



Prospective multicentric validation of a novel prediction model for paroxysmal atrial fibrillation

Constanze Schmidt^{1,2} · Sebastian Benda^{1,2} · Patricia Kraft³ · Felix Wiedmann^{1,2} · Sven Pleger^{1,3} · Antonius Büscher^{1,2} · Dierk Thomas^{1,2} · Rolf Wachter^{4,5} · Christian Schmid⁶ · Roland Eils^{7,8} · Hugo A. Katus^{1,2} · Stefan M. Kallenberger^{7,8}

Received: 15 August 2020 / Accepted: 27 October 2020 / Published online: 19 November 2020
© The Author(s) 2020

Abstract

Background The early recognition of paroxysmal atrial fibrillation (pAF) is a major clinical challenge for preventing thromboembolic events. In this prospective and multicentric study we evaluated prediction scores for the presence of pAF, calculated from non-invasive medical history and echocardiographic parameters, in patients with unknown AF status.

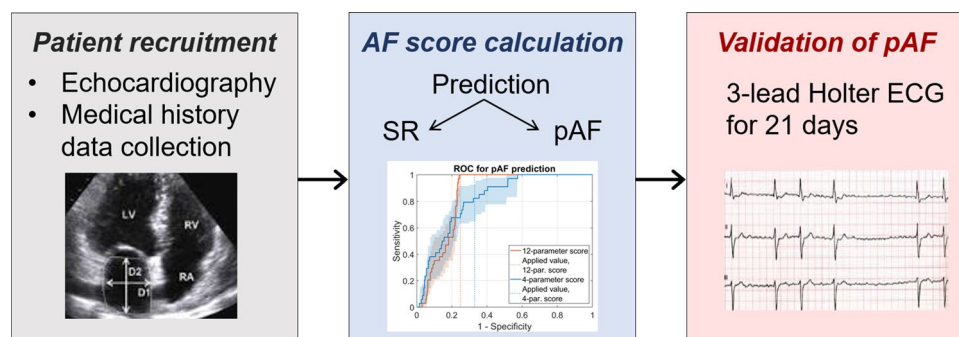
Methods The 12-parameter score with parameters age, LA diameter, aortic root diameter, LV,ESD, TDI A', heart frequency, sleep apnea, hyperlipidemia, type II diabetes, smoker, β -blocker, catheter ablation, and the 4-parameter score with parameters age, LA diameter, aortic root diameter and TDI A' were tested. Presence of pAF was verified by continuous electrocardiogram (ECG) monitoring for up to 21 days in 305 patients.

Results The 12-parameter score correctly predicted pAF in all 34 patients, in which pAF was newly detected by ECG monitoring. The 12- and 4-parameter scores showed sensitivities of 100% and 82% (95%-CI 65%, 93%), specificities of 75% (95%-CI 70%, 80%) and 67% (95%-CI 61%, 73%), and areas under the receiver operating characteristic (ROC) curves of 0.84 (95%-CI 0.80, 0.88) and 0.81 (95%-CI 0.74, 0.87). Furthermore, properties of AF episodes and durations of ECG monitoring necessary to detect pAF were analysed.

Conclusions The prediction scores adequately detected pAF using variables readily available during routine cardiac assessment and echocardiography. The model scores, denoted as ECHO-AF scores, represent simple, highly sensitive and non-invasive tools for detecting pAF that can be easily implemented in the clinical practice and might serve as screening test to initiate further diagnostic investigations for validating the presence of pAF.

Graphic abstract

Prospective validation of a novel prediction model for paroxysmal atrial fibrillation based on echocardiography and medical history parameters by long-term Holter ECG



Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00392-020-01773-z>) contains supplementary material, which is available to authorized users.

Extended author information available on the last page of the article

Keywords Paroxysmal atrial fibrillation · Atrial fibrillation detection · AF prediction model · ECHO-AF scores · Stroke prevention · Systems medicine

Introduction

Atrial fibrillation (AF) is the most frequent arrhythmia but often remains unrecognized. Early detection is highly relevant to reducing stroke, heart failure and mortality but represents a major clinical challenge [1]. Developing new non-invasive methods for early diagnosis of paroxysmal AF (pAF) represents an important task to improve prevention of adverse effects of AF such as stroke, heart failure or cardiomyopathy [2, 3].

As part of the Apple Heart Study and the Huawei Heart Study, the use of smart watches for AF detection was evaluated in several hundred thousand subjects [4, 5]. The studies, however, were conducted in unselected populations. Therefore, AF was only detected in 0.5% or 0.2% of the study population. Population-level screening for AF bears the risk of large numbers of false-positive results and might cause unnecessary additional investigations or treatments [6]. Therefore, screening investigations are particularly valuable in patient groups with increased risk for AF [6]. Several models were established to predict the risk for developing AF in the next 5 or 10 years based on clinical data and blood biochemistry measures using Cox regression and procedures for selecting predictive parameters [7–12]. Scores developed within the CHARGE-AF consortium were challenged in independent studies based on electronic medical record databases and further developed using machine learning methods [12, 13].

It is well established that echocardiography is informative about the hemodynamic and mechanical cardiac functions. Several echocardiographic measures are associated with an increased risk for developing AF or AF-dependent complications such as left atrial (LA) enlargement, an increase of the left ventricular (LV) wall thickness or reduced end-diastolic to end-systolic fractional shortening of the LV [14–16]. Screening for pAF in patients undergoing echocardiography is reasonable because the prevalence of pAF in this patient collective is high. In the echocardiographic laboratory of the Department of Cardiology at Heidelberg University (Heidelberg, Germany), a prevalence of about 16% was observed [17]. Therefore, AF screening in this patient population is appropriate.

