and interquartile range [Q1-Q3] as not normally distributed. The categorical variables are presented as values and percentag- es. Comparisons between groups were performed with the Mann–Whitney U-test. A receiver operating characteristic (ROC) curve analysis was performed to evaluate the ability of
AF risk scores to predict LVA. The area under the ROC curve (AUC) was used to evaluate the accuracy of their prognostic values [9]. The sensitivity, specificity, positive predictive value and negative predictive value were calculated with corresponding 95% confidence intervals (CI). Optimal cut-off points were determined by an analysis of sensitivity and specificity values derived from ROC curve data, prioritizing high sensitivity values. Hanley and McNeil's method was used to compare AUC for LVA. Statistical significance was accepted at P value < 0.05. The analysis was performed using Statistica software version 13.3 (StatSoft).

RESULTS

The baseline characteristics of the study population, which includes mean AF risk score values, are presented in Table 1. LVA (15 [7.5–31] cm²) were detected in 52% (79/152) of the patients. Twenty-eight percent of the patients with LVA (22/79) presented a severe global LVA burden >20% of the total LA surface area, whereas 56% of the patients (44/79), a disseminated pattern of remodeling, as 3 out of 5 LA segments were affected. Among the patients with detected LVA, 48% (38/79) had documented LVA on the septum (8 [3–12] cm²), 58% (46/79) on the anterior wall (6.5 [2–10.5] cm²), 77% (61/79) on the posterior wall (7.5 [4–12] cm²), 44% (35/79) on the inferior wall (6 [4–10] cm²), and 25% (20/79) on the lateral wall (3.5 [1.5–6] cm²). Patients with detected LVA and severe LVA pattern had higher values of ATLAS, CAAP-AF, DR-FLASH, and CHA2DS2-VASc scores than those without LVA (Table 2).

| Parameters | Value |
|------------|-------|
| AF duration, months | 24 (12–36) |
| Age, years | 63 (58–67) |
| Females, n (%) | 27 (18) |
| BMI, kg/m² | 30.1 (27.5–32) |
| Hypertension, n (%) | 123 (81) |
| Moderate mitral regurgitation, n (%) | 40 (26) |
| Moderate tricuspid regurgitation, n (%) | 29 (19) |
| Chronic coronary syndrome, n (%) | 34 (22) |
| Prior myocardial infarction, n (%) | 11 (7) |
| Heart failure, n (%) | 38 (25) |
| eGFR <60 ml/min/1.73 m², n (%) | 17 (11) |
| Diabetes, n (%) | 32 (21) |
| Hyperthyroidism, n (%) | 17 (11) |
| LVEF, % | 60 (55–65) |
| LA antero-posterior diameter, mm | 47 (44–50) |
| LAAI, cm²/m² | 13.0 (11.6–15.5) |
| LAVI, ml/m² | 47.0 (37.7–59.4) |
| Enlarged LA (LAVI >34 ml/m²), n (%) | 124 (82) |
| Severely enlarged LA (LAVI >48 ml/m²), n (%) | 65 (43) |

Table 1. Baseline characteristics of the study population (n = 152)

Table 2. Comparison of atrial fibrillation risk scores among different left atrial low voltage areas patterns

Predictive scores for severe LVA

On the ROC curve analyses, only ATLAS, CAAP-AF, DR-FLASH, and CHA2DS2-VASc scores showed significant predictive values for severe LVA (Table 4, Figure 2). However, all scores presented a high negative association. Areas under the ROC-curve, which were ATLAS, CAAP-AF, DR-FLASH and CHA2DS2-VASc score exhibited moderate discriminative power, but in the case of CHA2DS2-VASc the highest values. In the comparison of AUC among the predictive scores no significant difference was observed. It was noted that 7% (10/152) of the patients with ATLAS ≤7 (28/152), 32% (49/152) with CAAP-AF ≥7, 34% (52/152) with DR-FLASH ≥4, and 24% (37/152) with CHA2DS2-VASc ≥3 in the LVA cohort.
Table 3. Receiver operating characteristic curve analysis to compare the ability of atrial fibrillation risk scores to predict the presence of absolute left atrial low voltage areas

| Score       | AUC    | 95% CI       | P  | Best cut-off value | Accuracy | Sensitivity | Specificity | PPV   | NPV   |
|-------------|--------|--------------|----|--------------------|----------|-------------|-------------|-------|-------|
| APPLE       | 0.54   | 0.449–0.632  | 0.39| 2                  | 0.539    | 0.937       | 0.890       | 0.532 | 0.615 |
| ATLAS       | 0.633  | 0.543–0.723  | 0.004| 7                 | 0.632    | 0.844       | 0.612       | 0.604 | 0.707 |
| CAAP-AF     | 0.596  | 0.506–0.686  | 0.04| 7                 | 0.579    | 0.620       | 0.466       | 0.590 | 0.565 |
| DR-FLASH    | 0.61   | 0.521–0.699  | 0.02| 4                 | 0.586    | 0.658       | 0.493       | 0.591 | 0.578 |
| CHA2DS2-VASc| 0.657  | 0.571–0.743  | <0.001| 3                | 0.618    | 0.506       | 0.260       | 0.678 | 0.581 |
| HATCH       | 0.543  | 0.451–0.634  | 0.36| 1                 | 0.539    | 0.899       | 0.849       | 0.534 | 0.579 |

