



Distinguishing atrial fibrillation from sinus rhythm using commercial pulse detection systems: The non-interventional BAYathlon study

Digital Health
Volume 7: 1–11
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076211019620
journals.sagepub.com/home/dhj


Christian Müller¹ , Ulf Hengstmann², Michael Fuchs³, Martin Kirchner¹, Frank Kleinjung⁴, Harald Mathis³, Stephan Martin⁵, Ingo Bläse⁶ and Stefan Perings^{6,7}

Abstract

Objective: Early diagnosis of atrial fibrillation (AFib) is a priority for stroke prevention. We sought to test four commercial pulse detection systems (CPDSs) for ability to distinguish AFib from normal sinus rhythm using a published algorithm (Zhou et al., PLoS One 2015;10:e0136544), compared with visual diagnosis by electrocardiogram inspection.

Methods: BAYathlon was a prospective, non-interventional, single-centre study. Adult cardiology patients with documented AFib or sinus rhythm who were due to have a routine 5-min electrocardiogram were randomized to undergo a parallel 5-min pulse assessment with a Polar V800, eMotion Faros 360, TomTom heart rate monitor, or Adidas miCoach Smart Run.

Results: 144 patients (73 with AFib, 71 with sinus rhythm (based on electrocardiograms); median age: 73 years; 53.5% male) were analysed. Algorithm sensitivities (primary endpoint) and specificities for AFib when applied to CPDS recordings were 93.3% and 94.1% with the Polar V800, 90.0% and 84.2% with the eMotion Faros 360, and 0% and 100% with the other CPDSs (analysis period: 127 heart rate signals + 2 min). When applied to routine electrocardiograms, the algorithm correctly detected AFib in 71/73 patients. Different analysis periods (127 heart rate signals +1 or 3 min) only slightly changed the sensitivities with the Polar V800 and eMotion Faros 360 and had no effect on the sensitivities with the other CPDSs.

Conclusion: AFib screening using the applied algorithm is feasible with the Polar V800 and eMotion Faros 360 (which provide RR interval data) but not with the other CPDSs (which provide pre-processed heart rate time series). ClinicalTrials.gov identifier: NCT02875106

Keywords

Atrial fibrillation, electrocardiography, algorithms, sensitivity, observational study, sinus rhythm

Submission date: 5 November 2020; Acceptance date: 4 May 2021

¹Bayer Vital GmbH, Leverkusen, Germany

²Bayer AG, Leverkusen, Germany

³Fraunhofer-Institut für Angewandte Informationstechnik FIT, Sankt Augustin, Germany

⁴Bayer AG, Berlin, Germany

⁵Verbund Katholischer Kliniken Düsseldorf, Düsseldorf, Germany

⁶Cardio Centrum Düsseldorf, Düsseldorf, Germany

⁷Division of Cardiology, Pulmonology and Vascular Medicine, Department of Internal Medicine, Heinrich-Heine-University, Düsseldorf, Germany

Corresponding author:

Christian Müller, Bayer Vital GmbH, Medical Excellence & Innovation Management, Data Generation, Building K 56, 51368 Leverkusen, Germany. Email: christian.mueller4@bayer.com

Twitter: @Cjmueller4



Introduction

Atrial fibrillation (AFib) is common, with an estimated age-dependent prevalence of ~3% in adults aged >20 years, and is a major cause of stroke, heart failure, sudden death and cardiovascular morbidity.¹ Early diagnosis of AFib is a recognized priority for the prevention of strokes. AFib is detected as irregular RR intervals and absence of distinct P waves in an electrocardiogram (ECG).¹ However, AFib generally begins with paroxysmal episodes² which can be missed by short-term monitoring during office visits.^{3,4} Moreover, about one-third of patients with AFib are asymptomatic.⁵ The consequences of paroxysmal and asymptomatic AFib remain severe. The risk of thromboembolic events associated with paroxysmal AFib was found to be as high as that associated with persistent/permanent AFib in the Euro Heart Survey⁶ and the AFib Clopidogrel Trial With Irbesartan for Prevention of Vascular Events (ACTIVE W) trial.⁷ Outcomes in asymptomatic AFib have been reported to be similar to or worse than outcomes in symptomatic AFib.⁵

Long-term monitoring methods are therefore needed to improve the detection of asymptomatic and paroxysmal AFib in populations at risk of developing AFib or stroke (e.g. the elderly, those with hypertension, heart failure or diabetes mellitus, and especially those with ischaemic stroke^{1,5}). There is likely to be an optimal screening duration and frequency for the detection of paroxysmal AFib in at-risk populations. Implanted devices are used to achieve continuous monitoring in patients with cryptogenic stroke,⁸ but such invasive monitoring would not be suitable for AFib screening in broader populations. Many non-invasive mobile health devices have been evaluated as AFib screening tools, with results showing variable accuracy depending on the technology used, the application setting and the study population.⁹

An AFib screening algorithm that is compatible with multiple commercial pulse detection systems (CPDSs) may help to increase the accessibility of AFib screening by addressing different consumer/researcher needs and budgets. The aim of the BAYathlon study was to test four CPDSs (Polar V800 with a Polar H7 chest strap monitor (Polar Electro Oy, Kempele, Finland), eMotion Faros 360 set with belt (Mega Electronics Ltd, Kuopio, Finland; now marketed as Bittium Faros 360 by Bittium Corporation, Oulu, Finland), TomTom heart rate monitor (HRM; TomTom, Amsterdam, the Netherlands)) and Adidas miCoach Smart Run (Adidas AG, Herzogenaurach, Germany)) for their ability to distinguish AFib from normal sinus rhythm using a previously published AFib detection algorithm.¹⁰ The CPDSs were chosen for their ease of use, lossless data

recording during a wide range of physical activities, open access to the data, and exported data quality for use with AFib detection algorithms.

