European Journal of Internal Medicine 73 (2020) 72–75

Contents lists available at ScienceDirect

European Journal of Internal Medicine

journal homepage: www.elsevier.com/locate/ejim

Original article

Prospective blinded Evaluation of the smartphone-based AliveCor Kardia ECG monitor for Atrial Fibrillation detection: The PEAK-AF study

Felix K. Wegner[a,*], Simon Kochhäuser[a,b], Christian Ellermann[a], Philipp S. Lange[a], Gerrit Frommeyer[a], Patrick Leitz[a], Lars Eckardt[a], Dirk G. Dechering[a]

[a] Department of Electrophysiology, University Hospital Muenster, Muenster, Germany
[b] Current address: Department of Cardiology and Vascular Medicine, West German Heart and Vascular Center Essen, University Duisburg-Essen, Essen, Germany

ARTICLE INFO

Keywords:
Atrial fibrillation
Stroke, ECG
EHealth
Prevention

ABSTRACT

Introduction: The AliveCor Kardia ECG monitor (ACK) offers a smartphone-based one-lead ECG recording for the detection of atrial fibrillation. We compared ACK lead I recordings with the 12-lead ECG and introduce a novel parasternal lead (NPL).

Methods: Consecutive cardiac inpatients were recruited. In all patients a 12-lead ECG, ACK lead I and NPL were obtained. Two experienced electrophysiologists were blinded and separately evaluated all ECG. We calculated sensitivity, specificity, and predictive values of the ACK ECG compared to the 12-lead ECG.

Results: 296 ECG from 99 patients (38 female, age 64 ± 15 years, BMI 27.8 ± 5.1 kg/m^2) were analyzed. 20% of ACK lead I recordings contained a critical amount of artifact. The electrophysiologists' interpretation of the ACK recordings yielded a sensitivity of 100% and specificity of 94% for atrial fibrillation or flutter in lead I ($\kappa = 0.90$) and a sensitivity of 96% and specificity of 97% in the NPL ($\kappa = 0.92$). The ACK diagnostic algorithm displayed a significantly lower sensitivity (55–70%), specificity (60–69%), and accuracy ($\kappa = 0.4$–0.53) but a high negative predictive value (100%). Patients with atrial flutter ($n = 5$) and with ventricular stimulation ($n = 12$) had a high likelihood of being misclassified by the algorithm.

Conclusion: The AliveCor Kardia ECG monitor allows a highly accurate detection of atrial fibrillation by an interpreting electrophysiologist both in the standard lead I and a novel parasternal lead. The diagnostic algorithm offered by the system may be useful in screening recordings for further review. Diagnostic challenges present in atrial flutter and ventricular pacemaker stimulation.

1. Introduction

Stroke is the second leading cause of death worldwide and causes significant morbidity, especially if the source is a cardiogenic embolism [1,2]. Atrial fibrillation (AF) substantially increases an individual patient's risk for stroke while oftentimes remaining asymptomatic [3]. A diagnosis of AF can be reached using an ECG recording as short as 30 s and can reduce the burden of stroke by establishing oral anticoagulation. With the emergence of smartphones and smart watches being increasingly used worldwide, technicians and policy makers alike see the opportunity to increase the likelihood of early AF diagnosis and by this to reduce its sequelae [4–6].

The AliveCor Kardia (ACK) mobile ECG monitor was recently FDA-certified as a one-lead ECG recorder [7]. It is a portable device with two electrodes being able to write a single bipolar ECG lead utilizing a smartphone app. Its ease of use allows the ACK to be used for large-scale AF screening, especially outside of traditional health care settings. The app has been reported to use a Random Forest machine learning algorithm to differentiate sinus rhythm from AF and other rhythm disturbances [8].

The manufacturer-recommended lead I ACK recording may contain a critical amount of artifact, making an unequivocal rhythm diagnosis difficult. We therefore introduce an easily reproducible novel parasternal lead (NPL) by placing the ACK in a left parasternal position on a supine patient (Fig. 1). No data have been published systematically comparing different ACK recording vectors to a gold standard 12-lead ECG.