Recently, we systematically evaluated the predictive value of various echocardiographic parameters for AF and developed mathematical models and scores for predicting the presence of paroxysmal AF (pAF) using medical history and echocardiographic parameters that can be easily implemented in routine clinical practice [17]. The derivation cohort contained 47 clinical and echocardiographic

parameters from 1000 patients. The optimal score variant includes 12 echocardiographic and medical history parameters that are most predictive for classifying between pAF and sinus rhythm (SR). To further simplify the clinical application, a reduced score with four parameters, age, LA diameter, tissue Doppler imaging velocity during atrial contraction (TDI, A'), and aortic root diameter, was developed.

In this prospective and multicentric study, we tested the pAF prediction scores in 305 patients with unknown AF status by continuous ECG monitoring over a period of up to 21 days. Thereby, we could validate the developed prediction scores, subsequently termed 'ECHO-AF' scores, as non-invasive tool for detecting pAF that can be easily implemented in clinical practice and might serve as a screening test to initiate further diagnostic investigations for validating the presence of pAF.

Methods

Study population

Echocardiographic and additional clinical data of 305 patients without diagnosis of AF or atrial flutter (50.2% males, 49.8% females) who were interested in undergoing a screening investigation for AF were included in this study between May 2016 and November 2019 at the Department of Cardiology and the Department of Neurology of the University Hospital of Heidelberg (Germany), and four cardiology practices in Germany (Heidelberg, Lüneburg, Essen). Patient data were collected in a de-identified manner. Based on the observed prevalence of pAF in our echocardiographic laboratory, we estimated that about 300 subjects were necessary to estimate sensitivity values with confidence intervals of $\pm 10\%$ [18, 19]. The study protocol was approved by the ethics committee of the University of Heidelberg (Germany, Medical Faculty Heidelberg, S-491/2015). Written informed consent was obtained from all patients, and the study was conducted in accordance with the Declaration of Helsinki.

Collected score parameters consisted of medical history parameters (age, smoker, heart frequency, sleep apnea, hyperlipidemia, type 2 diabetes mellitus, catheter ablation), medication (beta-blocker), and echocardiographic parameters (aortic root diameter, left atrial diameter, left ventricular end-systolic diameter, TDI A' velocity). Catheter interventions were ablations of accessory pathways or atrioventricular nodal reentry tachycardia. Patients with a history of atrial fibrillation or atrial flutter ablation were excluded.

Echocardiography and long-term Holter ECG

Transthoracic echocardiography examinations were performed on commercially available ultrasound systems (GE Healthcare, Philips, Sony). Images included parasternal, apical and subxiphoidal views using 1.5–4.0 MHz phase-array transducers. All examinations were performed with 2D echocardiography for anatomic imaging and Doppler echocardiography for assessment of velocities. Left atrial size was determined as the maximal distance between the posterior aortic root wall and the posterior left atrial wall at the end of the systole. Aortic root and left ventricular end-systolic diameters (LV, ESD) were obtained in the parasternal long axis view. The TDI A' velocity was measured in the apical four-chamber view. Patients carried 3-lead 7-day Holter ECG devices (Mortara Instruments) for up to 3 weeks. A few patients carried the device for 4 weeks. ECG recordings were evaluated by a cardiologist blinded to the score parameters and the calculated score values. AF was diagnosed in case AF episodes longer than 20 s were documented.

Statistical methods

Continuous variables between SR and pAF groups were compared using one-way analysis of variance (ANOVA). Standard deviations are indicated by plus–minus signs (Table 1). Categorical variables between groups were compared with two-tailed Fisher exact test. Predictive scores were previously derived from logistic regression models calibrated with 47 parameters recorded in 1000 patients of the Department of Cardiology at Heidelberg University (Heidelberg, Germany) [17]. Of these 47 parameters, the most predictive 12 parameters were identified by sequential feature selection and likelihood-ratio testing. Scores were scaled between 0 and 100. The 12-parameter score is given by Eq. (1), and the 4-parameter score by Eq. (2):

$$L_{12} = -17.07 + 0.3359 \times \frac{\text{age}}{y} + 0.8700 \times \frac{\text{Ao,root}}{\text{mm}} + 0.7512 \times \frac{\text{LA}}{\text{mm}} - 0.3331 \times \frac{\text{LV,ESD}}{\text{mm}} - 1.570 \times \frac{\text{TDI,A'}}{\text{cm/s}} + 0.1527 \times \frac{\text{HF}}{1/\text{min}} + 10.98 \times \text{sleep apnea} + 4.172 \times \text{hyperlipidemia} - 0.1995 \times \text{type II diabetes} + 0.7565 \times \text{smoker} + 5.307 \times \beta\text{-blocker} + 21.39 \times \text{catheter ablation.} \quad (1)$$

$$L_4 = -22.96 + 0.4997 \times \frac{\text{age}}{y} + 0.9188 \times \frac{\text{Ao,root}}{\text{mm}} + 0.9459 \times \frac{\text{LA}}{\text{mm}} - 1.583 \times \frac{\text{TDI,A'}}{\text{cm/s}}. \quad (2)$$

In Eqs. (1) and (2), variables (age; Ao,root, aortic root diameter; LA, LA diameter; LV,ESD, LV end-systolic diameter; TDI,A', tissue Doppler imaging, late diastolic velocity of mitral annulus; HF, heart frequency) are divided by