Methods

Study design and patients

BAYathlon was a prospective, non-interventional, company-sponsored, feasibility study conducted in a single private cardiac practice (the Cardio Centrum Düsseldorf, Düsseldorf, Germany) which routinely conducts ECG recordings. Consecutive adult female and male patients were eligible for inclusion if their treating cardiologist had decided that they should have an ECG recorded as part of routine clinical practice, and if they had AFib or sinus rhythm documented in their medical history. Patients were excluded if they had a pacemaker or any comorbidities or abnormalities of heart function or rhythm which might, in the opinion of the investigator, interfere with the evaluation of study data. Participation in an investigational program with interventions outside of routine clinical practice was also an exclusion criterion.

The investigator collected demographic data from medical records if available, or by interviewing the patient. Enrolled patients were randomly assigned to undergo assessment with one of four CPDSs (Polar V800 with a Polar H7 chest strap monitor, eMotion Faros 360 set with belt, TomTom HRM and Adidas miCoach Smart Run) in parallel with a routine ECG recording using a certified medical device (Schiller CS-200 or AT-104; Schiller AG, Baar, Switzerland) (Figure 1). Randomization was stratified by patient population (AFib or normal sinus rhythm in medical history). The CPDS and routine ECG data were captured and exported via an open interface into the study database over a recording period of 5 min, during which the patients were at rest. As a reference, the routine ECG data were visually reviewed by the investigator to determine a diagnosis of AFib or sinus rhythm (without automated input).

The study protocol was approved by the Independent Ethics Committee (Ethik-Kommission der Ärztekammer Nordrhein, Düsseldorf; file reference: 2016/333). All participating patients provided signed informed consent. The study is registered on ClinicalTrials.gov (identifier: NCT02875106).

Data processing for automatic detection of AFib by the algorithm

Automatic classification of CPDS and ECG data as AFib or non-AFib required four processing steps, described below and summarized in Figure 2.

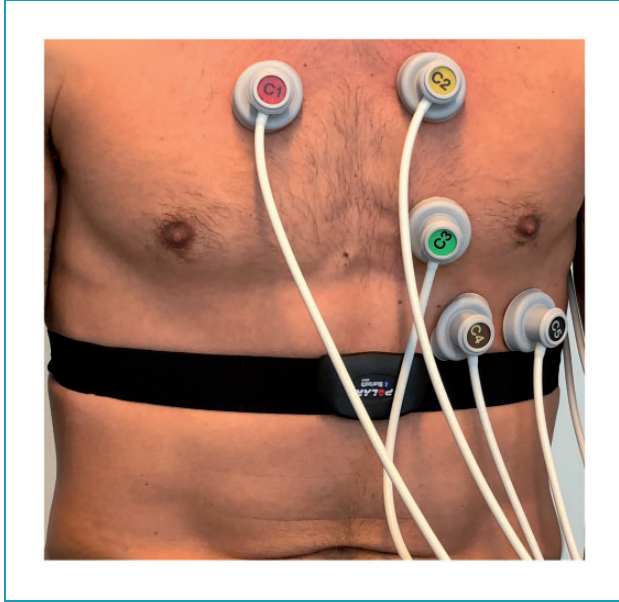


Figure 1. Parallel assessment by routine electrocardiogram and a commercial pulse detection system. Patients underwent parallel assessment by routine electrocardiogram and a commercial pulse detection system (the Polar V800 in this example) for 5 minutes.

1. Conversion of raw data to heart rate (HR) time series. The TomTom HRM and Adidas miCoach Smart Run deliver pre-processed pulse detection data as a HR time series, bypassing the need for an initial processing step.

The Polar V800 provides pulse detection data as a list of consecutive RR intervals (t_{RR}). This list was converted to a HR time series with the following formula:

$$HR/\text{min} = (60 \text{ s}/\text{min})/t_{RR} \text{ s}$$

where HR is the corresponding HR value for each RR interval.

The eMotion Faros 360 provides pulse detection data as a one-channel ECG record with a sampling rate of 1000 entries per second. RR intervals were extracted from the raw data using the Java software library OSEA-4-Java (Version 1.0.0, <https://github.com/MEDEVIT/OSA-4-Java>)¹¹ and converted to a HR time series using the formula described above for the Polar V800.

The Schiller CS-200 and AT-104 provide a 12-channel ECG record. For automatic classification via the algorithm, channel number 1 was analysed as described above for the eMotion Faros 360, using the Java software library OSEA-4-Java (for diagnosis by visual ECG inspection, the investigator reviewed all 12 channels).

To explore the effect of data pre-processing on the function of the algorithm, a moving average of the RR intervals from the Polar V800 and the eMotion Faros

360 was calculated for different window sizes (2–30 RR intervals) starting at the data entry with index $w/2$ and ending at index $N - w/2$, where w is the window size and N is the number of entries in the data sample. The moving average values were then converted to a HR time series as described above.

2. Definition of analysis periods. Due to the design and working mechanism of the algorithm, analysis periods were constructed by taking the first 127 HR signals plus the following 1, 2 and 3 min of the HR time series.

The effect of data pre-processing on the function of the algorithm was evaluated using data from the whole 5-min recording period ($N - w$ moving average values).