A recent study by William et al. [8] excluded patients with an "unclassified" diagnosis by the ACK app algorithm, which constituted about one in four patients, from a sensitivity and specificity analysis. Since "unclassified" recordings contain all heart rates above 100/min and all patients with ambiguous results, this may exclude the most

* Corresponding author.
E-mail address: felix.wegner@ukmuenster.de (F.K. Wegner).

https://doi.org/10.1016/j.ejim.2019.11.018
Received 11 July 2019; Received in revised form 24 November 2019; Accepted 26 November 2019
Available online 02 December 2019

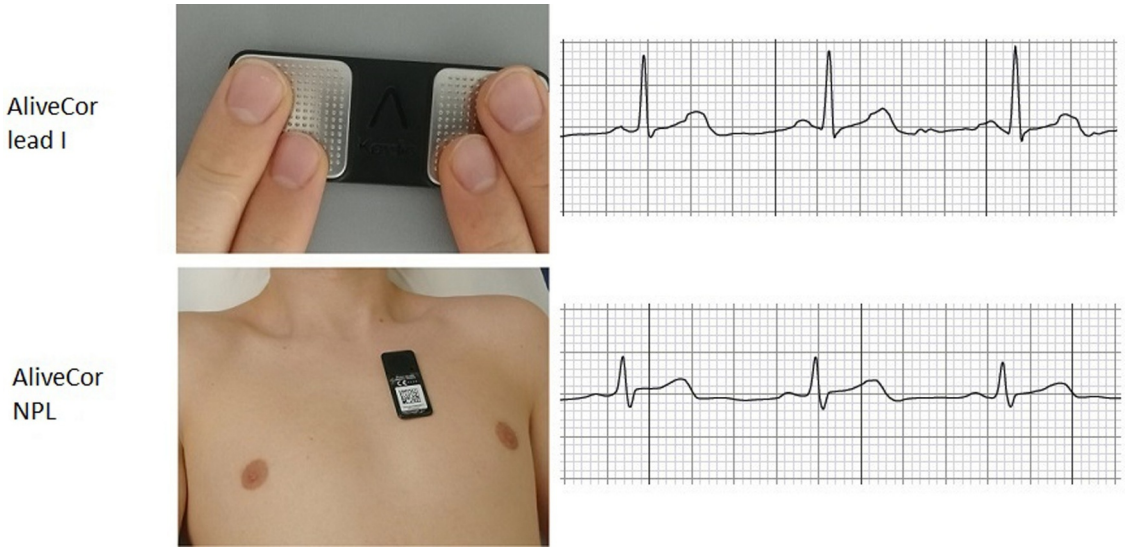

Fig.. 1. Device placement and representative ECG recordings in the manufacturer-recommended lead I and the novel parasternal lead (NPL) proposed by this study.

relevant portion of patients from AF detection. We therefore designed the present study to compare the diagnostic value of the ACK in a prospective patient cohort with no predefined exclusion criteria to elicit its use in everyday atrial fibrillation screening.

2. Methods

For the present investigator-initiated, unsponsored study we prospectively included 99 consecutive inpatients without further inclusion or exclusion criteria from our electrophysiology ward in a large tertiary-care university hospital. In each patient, a 12-lead ECG, an ACK lead I recording and a novel parasternal lead were recorded sequentially. The local ethics committee approved the study protocol and all patients gave informed consent.

2.1. ECG acquisition

Patients were instructed to record a lead I ACK ECG by placing two fingers of each hand on the electrodes of the ACK device. If a recording could not be obtained, different finger positions were allowed. The NPL was recorded by asking the patient to lie supine with an exposed chest and breathe normally. The ACK device was then placed in a left parasternal position so that a rhythm recording could be obtained (Fig. 1). The recordings were downloaded from the app and stored in the hospital IT system.

2.2. ECG analysis

All ECGs were blinded and evaluated separately by two experienced electrophysiologists. The electrophysiologists were asked to either classify the atrial rhythm as a) sinus rhythm or b) AF / atrial flutter. Additionally, the diagnoses of the Kardia smartphone app were recorded. These were a) sinus rhythm, b) possible AF, c) unclassified and d) no analysis. The diagnoses of the electrophysiologists using the standard 12-lead ECG were defined as the gold standard. For intrinsic validation, we first opted for interobserver agreement between the 12-lead ECG diagnoses of the two electrophysiologists. After that, we compared the individual smartphone diagnoses of the electrophysiologists and of the app algorithm to the 12-lead ECG for sensitivity and specificity. Additionally, the electrophysiologists were asked to rate the ACK recording as either adequate or inadequate in terms of the overall quality and interpretability (subjective signal-to-noise ratio etc.).