Table 1 Comparison of score parameters between patient groups

	SR (<i>n</i> = 271)	pAF (<i>n</i> = 34)
Parameters		
Age (years)	58.7 ± 18.6	69.3 ± 12.5**
Aortic root (mm)	28.8 ± 6.3	34.3 ± 4.5***
LA, (mm)	34.4 ± 7.4	40.4 ± 5.5***
LV ESD (mm)	44.3 ± 12.6	34.2 ± 10.1***
Sleep apnea [<i>n</i> (%)]	28 (10)	1(3)
Hyperlipidemia [<i>n</i> (%)]	89 (33)	14(41)
Diabetes mellitus [<i>n</i> (%)]	38 (14)	7 (21)
Smoker [<i>n</i> (%)]	62 (23)	2 (6)*
Beta-blocker [<i>n</i> (%)]	110 (41)	17 (50)
Catheter ablation [<i>n</i> (%)]	18 (7)	6 (18)*
TDI A' (cm/s)	8.2 ± 3.5	7.0 ± 4.2
Heart rate (1/min)	73.8 ± 12.6	76.8 ± 12.0

For continuous parameters, means and standard deviations are given, for categorical parameters with two levels, total counts and percentages are indicated

LA left atrium, LV ESD, left ventricular end-systolic diameter, TDI A', tissue Doppler imaging, late diastolic velocity of mitral annulus

p* < 0.05, *p* < 0.01, ****p* < 0.001 versus SR from ANOVA for continuous variables and from Fisher exact test for categorical variables

their units (y, years; mm, millimetres; cm/s, centimetres per second; 1/min, per minute). Categorical variables (sleep apnea; hyperlipidemia; type II diabetes; smoker; β-blocker, β-blocker intake; catheter ablation, status after catheter ablation) are set to 1 or 0 in case of their presence or absence. Using the 12-parameter score, presence of pAF was predicted in case of $L_{12} \geq 58.35$, and using the 4-parameter score, pAF was predicted in case of $L_4 \geq 63.32$. An online calculator was created to simplify application of the scores [20].

For comparison, we assessed the predictive performance of logistic regression models containing only subsets of variables that are part of the 12-parameter score. We tested the following

reduced models: (1) a model, reduced by all echocardiographic parameters, with parameters age, heart frequency, sleep apnea, hyperlipidemia, type II diabetes, smoker, beta-blocker, catheter ablation, (2) a model containing the parameters age,

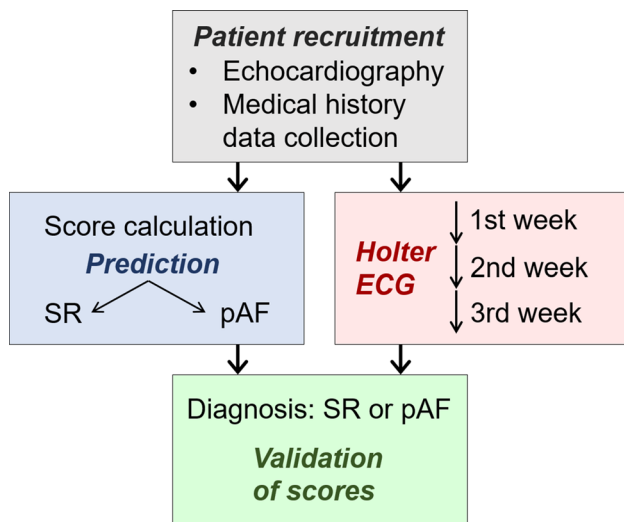


Fig. 1 Design of the multicentric prospective study to evaluate the developed pAF prediction scores. In $n=305$ patients, an echocardiographic investigation was conducted and medical history parameters were collected. Scores were calculated to predict the presence of SR or pAF. Holter ECG monitors were carried for up to 3 weeks and evaluated by medical personnel blinded to predictions based on scores. Diagnoses of pAF based on Holter ECG recordings were used to validate predictions from scores

gender, BMI, and (3) a model containing only the parameter age. Parameters of these reduced models were obtained from calibration of logistic regression models based on the dataset that was previously used to establish the 4-parameter and 12-parameter scores [17]. ROC curves of these models were obtained as previously described using 100-fold cross-validation [17]. We further compared the predictive performance of our scores with two previous prediction scores, the HAVOC and the ACTEL scores [21, 22]. These two scores were developed to predict AF in patients with cryptogenic stroke or transient ischemic attack (TIA) and use non-invasive clinical parameters as well. In these scores, 95% confidence intervals were estimated by bootstrapping with $n=1000$ samples.

To assess classification performance, area under the curve (AUC) values, sensitivities, specificities and precisions were analysed. We estimated 95% confidence intervals for sensitivity, specificity, precision and AUC values by bootstrapping with $n=1000$ samples. All analyses were performed based on pre-implemented functions and custom scripts in MATLAB (MathWorks).

Results

Cohort characteristics and study outcomes

Patients were included in the study at the Department of Cardiology and the Department of Neurology at

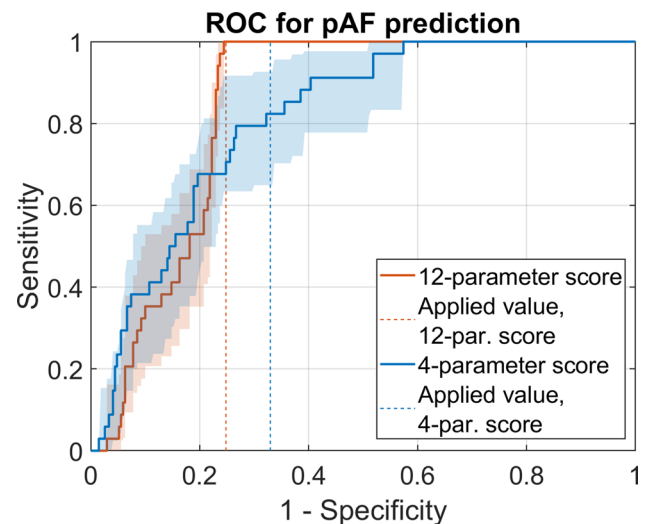


Fig. 2 Performance of predictive scores for classification between pAF and SR. ROC curves were calculated for pAF prediction scores containing 12 parameters or 4 parameters (areas: 95% confidence intervals estimated by bootstrapping). Dashed lines indicate specificity values (75% for 12-parameter score, 67% for 4-parameter score) and sensitivity values (100% for 12-parameter score, and 82% for 4-parameter score) obtained by the pre-defined score thresholds