3. Detection of AFib episodes. The data were searched for AFib episodes using the algorithm developed and published by Zhou et al.¹⁰ Briefly, the algorithm identifies AFib episodes in a moving window containing 127 consecutive HR signals, which roughly corresponds to 2 min if an average HR of 64 beats/min is assumed. Zhou et al. convert the symbol sequence into a word sequence using a novel operator. The probability distribution of the word sequence in the specified space is obtained, and a coarser version of Shannon entropy (SE) is next employed to discriminate the AFib arrhythmias. The novelty of this proposed method is based on the HR which was firstly introduced in the field of AFib detection.¹⁰

4. Determination of overall classification (AFib/non-AFib). If the recording interval contained ≥ 5 consecutive AFib episodes, the interval was classified as AFib. Otherwise, it was classified as non-AFib (sinus rhythm).

Endpoints

Taking the investigator's diagnosis based on visual assessment of the routine ECG data as the reference, the primary endpoint was the rate of correctly detected AFib using the CPDS based on an analysis period of 127 HR signals + 2 min (sensitivity of the CPDS for AFib).

The secondary endpoints were the rate of correctly detected non-AFib cases using the CPDS based on the analysis period of 127 HR signals + 2 min (specificity of the CPDS for AFib), the rate of correctly detected AFib when applying the AFib algorithm to the ECG data over the whole 5-min recording period (sensitivity of the algorithm for AFib), and the rate of correctly detected AFib using the CPDS based on analysis periods of 127 HR signals + 1 min and + 3 min (sensitivity analysis in different time frames).

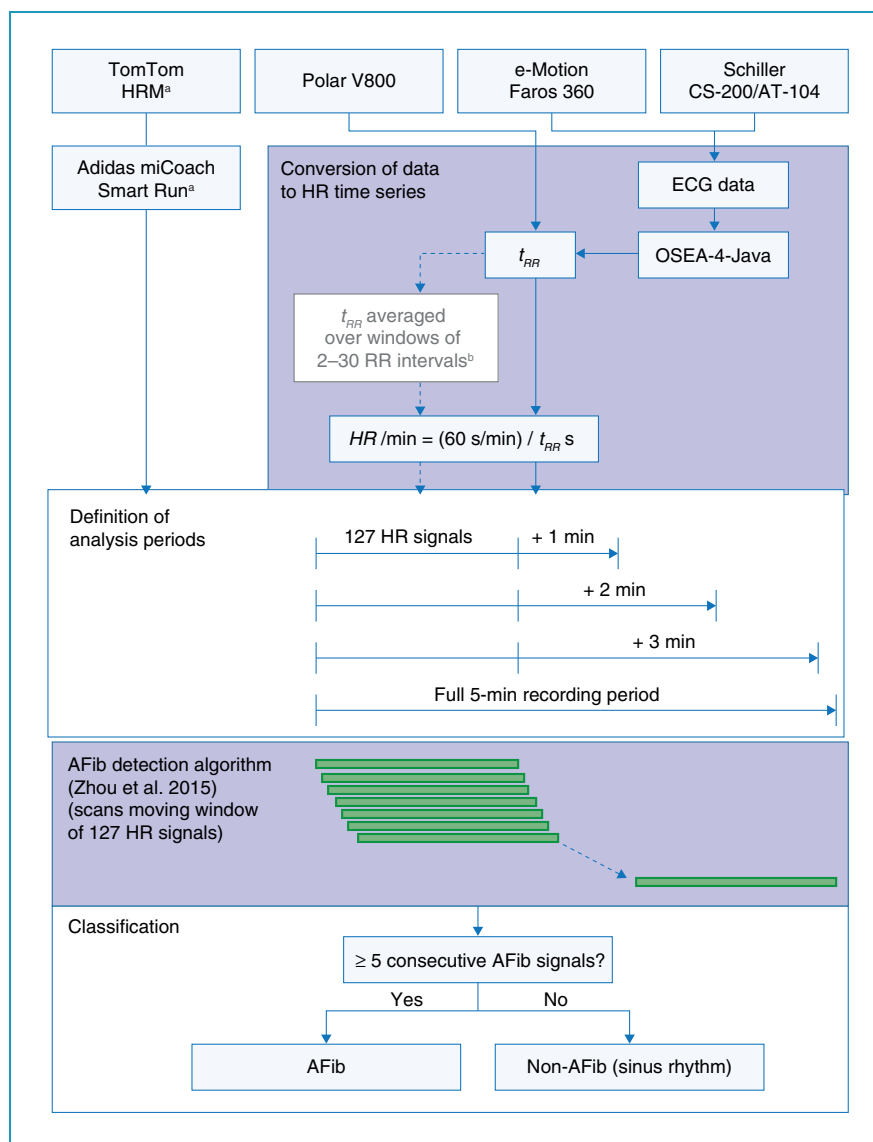


Figure 2. Overview of data processing steps. Patients with AFib or sinus rhythm underwent parallel assessment by routine ECG and one of four commercial pulse detection systems. Data were analysed using the AFib detection algorithm of Zhou et al.¹⁰

^aData from these devices are already delivered as HR time series.

^bMoving average RR intervals over different window sizes were calculated from Polar V800 and eMotion Faros 360 data in patients with AFib (diagnosed by the investigator based on the routine ECG) to explore the effect of data pre-processing on the function of the algorithm.

AFib: atrial fibrillation; ECG: electrocardiogram; HR: heart rate; HRM: heart rate monitor; OSEA-4-Java: open source ECG analysis software for Java (Version 1.0.0, <https://github.com/MEDEVIT/OSEA-4-Java>); t_{RR} : RR intervals.