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value

Table 4. Receiver operating characteristic curve analysis to compare the ability of atrial fibrillation risk scores to predict the presence of severe left atrial low voltage areas

| Score       | AUC    | 95% CI       | P  | Best cut-off value | Accuracy | Sensitivity | Specificity | PPV   | NPV   |
|-------------|--------|--------------|----|--------------------|----------|-------------|-------------|-------|-------|
| APPLE       | 0.491  | 0.359–0.623  | 0.89| 5                  | 0.859    | 0.045       | 0           | 1.0   | 0.858 |
| ATLAS       | 0.693  | 0.58–0.8     | <0.001| 8                 | 0.535    | 0.810       | 0.512       | 0.215 | 0.937 |
| CAAP-AF     | 0.672  | 0.542–0.789  | 0.005| 9                 | 0.810    | 0.333       | 0.107       | 0.350 | 0.885 |
| DR-FLASH    | 0.696  | 0.594–0.81   | 0.001| 4                 | 0.528    | 0.875       | 0.529       | 0.20  | 0.950 |
| CHA2DS2-VASc| 0.748  | 0.653–0.878  | <0.001| 3                | 0.683    | 0.714       | 0.322       | 0.278 | 0.932 |
| HATCH       | 0.529  | 0.413–0.644  | 0.62| 1                 | 0.275    | 1.0         | 0.851       | 0.169 | 1.0   |

Abbreviations: see Table 3

Predictive scores for a disseminated pattern of LVA

On the ROC curve analyses, only APPLE and CHA2DS2-VASc scores showed significant predictive values for a disseminated pattern of LVA (Table 5, Figure 3). Each score exhibited a moderately negative predictive ability for LVA. No significant difference was observed between the scores when AUC were compared. It was noted that 7% (10/152) of the patients with APPLE ≤3, and 3% (5/152) with CHA2DS2-VASc ≤2 in this group.

DISCUSSION

Many studies have demonstrated that there is a strong association between LVA and AF duration and adverse ablation outcomes [1]. Therefore, the persistence of AF is a common component of scores to assess arrhythmia recurrences following AF ablation, including APPLE, ATLAS, CAAP-AF, and DR-FLASH scores. However, as noted here, LSPAF diagnosis does not necessarily equate to extensive voltage-derived LA remodeling. The present study focused on the evaluation of multiple AF risk scores, and whether they have any predictive ability to detect voltage-based LA substrate among large, unselected LSPAF cohort. It was found that:

- CAAP-AF ≥7, DR-FLASH ≥4, and CHA2DS2-VASc ≥3 predicted the presence of LVA whereas ATLAS ≤7, the absence of LVA. Among the scores, the ATLAS score showed the highest sensitivity and specificity;
The ATLAS ≤8, CAAP-AF ≤9, DR-FLASH ≤4, and CHA2DS2-VASc ≤3 predicted the absence of severe LVA. Among the scores, the DR-FLASH score showed the highest sensitivity and specificity.

The APPLE ≤3 and CHA2DS2-VASc ≤2 predicted the absence of a disseminated pattern of LVA. Among the scores, the CHA2DS2-VASc score showed the highest sensitivity and specificity.

The HATCH score neither predicted LVA, nor severe or the disseminated pattern of LVA.

To date, only the DR-FLASH score has been developed primarily to directly predict LVA in AF patients [6]. This has been also recently verified [10]. The most notable increase in the dimension of LVA was observed in patients with DR-FLASH scores >3 [6] and a mean DR-FLASH score of 5 was observed among patients with LVA [10]. Moreover, it was recently shown that the APPLE score can be useful to detect LVA [10, 11]. A mean APPLE score of 5 was observed among patients with LVA [10]. Of note, all of the abovementioned research included paroxysmal and persistent, but not long-standing persistent AF patients [6, 10, 11]. This seems to be a main cause for the diverse or discrepant results achieved in our study. The predictive ability of other AF risk scores, such as ATLAS, CAAP-AF, CHA2DS2-VASc, and HATCH for the estimation of the electro-anatomical substrate was yet to be used to date.