Heidelberg University as well as four cardiology practices (age between 18 and 93 years, median age: 63, 50.2% males, 49.8% females; Supplementary Fig. S1). Patients undergoing an ambulatory echocardiographic investigation as well as healthy volunteers were offered participation in the study (Fig. 1). Patients were asked to wear Holter ECG devices for up to 3 weeks. Patients carried Holter ECG devices on average for 12.4 days (newly diagnosed pAF patients: 12.1 ± 7.8 days; SR patients: 12.4 ± 7.0 days; Supplementary Fig. S2). Table 1 gives an overview of the score parameters in SR and pAF patients. In addition, we compared measurements of a selection of biomarkers that were available for our patient collective (GFR, creatinine, Troponin-T hs, NT-proBNP). These parameters were previously associated with an increased risk for AF [11, 23, 24]. In our study sample, values of these biomarkers did, however, not significantly differ between SR and AF patients (Supplementary Table S1).

Patients with newly diagnosed pAF had significantly higher age, larger aortic root and left atrial diameters, smaller end-systolic diameters, were less frequently smokers and had more frequently undergone a catheter ablation (accessory pathway or atrioventricular nodal reentry tachycardia ablation). Based on our previous study, threshold values of the developed 12-parameter ($L_{12} \geq 58.35$) and 4-parameter ($L_4 \geq 63.32$) ECHO-AF scores that allowed prediction of pAF with 80% sensitivity were selected [17]. These score threshold values were applied in this study to

prospectively evaluate the predictive performance of the developed predictive scores.

Prospective validation of the ECHO-AF scores

In this study, the 12-parameter and the 4-parameter scores showed sensitivities of 100% and 82% (95%-CI 65%, 93%), specificities of 75% (95%-CI 70%, 80%) and 67% (95%-CI: 61%, 73%), and areas under the receiver operating characteristic (ROC) curves of 0.84 (95%-CI: 0.80, 0.88) and 0.81 (95%-CI 0.74, 0.87) (Fig. 2). No confidence interval can be reported for the sensitivity of the 12-parameter score because all 34 pAF patients of a total of 305 patients were detected by its application, resulting a sensitivity of 100%. Values for sensitivity, specificity, and area under the ROC curve of 12-parameter and 4-parameter scores were higher than expected from the results of our previous retrospective study (Table 2) [17].

The precision of the 12-parameter score, defined as the number of true positives divided by the sum of true and false positives, was 34% (95%-CI 25%, 44%). This value is comparable to the precision value of 36% that was previously estimated based on a retrospective dataset [17]. The 4-parameter score showed a precision of 24% (95%-CI 17%, 32%).

Taken together, the capabilities of the developed 12-parameter and 4-parameter ECHO-AF scores for pAF prediction could be validated indicating that the developed model scores represent a simple, highly sensitive and non-invasive tool for detecting pAF.

Predictive performance of the ECHO-AF scores

To put these results into perspective and assess the predictive value of echocardiographic imaging parameters, we retrospectively tested reduced and modified model variants based on the dataset that was originally used to derive the 12- and 4-parameter scores [17]. Excluding echocardiographic imaging parameters clearly reduced the predictive performance in the original dataset, comprising data from 1000 patients, indicated by an AUC value for the ROC curve

of 0.70 (95%-CI 0.65, 0.75) compared to an AUC value of 0.80 (95%-CI 0.76, 0.84) for the full 12-parameter score (Fig. 3a). Other previous screening studies pointed out the predictive value of the parameters *age*, *gender* and *BMI* or were only based on the subject age [23, 25]. Accordingly, we tested logistic regression models reduced to these parameters. Whereas the model based on *age*, *gender* and *BMI* showed an AUC value under the ROC curve of 0.65 (95%-CI 0.60, 0.69; Fig. 3b), classification only based on age resulted in an AUC value of 0.64 (95%-CI 0.59, 0.69; Fig. 3c). For comparison, we further applied two previous scores (HAVOC and ACTEL) originally developed for detecting AF in patients after stroke or TIA that also resulted in ROC curves with comparably small AUC values (0.66 and 0.64; Supplementary Fig. S3A and B) [21, 22]. It was concluded that the predictive performance of the 12- and 4-parameter scores relies on echocardiographic parameters and that they should not be replaced by reduced variants.

AF detection and characteristics of AF episodes

To evaluate pAF score models, we collected extensive 3-lead Holter ECG recordings that are informative about characteristics of AF episodes in pAF patients. We characterized properties of AF episodes and analysed Holter ECG carrying durations that were necessary to detect pAF. In 25 of a total of 34 patients with AF, up to 3 weeks of time-resolved Holter ECG data were available (Fig. 4). In other pAF patients, pAF was detected by pacemaker devices or 12-lead ECG recorded after including a patient in the study. In several patients, AF was detected in the second or third week. Profiles of AF episodes strongly differed between AF patients. In 9 patients, only episodes of less than one hour were observed. In three patients, AF was persistent. We did not observe circadian rhythmicity of AF episode frequency (Supplementary Fig. S4). Score values were not significantly correlated with AF burden (Supplementary Fig. S5).