Statistical analysis

Statistical analyses were exploratory and descriptive. The script language Python 3.6.7 (<https://www.python.org/>; Python Software Foundation, Beaverton, Oregon, USA) was used with the packages NumPy 1.13.3 (<https://numpy.org/>) and Pandas 0.23.4 (<https://pandas.pydata.org/>). All analyses were performed in the total study population. Missing data

were not replaced. Sensitivity and specificity for AFib were defined as

$$\text{Sensitivity} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

$$\text{Specificity} = \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}}$$

Sensitivity and specificity are presented with 95% confidence intervals (Wilson confidence limits). Cohen's κ^{12} was calculated to judge the level of agreement with the investigator's diagnosis based on the routine ECG data.

For an expected sensitivity and specificity of 100.0% for the diagnosis of AFib, the sensitivity and specificity can be estimated with a precision of 16.1% (95% Wilson confidence limits). Based on the expected sensitivity and specificity and the width of the confidence interval in the PULSE-SMART study,¹³ a sample size of 20 patients with AFib and 20 patients with sinus rhythm per CPDS was chosen.

The effect of data pre-processing on the algorithm output was evaluated descriptively in patients with AFib (diagnosed by the investigator based on the routine ECG) who were assessed using the Polar V800 or the eMotion Faros 360.

Results

Patients

In total, 163 patients were screened, 161 patients were enrolled, and 144 had full data available for analysis (Figure 3). In each of the four CPDS groups, the median age was higher in patients with a medical history of AFib than in those with sinus rhythm (Table 1).

In three of 74 patients with AFib documented in their medical history, AFib was detected in the ECG data by the investigator but not by the algorithm. In a post hoc analysis, the ECGs from these three patients were re-evaluated by the investigators and one was found to have been incorrectly classified by them. This patient was re-classified as having sinus rhythm, giving a total of 73 patients with AFib and 71 with sinus rhythm diagnosed by the investigator based on the routine ECG.

Sensitivity and specificity of CPDSs and algorithm for AFib

The sensitivity (primary endpoint) and specificity of the algorithm for correctly detected AFib when applied to CPDS recordings for an analysis period of 127 HR signals + 2 min were 93.3% and 94.1%, respectively, with the Polar V800, 90.0% and 84.2%, respectively, with the eMotion Faros 360, and 0% and 100%, respectively, with the other CPDSs (Table 2).

Pre-processing of the RR interval data affected the algorithm output. As the window size for averaging of RR intervals increased, the rate of correct detection of AFib decreased, falling to zero at window sizes of 15 and 20 RR intervals for the Polar V800 and the eMotion Faros 360, respectively (Supplemental material online, Table S1). In one patient with AFib, the

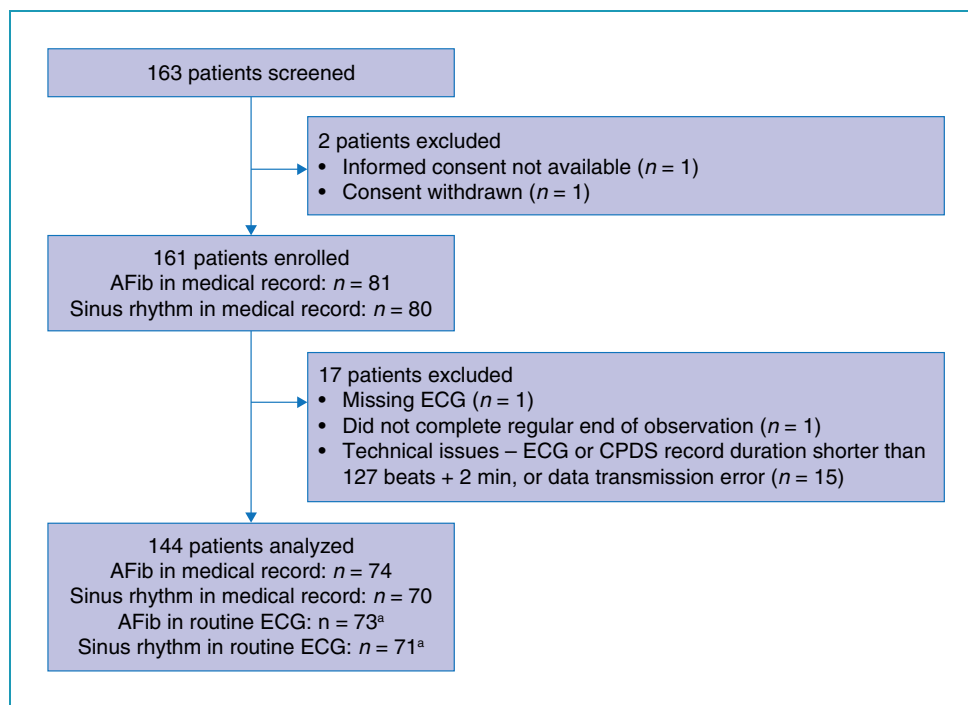


Figure 3. Patient disposition. Of 163 patients who were screened, 161 were enrolled and 144 were analysed.

^aRoutine ECG data obtained during the study were visually reviewed by the investigator to determine a diagnosis of AFib or sinus rhythm. AFib: atrial fibrillation; CPDS: commercial pulse detection system; ECG: electrocardiogram.

Table 1. Patient demographics.