All assessed AF risk scores, excluding the HATCH score, were useful to detect voltage-based LA remodeling. However, individual scores applying distinct components were able to detect the various extent of LVA. In addition, many of these presented high negative predictive values, meaning that they perform better as part of a “rule-out” test. It seemed clear that such variations between the scores, in terms of their predictive value, show that the development and progression of LVA are most likely multifactorial with a potential interplay between contributing factors. That is the reason why there is no universal risk score to predict the presence and extent of LVA. The only common parameter included in all 6 scores was patient age (60–75 years). We believe that a major reason that resulted in a HATCH score was the unhelpful fact of unmasking patients with LVA, which did not include female sex and/or LA size, which were integral components of the other scoring system (however, female sex was not incorporated into the APPLE score). Female sex was recently considered a strong risk factor in the development of LA substrate in AF patients. Females might probably present with clinical AF at a later state of fibro-fatty infiltration, which could explain the higher presence of electro-anatomical substrate among them [2].

In the current study, we found that some AF risk scores significantly predicted LA LVA before catheter ablation among LSPAF patients. This could help to identify individuals who require PVI alone, minor substrate modification, or extensive substrate modification in addition to PVI when an ablation strategy is chosen. This finding may essentially contribute to tailored AF therapy when considering a catheter ablation as a potential treatment strategy.

Limitations
The accuracy of LA voltage mapping might have been influenced by several factors, such as mapping during CS pacing, following PVI, using voltage cut-off values <0.5 mV, or due to functional voltage reduction related to the electrical stunning caused by long-lasting AF.

We cannot exclude the possibility that the overall LVA burden might have been altered due to the exclusion

Table 5. Receiver operating characteristic curve analysis to compare the ability of AF risk scores to predict the presence of a disseminated pattern of left atrial low voltage areas

| Score       | AUC       | 95% CI      | P     | Best cut-off value | Accuracy | Sensitivity | Specificity | PPV    | NPV    |
|-------------|-----------|-------------|-------|--------------------|----------|-------------|-------------|--------|--------|
| APPLE       | 0.69      | 0.526–0.864 | 0.02  | 3                  | 0.572    | 0.769       | 0.446       | 0.139  | 0.963  |
| ATLAS       | 0.557     | 0.397–0.717 | 0.48  | 7                  | 0.342    | 0.923       | 0.712       | 0.108  | 0.976  |
| CAAP-AF     | 0.546     | 0.378–0.715 | 0.59  | 8                  | 0.638    | 0.462       | 0.346       | 0.111  | 0.969  |
| DR-FLASH    | 0.539     | 0.367–0.71  | 0.66  | 6                  | 0.868    | 0.154       | 0.065       | 0.182  | 0.922  |
| CHA2DS2-VASc| 0.644     | 0.518–0.77  | 0.03  | 2                  | 0.388    | 0.923       | 0.662       | 0.115  | 0.979  |
| HATCH       | 0.506     | 0.335–0.676 | 0.95  | 2                  | 0.632    | 0.385       | 0.345       | 0.094  | 0.919  |

Abbreviations: see Table 3

Figure 3. A comparison of atrial fibrillation risk scores for the prediction of a disseminated pattern of low voltage areas

- The ATLAS ≤8, CAAP-AF ≤9, DR-FLASH ≤4, and CHA2DS2-VASc ≤3 predicted the absence of severe LVA. Among the scores, the DR-FLASH score showed the highest sensitivity and specificity;
- The APPLE ≤3 and CHA2DS2-VASc ≤2 predicted the absence of a disseminated pattern of LVA. Among the scores, the CHA2DS2-VASc score showed the highest sensitivity and specificity;
- The HATCH score neither predicted LVA, nor severe or the disseminated pattern of LVA.

To date, only the DR-FLASH score has been developed primarily to directly predict LVA in AF patients [6]. This has been also recently verified [10]. The most notable increase in the dimension of LVA was observed in patients with DR-FLASH scores >3 [6] and a mean DR-FLASH score of 5 was observed among patients with LVA [10]. Moreover, it was recently shown that the APPLE score can be useful to detect LVA [10, 11]. A mean APPLE score of 5 was observed among patients with LVA [10]. Of note, all of the abovementioned research included paroxysmal and persistent, but not long-standing persistent AF patients [6, 10, 11]. This seems to be a main cause for the diverse or discrepant results achieved in our study. The predictive ability of other AF risk scores, such as ATLAS, CAAP-AF, CHA2DS2-VASc, and HATCH for the estimation of the electro-anatomical substrate was yet to be used to date.