We evaluated how long ECG monitoring was necessary to detect AF in patients in case of our study. To this end, we extracted ECG monitoring times before the first AF episode and calculated cumulative fractions of AF diagnoses

Table 2 Comparison between expected and observed performance of pAF prediction scores

	Results from this study		Values from retrospective study	
	12-Parameter score	4-Parameter score	12-Parameter score	4-Parameter score
Sensitivity	100%	82% (65%, 93%)	80% (73%, 87%)	80% (73%, 87%)
Specificity	75% (70%, 80%)	67% (61%, 73%)	68% (73%, 63%)	59% (65%, 52%)
AUC	0.84 (0.80, 0.88)	0.81 (0.74, 0.87)	0.80 (0.76, 0.84)	0.77 (0.72, 0.81)

In brackets, 95% confidence intervals are indicated (AUC, area under the ROC curve). For score development and selection of score thresholds for pAF prediction, confidence intervals were estimated by 100-fold cross-validation. In this study, these score thresholds were prospectively applied. Confidence intervals were estimated from bootstrapping with $n = 1000$ samples

Fig. 3 Comparison with ROC curves for reduced score variants: **a** the 12-parameter score without imaging parameters [AUC 0.70 (95%-C.I.: 0.65, 0.75)]; the original 12-parameter score with AUC 0.80 (95%-CI 0.76, 0.84) is indicated in grey; **b** a score for pAF prediction comprising the variables *age*, *gender* and *BMI* [AUC 0.65 (95%-CI 0.60, 0.69)], (C) a score for pAF prediction comprising only *age* as variable [AUC 0.64 (95%-CI 0.59, 0.69)]; the ‘x’ symbol indicates values of sensitivity and 1-specificity for a cut-off value of 75 years

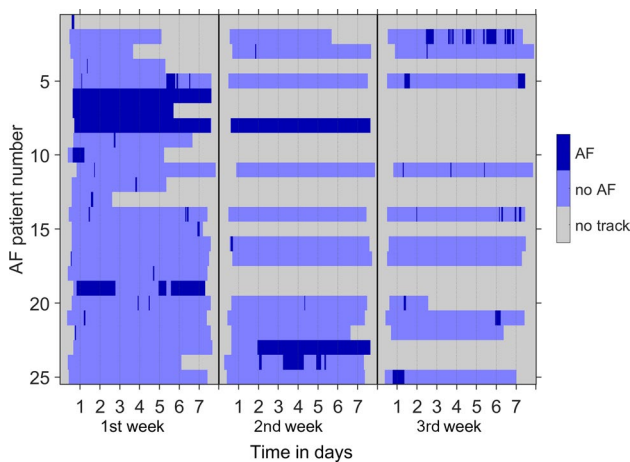
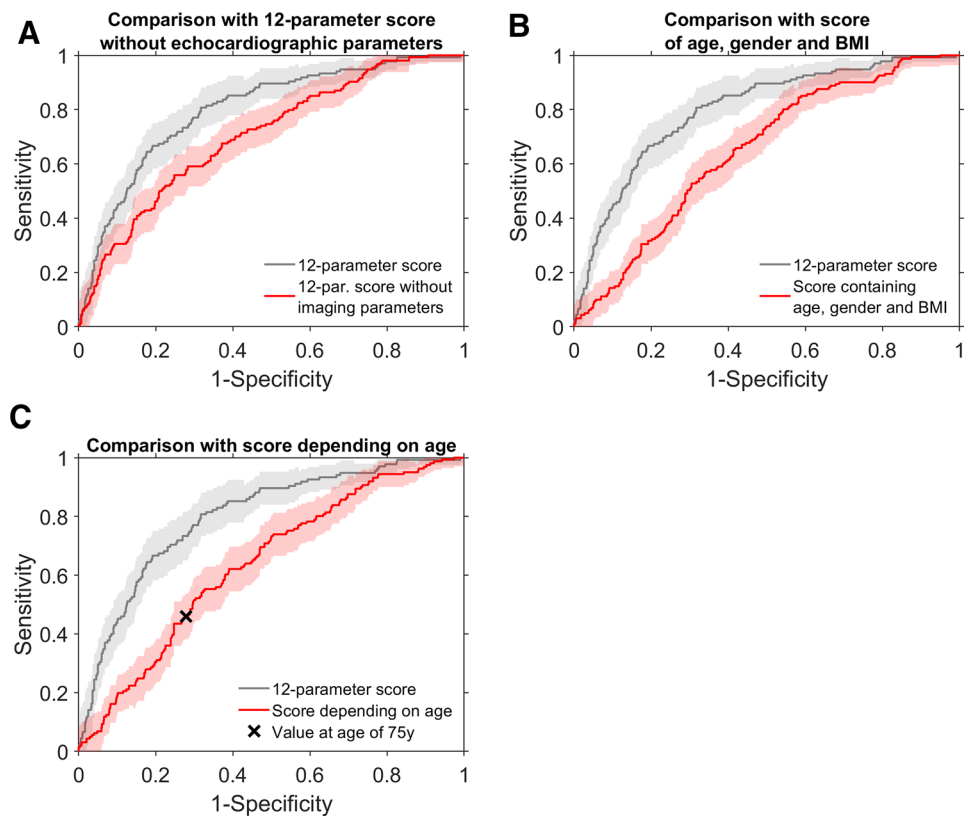


Fig. 4 Temporal profiles of AF episodes. In 25 of a total of 34 patients with AF, up to 3 weeks of time-resolved ECG data were available (y-axis). Episodes are indicated by dark blue colour in ECG traces displayed parallel to the x-axis

dependent on ECG monitoring times (Fig. 5a). Half of the AF patients were diagnosed within 1 day of ECG recording, whereas, 75% were diagnosed within 4.5 days ECG recording (dashed red lines in Fig. 5a). All pAF patients were diagnosed within 14.4 days of ECG monitoring. Figure 5b visualizes AF episode durations and heart rates during AF episodes for different patients. Interestingly, for patients with

non-permanent AF (n = 22), heart rate in the presence of AF was weakly positively correlated with the AF episode duration (Spearman rank-order correlation $\rho = 0.24$, $p = 0.014$).