	Polar V800	eMotion Faros 360	TomTom HRM	Adidas miCoach Smart Run	All
Atrial fibrillation in medical record					
<i>n</i>	16	20	20	18	74
Male/female, <i>n/n</i>	12/4	9/11	13/7	10/8	44/30
Age, years	76 (55–85)	80 (64–92)	80 (55–88)	78 (56–88)	78 (55–92)
Sinus rhythm in medical record					
<i>n</i>	16	19	17	18	70
Male/female, <i>n/n</i>	10/6	7/12	6/11	10/8	33/37
Age, years	66 (25–88)	72 (48–82)	61 (44–87)	62 (22–80)	66 (22–88)
All					
<i>n</i>	32	39	37	36	144
Male/female, <i>n/n</i>	22/10	16/23	19/18	20/16	77/67
Age, years	72 (25–88)	75 (48–92)	73 (44–88)	72 (22–88)	73 (22–92)

Age is reported as median (minimum–maximum).

HRM: heart rate monitor.

Polar V800 provided HR rather than RR interval data because of technical issues during data acquisition – this patient was also not correctly identified by the algorithm.

The sensitivity of the algorithm for correctly detected AFib based on 5 min of ECG data was 97.3% (Table 2). The two undetected cases showed, according to the investigators, pseudo-arrhythmic AFib (Figure 4(a)) and AFib with partly regular transition (Figure 4(b)). A Poincaré plot from a patient with correctly detected AFib is shown in Figure 4(c). The associated Poincaré plots for the undetected cases are presented in Figure 4(d) (pseudo-arrhythmic AFib) and 4(e) (partly regular transition).

The sensitivities and specificities of the algorithm for correctly detected AFib based on recordings from the Polar V800 and eMotion Faros 360 showed only slight changes when different analysis periods were used (Table 2). The use of different analysis periods had no effect on the sensitivity of the algorithm for AFib when applied to recordings from the other CPDSs. The level of agreement with the investigator's diagnosis (Cohen's κ) also showed little change when different analysis periods were used (Table 3).

Discussion

In the BAYathlon study, the AFib screening algorithm of Zhou et al.¹⁰ showed good sensitivity for AFib when used with CPDSs that provided raw RR interval data (Polar V800 and eMotion Faros 360) but not when used with CPDSs that provided pre-processed HR time series (TomTom HRM and Adidas miCoach Smart Run). The algorithm also correctly identified 71/73 patients with AFib when applied to routinely collected ECG data. As the algorithm used a moving time window of 127 consecutive HR signals, it showed no large differences in sensitivity for AFib when used over different time intervals.

Algorithms for automated AFib detection have mostly been based on RR interval irregularity in ECGs.¹⁰ These algorithms have limitations and in some cases their complexity renders them unsuitable for use in wearable devices. The algorithm developed by Zhou et al. and used in the BAYathlon study provides a real-time and low-complexity method for discriminating AFib episodes based on HR instead of RR interval irregularity.¹⁰ However, the performance of the algorithm (similar to other algorithms for AFib screening) is dependent on the training data set and

Table 2. Sensitivity and specificity of AFib detection algorithm when applied to recordings from CPDSs and ECGs.

	Sensitivity and specificity, % (95% confidence interval) [n/n]				
	CPDSs providing RR interval data		CPDSs providing pre-processed HR time series		
	Polar V800 ^a	eMotion Faros 360	TomTom HRM	Adidas miCoach Smart Run	
Primary endpoint				ECG with algorithm	
Sensitivity for AFib	93.3 (70.2–98.8) [14/15]	90.0 (69.9–97.2) [18/20]	0.0 (0.0–16.1) [0/20]	0.0 (0.0–17.6) [0/18]	-
Secondary endpoints					
Specificity for AFib	94.1 (73.0–99.0) [16/17]	84.2 (62.4–94.5) [16/19]	100.0 (81.6–100.0) [17/17]	100.0 (82.4–100.0) [18/18]	-
Sensitivity for AFib	-	-	-	-	97.3 (90.5–99.2) [71/73]
Sensitivity for AFib by analysis period:					
127 HR signals + 1 min	93.3 (70.2–98.8) [14/15]	90.0 (69.9–97.2) [18/20]	-	-	-
127 HR signals + 3 min	100.0 (74.1–100.0) [11/11]	94.1 (73.0–99.0) [16/17]	-	-	-

Reference is investigator's diagnosis of AFib/sinus rhythm based on manual review of the ECG data. Analysis period is 5 min for the routine ECG and 127 HR signals + 2 min for the CPDSs, except where otherwise specified.

AFib: atrial fibrillation; CPDSs: commercial pulse detection systems; ECG: electrocardiogram; HR: heart rate; HRM: heart rate monitor.

^aFive recordings with the Polar V800 (one in a patient with AFib and four in patients with sinus rhythm) provided HR rather than RR intervals because of technical issues during data acquisition.