All assessed AF risk scores, excluding the HATCH score, were useful to detect voltage-based LA remodeling. However, individual scores applying distinct components were able to detect the various extent of LVA. In addition, many of these presented high negative predictive values, meaning that they perform better as part of a “rule-out” test. It seemed clear that such variations between the scores, in terms of their predictive value, show that the development and progression of LVA are most likely multifactorial with a potential interplay between contributing factors. That is the reason why there is no universal risk score to predict the presence and extent of LVA. The only common parameter included in all 6 scores was patient age (60–75 years). We believe that a major reason that resulted in a HATCH score was the unhelpful fact of unmasking patients with LVA, which did not include female sex and/or LA size, which were integral components of the other scoring system (however, female sex was not incorporated into the APPLE score). Female sex was recently considered a strong risk factor in the development of LA substrate in AF patients. Females might probably present with clinical AF at a later state of fibro-fatty infiltration, which could explain the higher presence of electro-anatomical substrate among them [2].

In the current study, we found that some AF risk scores significantly predicted LA LVA before catheter ablation among LSPAF patients. This could help to identify individuals who require PVI alone, minor substrate modification, or extensive substrate modification in addition to PVI when an ablation strategy is chosen. This finding may essentially contribute to tailored AF therapy when considering a catheter ablation as a potential treatment strategy.

Limitations
The accuracy of LA voltage mapping might have been influenced by several factors, such as mapping during CS pacing, following PVI, using voltage cut-off values <0.5 mV, or due to functional voltage reduction related to the electrical stunning caused by long-lasting AF.

We cannot exclude the possibility that the overall LVA burden might have been altered due to the exclusion
of patients unable to maintain sinus rhythm, presenting LAA thrombus, or if another method of LVA detection had been applied.

Women were underrepresented in this study.

**CONCLUSIONS**

Atrial fibrillation risk stratification with specific scoring systems can noninvasively unmask the presence of voltage-derived LA remodeling and lead to more tailored management in the LSPAF population. Among several predictive scores, ATLAS, DR-FLASH, and CHA₂DS₂-VASc scores were the best predictors for the absence of LVA, severe LVA, and the disseminated pattern of LVA, respectively.

**REFERENCES**

1. Blandino A, Bianchi F, Grossi S, et al. Left atrial substrate modification targeting low-voltage areas for catheter ablation of atrial fibrillation: a systematic review and meta-analysis. Pacing Clin Electrophysiol. 2017; 40(2): 199–212, doi: 10.1111/pace.13015, indexed in Pubmed: 28054377.

2. Kiedrowicz RM, Wielusinski M, Wojtarowicz A, et al. Predictors of the voltage derived left atrial fibrosis in patients with long-standing persistent atrial fibrillation. Cardiol J. 2020 [Epub ahead of print], doi: 10.5603/CJ.a2020.0069, indexed in Pubmed: 32419127.

3. Kornej J, Hindricks G, Schoenmaker 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(10): 871–876, doi: 10.1007/s00392-015-0856-x, indexed in Pubmed: 25876528.

4. 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. Europace. 2018; 20(FI_3): f428–f435, doi: 10.1093/eurheartj/eux265, indexed in Pubmed: 29016770.

5. Winkle RA, Jarman JWE, Mead RH, et al. Predicting atrial fibrillation ablation outcome: The CAAP-AF score. Heart Rhythm. 2016; 13(11): 2119–2125, doi: 10.1016/j.hrthm.2016.07.018, indexed in Pubmed: 27435586.

6. Kosiuk J, Dinov B, Kornej J, et al. Prospective, multicenter validation of a clinical risk score for left atrial arrhythmogenic substrate based on voltage analysis: DR-FLASH score. Heart Rhythm. 2015; 12(11): 2207–2212, doi: 10.1016/j.hrthm.2015.07.003, indexed in Pubmed: 26144350.

7. Lip GYH, Nieuwlaat R, Pisters R, et al. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest. 2010; 137(2): 263–272, doi: 10.1378/chest.09-1584, indexed in Pubmed: 19762550.

8. de Vos CB, Pisters R, Nieuwlaat R, et al. Progression from paroxysmal to persistent atrial fibrillation clinical correlates and prognosis. J Am Coll Cardiol. 2010; 55(8): 725–731, doi: 10.1016/j.jacc.2009.11.040, indexed in Pubmed: 20170808.

9. Hanley JA. Receiver operating characteristic (ROC) methodology: the state of the art. Crit Rev Diagn Imaging. 1989; 29(3): 307–335, indexed in Pubmed: 2667567.

10. Kornej J, Schumacher K, Dinov B, et al. Prediction of electro-anatomical substrate and arrhythmia recurrences using APPLE, DR-FLASH and MB-LATER scores in patients with atrial fibrillation undergoing catheter ablation. Sci Rep. 2018; 8(1): 12686, doi: 10.1038/s41598-018-31133-x, indexed in Pubmed: 30139967.

11. Kornej J, Büttnner P, Sommer P, et al. Prediction of electro-anatomical substrate using APPLE score and biomarkers. Europace. 2019; 21(1): 54–59, doi: 10.1093/europace/euy120, indexed in Pubmed: 29893827.