Discussion

In this study, the capabilities of the developed 12-parameter and 4-parameter ECHO-AF scores for pAF prediction could be prospectively validated. In 34 study patients, pAF could be newly diagnosed. Further diagnostic validation and a positive CHA₂DS₂-VASc score confirmed that oral anticoagulation was necessary in these patients.

The moderate precision values (12-parameter score: 34%, 4-parameter score: 24%) result from the relatively small fraction of pAF patients (11%) relative to the fraction of SR patients in the study sample. Therefore, the scores cannot replace documentation of AF by ECG measurements but represent highly sensitive screening tests to select patients for further diagnostic validation of the presence of pAF by long-term Holter ECG measurements. Our investigation showed that the scores can be used to select patients, in which a long-term ECG monitoring should be conducted. For example, a positive score result could be coupled to further screening tools such as carrying a smart watch that can measure photoplethysmography signals or long-term Holter ECG monitoring. Using the scores to narrow the

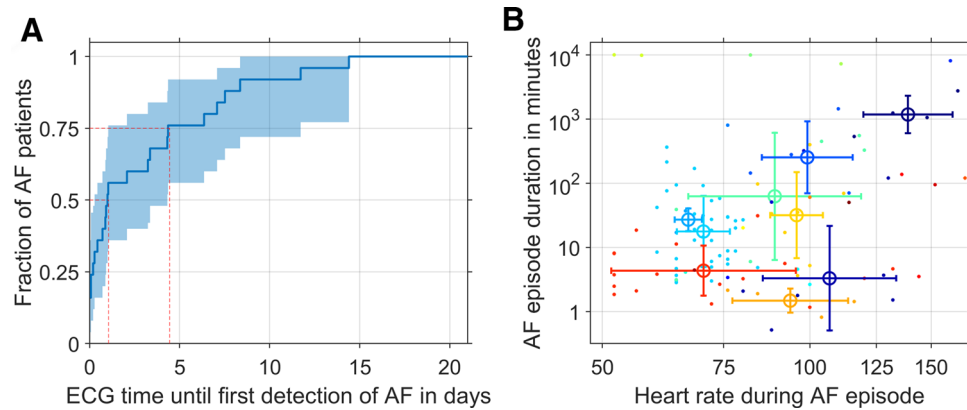


Fig. 5 Characteristics of AF episodes. **a** Cumulative fractions of AF diagnoses for ECG monitoring durations (dashed red lines: ECG monitoring times until 50% or 75% AF diagnoses; shaded areas: 95% confidence intervals estimated by bootstrapping with $n=1000$ sam-

ples). **b** AF episode durations and heart rates during AF episodes. Values of different patients are indicated by colours. Averages and standard deviation of log-scaled values are shown with error bars for all patients with more than three episodes ($n=9$)

patient group, in which further AF screening is performed, can lead to a cost optimization [6].

Previously, several clinical studies defined AF prediction scores based on blood serum parameters [8–12, 23]. In contrast, our study focused on non-invasive clinical parameters, particularly on echocardiographic parameters. These are often available in cardiology practices, whereas, several serum biomarkers predictive for AF are frequently unavailable in patients not treated in a hospital setting. Furthermore, echocardiographic parameters reflecting measures of the cardiac anatomy, blood flow and tissue velocities are immediate indicators of the cardiac function. Therefore, it is likely that these are predictive for arrhythmias with pathophysiological consequences. Comparing the developed scores with variants without imaging parameters underlined that echocardiographic parameters are important for predicting the presence of AF.

We compared the ECHO-AF scores with the previously established HAVOC and ACTEL scores. ROC curves of these scores showed comparably small AUC values compared to the ECHO-AF scores. This discrepancy can be probably attributed to differences in patient collectives that were taken into account—the HAVOC and ACTEL scores were developed using clinical data of patients after stroke or TIA. In these patients, the prevalence of AF was substantially higher compared to our study population. As in case of the HAVOC and ACTEL scores, the C2HEST score represents an easy applicable tool for predicting AF by summing up integer numbers associated with the risk factors coronary artery disease, COPD, hypertension, age above 75 years, systolic heart failure and hyperthyroidism [26]. An internal validation of the score showed an AUC value of 0.75, whereas the external application showed an AUC value of 0.65, which is modest as compared to the ECHO-AF scores evaluated in this study. In general, scores calculated from

summing up integer numbers associated with clinical features can be easily calculated but are less precise than scores represented by more complex equations, as in case of the ECHO-AF scores.

In contrast to the Apple Heart Study and the Huawei Heart Study, this screening trial was performed in a more selective patient cohort with higher pAF prevalence [4, 5]. In this context, it should be noted that the detection of an irregular rhythm by the algorithm used in a smart watch is not equivalent to the detection of AF, which has consequences for the false-positive detection rate of smart watches. It is important to note that rare arrhythmia episodes detected by a smart watch do not imply the same stroke risk as clinically diagnosed AF.

Other scores were developed to predict the 5-year or 10-year risk for developing pAF that require parameters measured by cardiac computed tomography or specialized laboratory tests (NT-proBNP, troponin T) [10, 12]. In contrast to these diagnostic procedures, echocardiography is routinely performed in many cardiological patients. Applying the developed pAF prediction scores in this patient population further increases the clinical value of echocardiographic parameters. Analysing cumulative fractions of AF diagnoses depending on the ECG monitoring duration indicated that an extended ECG monitoring for more than 1 week is required to reliably test for presence of pAF in accordance with previous long-term ECG monitoring studies [27, 28].