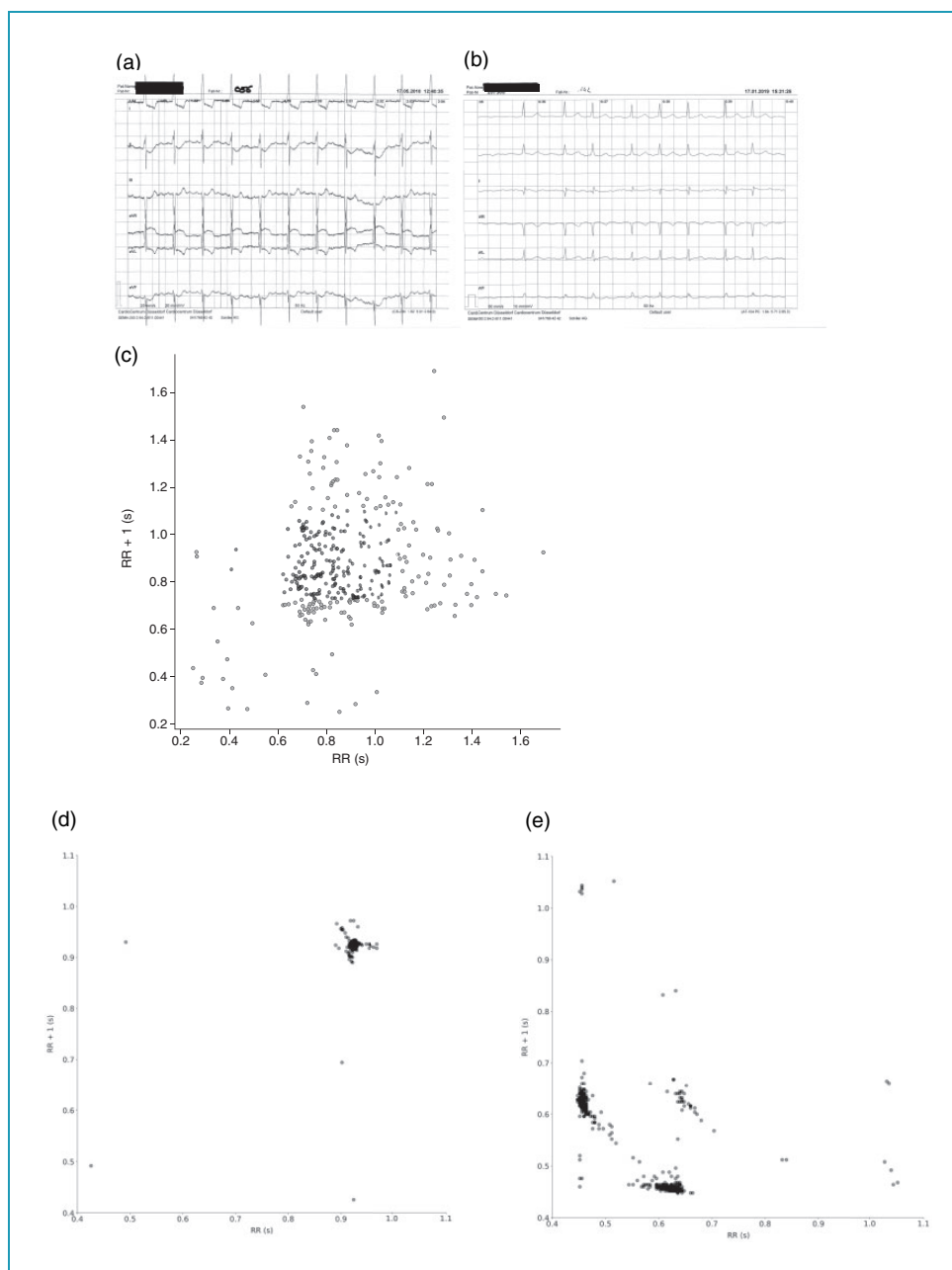


Figure 4. (a)–(b) ECGs from two patients with AFib who were incorrectly classified by the algorithm based on their routine ECG data, and (c) a Poincaré plot of correctly detected AFib. The ECGs and Poincaré plots of the two incorrectly classified patients showed (a) and (d) pseudo-arrhythmic AFib and (b) and (e) AFib with partly regular transition. AFib: atrial fibrillation; ECG: electrocardiogram.

may be susceptible to errors in certain circumstances, for example in the presence of extra systoles, intracardiac devices and/or atrial flutter.

In the BAYathlon study, the algorithm missed two cases of AFib when applied to routine ECG data. After post hoc review of the ECGs, these cases were considered to be closer to atrial flutter than AFib. Irregular RR intervals are common in atrial flutter,¹⁵ but the lack of a well-defined QRS complex in atrial flutter

may prevent the correct detection of QRS complexes by the OSEA-4-Java library, which may in turn prevent the algorithm of Zhou et al. from detecting the irregularities. Re-optimizing the complexity threshold of the algorithm by training on AFib and atrial flutter series may also increase the specificity of AFib detection in the presence of other arrhythmias.

When the algorithm was applied to read-outs from the Adidas miCoach Smart Run and the TomTom

Table 3. Cohen's κ^a for agreement of AFib detection algorithm (applied to recordings from CPDS and ECG) with investigator's diagnosis of AFib/sinus rhythm (based on visual assessment of the ECG data).

Analysis period	CPDSs providing RR interval data		CPDSs providing pre-processed HR time series		
	Polar V800 ^b	eMotion Faros 360	TomTom HRM	Adidas miCoach Smart Run	ECG with algorithm
127 HR signals					
+ 1 min	0.94	0.74	0.00	0.00	0.94
+ 2 min	0.87	0.74	0.00	0.00	0.94
+ 3 min	0.91	0.77	-	-	0.98

AFib: atrial fibrillation; CPDSs: commercial pulse detection systems; ECG: electrocardiogram; HR: heart rate; HRM: heart rate monitor.

^aThe strength of $\kappa > 0.60$ and ≤ 0.80 is considered 'substantial', and $\kappa > 0.80$ is considered 'almost perfect'.¹⁴

^bFive of the 32 recordings with the Polar V800 (one in patients with AFib and four in patients with sinus rhythm) provided HR rather than RR intervals because of technical issues during data acquisition.

HRM, no patients with AFib could be detected. These results suggest that the algorithm does not detect AFib when used with devices that provide pre-processed HR time series rather than RR intervals. Although it is unclear precisely how the Adidas miCoach Smart Run and the TomTom HRM pre-processed their data, we found that averaging RR intervals over a window of 15 or 20 RR intervals prevented the algorithm from detecting AFib. In the single case in which the Polar V800 provided HR rather than RR interval data (resulting in incorrect classification as non-AFib (sinus rhythm)), the data output was also most likely averages of consecutive HR values.