This study has limitations that have to be acknowledged. First, carrying durations strongly varied between subjects. Therefore, several cases of pAF might have been unrecognized. Our scores were evaluated in patients of different age groups, with different health profiles, and in hospital patients as well as in outpatient setting, which supports its general applicability. More selected subgroup analyses

could improve the performance of the prediction tool. For example, subanalyses could be conducted to compare between hospital patients and outpatients, patients without or after thromboembolic events, or without and after heart surgical interventions. Additional parameters that were not included in the 47 parameters used to develop the pAF prediction scores, such as biomarkers (BNP, FGF-23, GDF-15), could further increase predictive performance [17, 23]. Validation of our developed pAF prediction scores in a larger study population would further reduce uncertainties of the parameters describing the predictive performance. An independent external prospective validation will be important for testing the general applicability of the ECHO-AF scores.

Conclusion

The novel risk prediction scores adequately predicted pAF based on variables readily available during routine cardiac check-up and echocardiography. Thereby, the value of clinical parameters that are often known in cardiological patients can be increased and the early detection of pAF can be improved. Screening for pAF in patients undergoing echocardiography is reasonable because of a comparably high prevalence of pAF in this patient group. To simplify application of the scores, an online calculator is provided [20]. Collectively, the developed model scores represent a simple, highly sensitive and non-invasive tool for detecting pAF that can be easily implemented in clinical practice and might serve as a screening test to initiate further diagnostic investigations for documenting the presence of pAF.

Acknowledgements We thank Uschi von der Osten (Bayer Vital GmbH) for her continuous support of our project. Finally, we thank all subjects who took part in this study.

Author contributions CS and SMK conceived the research design; CS, SB, PK, FW, SP, AB, ChS were involved in data acquisition; CS, SB, DT, RW, SMK analysed the data; CS and SMK drafted the original version of the manuscript, CS, DT, RW, RE, HK, SMK edited the manuscript; all authors participated in the review of the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by research grants from the University of Heidelberg, from the German Center for Cardiovascular Research (DZHK) [Excellence Grant to C.S., Postdoc Startup Grant to F.W.], from the German Cardiac Society (DGK) [Research Scholarship DGK082018 to F.W.], from the German Heart Foundation/German Foundation of Heart Research [F/15/18 to F.W., F/41/15 and F/03/19 to C.S.], from the Joachim-Herz Foundation [Add-On-Fellowship to F.W.], from the German Research Foundation (DFG) [SCHM 3358/1–1 to C.S., TH 1120/7–1 to D.T.], and from the Else-Kröner-Fresenius Foundation to C.S. [2019_A106]. S.M.K. was supported by the BMBF-funded Heidelberg Center for Human Bioinformatics (HD-HuB) within the German Network for Bioinformatics Infrastructure (de.NBI).

Compliance with ethical standards

Conflict of interest This study was supported by Bayer Vital GmbH (purchase of long-term ECG Holter devices, personnel costs for working with study subjects) as part of a research grant. The funding source had no role in the design of this study, its execution, analyses, interpretation of the data, or publication of the results.

Ethics approval This study was performed in line with the principles of the Declaration of Helsinki. The study protocol was approved by the ethics committee of the University of Heidelberg (Germany, Medical Faculty Heidelberg, S-491/2015).


Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Kotecha D, Breithardt G, Camm AJ et al (2018) Integrating new approaches to atrial fibrillation management: the 6th AFNET/EHRA Consensus Conference. *Europace* 20:395–407. <https://doi.org/10.1093/europace/eux318>
2. January CT, Wann LS, Writing Group Members et al (2019) AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. *Heart Rhythm* 16:e66–e93. <https://doi.org/10.1016/j.hrthm.2019.01.024>
3. Kirchhof P, Breithardt G, Bax J et al (2016) A roadmap to improve the quality of atrial fibrillation management: proceedings from the fifth Atrial Fibrillation Network/European Heart Rhythm Association consensus conference. *Europace* 18:37–50. <https://doi.org/10.1093/europace/euv304>
4. Perez MV, Mahaffey KW, Hedlin H et al (2019) Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med* 381:1909–1917. <https://doi.org/10.1056/NEJMoa1901183>
5. Guo Y, Wang H, Zhang H et al (2019) Mobile photoplethysmographic technology to detect atrial fibrillation. *J Am Coll Cardiol* 74:2365–2375. <https://doi.org/10.1016/j.jacc.2019.08.019>
6. Jones NR, Taylor CJ, Hobbs FDR et al (2019) Screening for atrial fibrillation: a call for evidence. *Eur Heart J*. <https://doi.org/10.1093/eurheartj/ehz834>
7. Schnabel RB, Sullivan LM, Levy D et al (2009) Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. *The Lancet* 373:739–745. [https://doi.org/10.1016/S0140-6736\(09\)60443-8](https://doi.org/10.1016/S0140-6736(09)60443-8)
8. Chamberlain AM, Agarwal SK, Folsom AR et al (2011) A Clinical Risk Score for Atrial Fibrillation in a Biracial Prospective Cohort (from the Atherosclerosis Risk In Communities [ARIC] Study). *Am J Cardiol* 107:85–91. <https://doi.org/10.1016/j.amjcard.2010.08.049>