According to a European Heart Rhythm Association consensus document, screening for AFib is advised in high risk populations because of its cost-effectiveness.⁵ The use of consumer devices in combination with an AFib detection app may provide new insights that could improve the efficiency of health service provision. For example, the recent Huawei Heart study in China showed geographical variation in the proportion of detected AFib (using Huawei wristbands/watches with a photoplethysmography algorithm), suggesting a need for different prevention approaches in different regions.¹⁶ To offer clinicians an alternative for AFib screening, Fraunhofer and Bayer have developed the BAYathlon app (Supplemental material online, Figure S1), which uses the same source data (RR intervals) and algorithm as described in the current study (but was not used as a data collection tool in this study). The app performs ongoing HR and activity plotting and alerts the user if and to what extent AFib is detected. In contrast to the Apple and Huawei AFib detection systems which are dedicated to specific devices, the BAYathlon app aims to have broad compatibility with devices with open interfaces that provide raw RR interval data.

The BAYathlon app may therefore facilitate use of the AFib screening algorithm of Zhou et al.¹⁰ in a real-life clinical setting.

While further studies are required to validate the use of the algorithm with other CPDSs, our results suggest that the algorithm may be used with the Polar V800 and the eMotion Faros 360 to screen for AFib. Both devices used electrode-containing chest strap monitors. The Polar V800 chest strap monitor (Polar H7) was previously shown to have high sensitivity (96.3% (95% confidence interval: 89.7–99.2%)) and specificity (98.2% (95% confidence interval: 96.2–99.3%)) for the detection of AFib (using an algorithm that employed turning points, clustering within Lorentz plots and Shannon entropy) when compared with a 12-lead ECG interpreted by a panel of cardiologists.^{17,18} The Polar H7 also had better agreement with ECG data during exercise than eight different optically-based forearm- or wrist-worn HR monitors across two studies, suggesting that electrode-containing chest monitors may be more suitable than optical sensors for accurate assessment of HR during exercise.^{19,20} Athletes might therefore prefer chest-worn ECG devices such as the Polar V800 and the eMotion Faros 360 over optical sensors for HR measurement. As athletes are at significantly increased risk of AFib compared with the general population,^{21,22} addition of an AFib screening algorithm to chest-worn ECG devices may help to target AFib screening to an at-risk population.

Study limitations

Limitations of the BAYathlon study include its single-site design and limited sample size. There was an imbalance in age distribution across the device groups, particularly in the subset with sinus rhythm. The use of medications which may have influenced HR was not

recorded. As BAYathlon was a feasibility study, further real-life studies are needed to provide more generalizable results. For example, all included patients had either AFib or sinus rhythm; the ability of the algorithm to distinguish AFib from other arrhythmias was not investigated. ECG and CPDS recordings were taken at rest, so the influence of exercise on the results is not known. The complexity threshold of the algorithm (originally optimized for unprocessed RR series) was not adapted to the different device data evaluated in the current study. Two devices were not able to detect AFib; to compare this with the devices which were able to detect AFib, we calculated the moving average RR intervals over different window sizes (2–30 RR intervals; the window size of 1 indicates raw RR interval data without averaging) for the Polar V800 and the eMotion Faros 360 devices. The averaging of four RR intervals reduced the ability of AFib detection by half (Supplemental material online, Table S1), therefore the unprocessed recordings can be used for AFib detection with that algorithm and processed (averaged values) are not sufficient. Finally, the results of this study apply only to the four tested CPDSs; further studies are required to validate the use of the algorithm with other CPDSs that generate raw RR interval data.

Conclusions

The results of BAYathlon suggest that screening for AFib using the applied algorithm in combination with the Polar V800 and the eMotion Faros 360 devices is feasible when performed in resting individuals for a short interval of 127 HR signals + 1 to 3 min. Further studies are needed to evaluate the performance of the algorithm with other CPDSs that provide raw RR interval data and with long-term continuously collected CPDS data from active individuals. The BAYathlon app may facilitate such studies. Use of averaged rather than individual RR intervals reduces the ability of the algorithm to detect AFib, and the algorithm does not detect AFib when used with devices that provide pre-processed HR time series rather than RR intervals (TomTom HRM and Adidas miCoach Smart Run).

Acknowledgements: The authors thank Daniel Wolff (Bayer AG, Wuppertal, Germany) for data management and Miriam Tamm (Bayer AG, Wuppertal, Germany) for statistical support. Editorial assistance was provided by Dr Claire Mulligan and Dr Paul Overton (Beacon Medical Communications Ltd, Brighton, UK), funded by Bayer Vital GmbH (Leverkusen, Germany), in accordance with Good Publication Practice (GPP3) guidelines (<http://www.ismpp.org/gpp3>).

Contributorship: CM and UH are responsible for the study design and drafting of this article, and contributed to the

protocol, set-up, and conduct of the study. MF is responsible for the design and development of the data acquisition software for the study. MF contributed to the statistical analysis of the study, the study database administration and drafting of this article. MK is responsible for administration and drafting of this article and contributed to the conduct of the study. FK is responsible for statistical analysis and drafting of this article and contributed to conduct of the study. HM contributed to the concept and design of the study protocol, analysis and interpretation of data, and discussion of the manuscript content. SM contributed to the concept and design of the study protocol, analysis and interpretation of data, and discussion of the manuscript content. IB contributed to the conduct of the study, analysis and interpretation of data, and discussion of the manuscript content. SP contributed to the conduct of the study, analysis and interpretation of data, and discussion of the manuscript content.