9. Alonso A, Krijthe BP, Aspelund T et al (2013) Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *J Am Heart Assoc* 2:e000102. <https://doi.org/10.1161/JAHA.112.000102>
10. Sinner MF, Stepas KA, Moser CB et al (2014) B-type natriuretic peptide and C-reactive protein in the prediction of atrial fibrillation risk: the CHARGE-AF Consortium of community-based cohort studies. *Europace* 16:1426–1433. <https://doi.org/10.1093/europace/euu175>
11. Alonso A, Roetker NS, Soliman EZ et al (2016) Prediction of atrial fibrillation in a racially diverse cohort: the multi-ethnic study of atherosclerosis (MESA). *J Am Heart Assoc*. <https://doi.org/10.1161/JAHA.115.003077>
12. Bundy JD, Heckbert SR, Chen LY et al (2020) Evaluation of risk prediction models of atrial fibrillation (from the Multi-Ethnic Study of Atherosclerosis [MESA]). *Am J Cardiol* 125:55–62. <https://doi.org/10.1016/j.amjcard.2019.09.032>
13. Kolek MJ, Graves AJ, Xu M et al (2016) Evaluation of a prediction model for the development of atrial fibrillation in a repository of electronic medical records. *JAMA Cardiol* 1:1007–1013. <https://doi.org/10.1001/jamacardio.2016.3366>
14. Vaziri SM, Larson MG, Benjamin EJ, Levy D (1994) Echocardiographic predictors of nonrheumatic atrial fibrillation. Framingham Heart Study *Circ* 89:724–730. <https://doi.org/10.1161/01.cir.89.2.724>
15. (1992) Predictors of thromboembolism in atrial fibrillation: II. Echocardiographic features of patients at risk. The Stroke Prevention in Atrial Fibrillation Investigators. *Ann Intern Med* 116:6–12. <https://doi.org/10.7326/0003-4819-116-1-6>
16. Rosenberg MA, Gottdiener JS, Heckbert SR, Mukamal KJ (2012) Echocardiographic diastolic parameters and risk of atrial fibrillation: the Cardiovascular Health Study. *Eur Heart J* 33:904–912. <https://doi.org/10.1093/eurheartj/ehr378>
17. Kallenberger SM, Schmid C, Wiedmann F et al (2016) A simple, non-invasive score to predict paroxysmal atrial fibrillation. *PLoS ONE* 11:e0163621. <https://doi.org/10.1371/journal.pone.0163621>
18. Diamond GA (1989) Limited assurances. *Am J Cardiol* 63:99–100. [https://doi.org/10.1016/0002-9149\(89\)91084-9](https://doi.org/10.1016/0002-9149(89)91084-9)
19. Buderer NM (1996) Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad Emerg Med* 3:895–900. <https://doi.org/10.1111/j.1553-2712.1996.tb03538.x>
20. Online calculator for determining pAF prediction scores. <https://www.hidih.org/paf-score-calculator>
21. Kwong C, Ling AY, Crawford MH et al (2017) A clinical score for predicting atrial fibrillation in patients with cryptogenic stroke or transient ischemic attack. *Cardiology* 138:133–140. <https://doi.org/10.1159/000476030>
22. Muscari A, Barone P, Faccioli L et al (2020) Usefulness of the ACTEL score to predict atrial fibrillation in patients with cryptogenic stroke. *Cardiology* 145:168–177. <https://doi.org/10.1159/000505262>
23. Chua W, Purmah Y, Cardoso VR et al (2019) Data-driven discovery and validation of circulating blood-based biomarkers associated with prevalent atrial fibrillation. *Eur Heart J* 40:1268–1276. <https://doi.org/10.1093/eurheartj/ehy815>
24. Kaura A, Arnold AD, Panoulas V et al (2020) Prognostic significance of troponin level in 3121 patients presenting with atrial fibrillation (The NIH Health Informatics Collaborative TROP-AF study). *J Am Heart Assoc* 9:e013684. <https://doi.org/10.1161/JAHA.119.013684>
25. Emma S, Johan E, Faris A-K et al (2015) Mass screening for untreated atrial fibrillation. *Circulation* 131:2176–2184. <https://doi.org/10.1161/CIRCULATIONAHA.114.014343>
26. Li Y-G, Pastori D, Farcomeni A et al (2019) A simple clinical risk score (C2HEST) for predicting incident atrial fibrillation in Asian subjects: derivation in 471,446 Chinese subjects, with internal validation and external application in 451,199 Korean Subjects. *Chest* 155:510–518. <https://doi.org/10.1016/j.chest.2018.09.011>
27. Sanna T, Diener H-C, Passman RS et al (2014) Cryptogenic stroke and underlying atrial fibrillation. *N Engl J Med* 370:2478–2486. <https://doi.org/10.1056/NEJMoa1313600>
28. Gladstone DJ, Spring M, Dorian P et al (2014) Atrial fibrillation in patients with cryptogenic stroke. *N Engl J Med* 370:2467–2477. <https://doi.org/10.1056/NEJMoa1311376>

Affiliations

Constanze Schmidt^{1,2}  · Sebastian Benda^{1,2} · Patricia Kraft³ · Felix Wiedmann^{1,2} · Sven Pleger^{1,3} · Antonius Büscher^{1,2} · Dierk Thomas^{1,2} · Rolf Wachter^{4,5} · Christian Schmid⁶ · Roland Eils^{7,8} · Hugo A. Katus^{1,2} · Stefan M. Kallenberger^{7,8}

✉ Constanze Schmidt
Constanze.Schmidt@med.uni-heidelberg.de

¹ Department of Cardiology, University Hospital Heidelberg, University of Heidelberg, Im Neuenheimer Feld 410, 69120 Heidelberg, Germany

² DZHK (German Center for Cardiovascular Research), Partner Site Heidelberg/Mannheim, University of Heidelberg, Im Neuenheimer Feld 410, 69120 Heidelberg, Germany

³ Kardiologen Am Brückenkopf, Cardiology Practice, Brückenkopfstraße 1/2, 69120 Heidelberg, Germany

⁴ Clinic and Policlinic for Cardiology, University Hospital Leipzig, Liebigstraße 18, 04103 Leipzig, Germany

⁵ Clinic for Cardiology and Pneumology, University Medicine Göttingen, 37099 Göttingen, Germany

⁶ Department of Internal Medicine, GPR Klinikum Rüsselsheim, August-Bebel-Straße 59, 65428 Rüsselsheim am Main, Germany

⁷ Digital Health Center, Berlin Institute of Health (BIH) and Charité, Anna-Louisa-Karsch-Straße 2, 10178 Berlin, Germany

⁸ Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 267, 69120 Heidelberg, Germany