Declaration of conflicting interests: CM and MK are employees of Bayer Vital GmbH (Leverkusen, Germany). UH is an employee of Bayer AG (Leverkusen, Germany). CM and UH are inventors for a patent application of the BAYathlon app on behalf of Bayer AG (not yet granted). MF has no relationships relevant to the content of this paper to disclose. FK is an employee of Bayer AG (Berlin, Germany). SM received consultancy fees from Bayer Vital GmbH. HM and IB have no relationships with industry that might pose a conflict of interest in connection with the submitted article. SP is Editor-in-Chief of *www.kardiologie.org* which is partly sponsored by Bayer Vital GmbH.

Ethical approval: The study protocol was approved by the Independent Ethics Committee (Ethik-Kommission der Ärztekammer Nordrhein, Düsseldorf; file reference: 2016/333). All participating patients provided signed informed consent. Written consent was provided by the patient for publication of the photograph shown in Figure 1.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Bayer Vital GmbH, Leverkusen, Germany.

Guarantor: CM.

Data availability statement: The data underlying this article will be shared on reasonable request to the corresponding author.

Peer review: This manuscript was reviewed by reviewers who have chosen to remain anonymous.

ORCID iD: Christian Müller  <https://orcid.org/0000-0002-3722-7685>

Supplemental material: Supplemental material for this article is available online.

References

1. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016; 37: 2893–2962.
2. Heijman J, Voigt N, Nattel S, et al. Cellular and molecular electrophysiology of atrial fibrillation initiation, maintenance, and progression. *Circ Res* 2014; 114: 1483–1499.
3. Kirchhof P, Bax J, Blomstrom-Lundquist C, et al. Early and comprehensive management of atrial fibrillation: executive summary of the proceedings from the 2nd AFNET-EHRA consensus conference ‘research perspectives in AF’. *Eur Heart J* 2009; 30: 2969–2977c.
4. McConnell MV, Turakhia MP, Harrington RA, et al. Mobile health advances in physical activity, fitness, and atrial fibrillation: moving hearts. *J Am Coll Cardiol* 2018; 71: 2691–2701.
5. Mairesse GH, Moran P, Van Gelder IC, ESC Scientific Document Group, et al. Screening for atrial fibrillation: a European Heart Rhythm Association (EHRA) consensus document endorsed by the Heart Rhythm Society (HRS), Asia Pacific Heart Rhythm Society (APHRS), and sociedad latinoamericana de estimulación cardiaca y electrofisiología (SOLAECE). *Europace* 2017; 19: 1589–1623.
6. Nieuwlaat R, Dinh T, Olsson SB, et al. Should we abandon the common practice of withholding oral anticoagulation in paroxysmal atrial fibrillation? *Eur Heart J* 2008; 29: 915–922.
7. Hohnloser SH, Pajitnev D, Pogue J, et al. Incidence of stroke in paroxysmal versus sustained atrial fibrillation in patients taking oral anticoagulation or combined antiplatelet therapy: an ACTIVE W substudy. *J Am Coll Cardiol* 2007; 50: 2156–2161.
8. Giancaterino S, Lupercio F, Nishimura M, et al. Current and future use of insertable cardiac monitors. *JACC Clin Electrophysiol* 2018; 4: 1383–1396.
9. Giebel GD and Gissel C. Accuracy of mHealth devices for atrial fibrillation screening: systematic review. *JMIR Mhealth Uhealth* 2019; 7: e13641.
10. Zhou X, Ding H, Wu W, et al. A Real-Time atrial fibrillation detection algorithm based on the instantaneous state of heart rate. *PLoS One* 2015; 10: e0136544.
11. Hamilton PS. Open source ECG analysis software documentation, www.eplimited.com/osea13.pdf (2002, accessed 26 August 2019).
12. Kwicien R, Kopp-Schneider A and Blettner M. Concordance analysis: part 16 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2011; 108: 515–521.
13. McManus DD, Chong JW, Soni A, et al. PULSE-SMART: pulse-based arrhythmia discrimination using a novel smartphone application. *J Cardiovasc Electrophysiol* 2016; 27: 51–57.
14. Landis JR and Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
15. Krummen DE, Feld GK and Narayan SM. Diagnostic accuracy of irregularly irregular RR intervals in separating atrial fibrillation from atrial flutter. *Am J Cardiol* 2006; 98: 209–214.
16. Guo Y, Wang H, Zhang H, et al. Mobile photoplethysmographic technology to detect atrial fibrillation. *J Am Coll Cardiol* 2019; 74: 2365–2375.
17. Lown M, Yue AM, Shah BN, et al. Screening for atrial fibrillation using economical and accurate technology (from the SAFETY study). *Am J Cardiol* 2018; 122: 1339–1344.
18. Lown M, Yue A, Lewith G, et al. Screening for Atrial Fibrillation using Economical and accurate Technology (SAFETY) – a pilot study. *BMJ Open* 2017; 7: e013535.
19. Gillinov S, Etiwy M, Wang R, et al. Variable accuracy of wearable heart rate monitors during aerobic exercise. *Med Sci Sports Exerc* 2017; 49: 1697–1703.
20. Wang R, Blackburn G, Desai M, et al. Accuracy of wrist-worn heart rate monitors. *JAMA Cardiol* 2017; 2: 104–106.
21. Ayinde H, Schweizer ML, Crabb V, et al. Age modifies the risk of atrial fibrillation among athletes: a systematic literature review and meta-analysis. *Int J Cardiol Heart Vasc* 2018; 18: 25–29.
22. Li X, Cui S, Xuan D, et al. Atrial fibrillation in athletes and general population: a systematic review and meta-analysis. *Medicine (Baltimore)* 2018; 97: e13405.