2.3. Statistical analysis

For database management and statistical analysis, we used IBM SPSS Version 25 (IBM Corporation, Somers, NY, USA). Departure from the mean was tested with Shapiro–Wilk's statistic. For a comparison of means, the student's *t*-test was used. Due to sample size, statistical analysis of binary variables was conducted using the Fisher exact test. Where applicable, statistical significance was defined as a two-sided *p*-value of <0.05. κ values were calculated using the Spearman rank-order correlation. A $\kappa > 0.80$ was defined as excellent agreement [9].

3. Results

38 of the 99 included patients were female. Mean age was 64 ± 15 years and mean BMI 27.8 ± 5.1 kg/m^2. The NPL was able to be recorded in all patients. Seven patients (7%) were excluded from the sensitivity and specificity analysis: In 6 patients, the 12-lead ECG was ambiguous (in all: continuous ventricular pacemaker stimulation with possible AF as atrial rhythm). In one further patient, there was disagreement as to the 12-lead ECG rhythm diagnosis (electrophysiologist 1: AF, electrophysiologist 2: sinus rhythm). The remaining 275 ECG of 92 patients were used for analysis. 65 patients exhibited sinus rhythm according to the 12-lead ECG. 22 patients had AF and five had atrial flutter, resulting in an arrhythmia prevalence in the study population of 29.4%.

3.1. Electrophysiologists' diagnosis

Analyzing the ACK lead I, the blinded electrophysiologists characterized 74 ECG as being of adequate quality and 18 ECG (20%) as being of inadequate quality (Fig. 2). Using the NPL, 20 ECG (22%) were described as being of inadequate quality. Of note, inadequate quality of an ACK lead I recording was not associated with inadequate quality of the corresponding NPL recording and vice-versa ($\kappa = 0.23$).

A cross-tabulation of the electrophysiologists' diagnoses based on the respective ACK recordings with the gold-standard 12-lead ECG diagnoses is shown in Table 1. This resulted in a sensitivity of the electrophysiologists for the ACK lead I of 100% (27 of 27 patients with AF/atrial flutter detected) and a specificity of 94% (61 of 65 patients with sinus rhythm correctly identified) with a κ of 0.90. Four patients were mistakenly identified as being in AF on ACK lead I analysis while displaying sinus rhythm in the 12-lead ECG. Of these patients, two were in sinus rhythm with continuous ventricular stimulation and two patients had a left bundle branch block (LBBB).

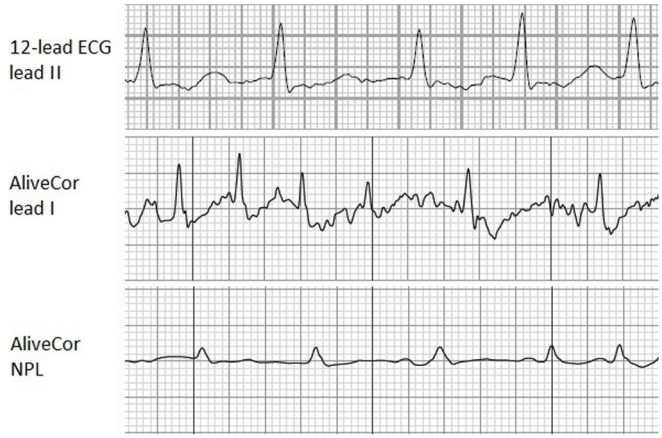

Fig.. 2. ECGrecordings of a patient with a critical amount of artifact in the AliveCor lead I showing no artifact in the novel parasternal lead (NPL).

Table 1
Electrophysiologists' diagnoses of the ACK ECG recordings compared to the gold-standard 12-lead ECG diagnosis. EP = electrophysiologists, AF = atrial fibrillation, Afl = atrial flutter.

Table 1a		12-lead ECG sinus rhythm (n = 65)	AF/Afl (n = 27)
EP diagnosis of **lead I**	sinus rhythm	61	0
	AF/Afl	4	27
Sensitivity	100%		
Specificity	94%		

Table 1b		12-lead ECG sinus rhythm (n = 65)	AF/Afl (n = 27)
EP diagnosis of **novel parasternal lead**	sinus rhythm	63	1
	AF/Afl	2	26
Sensitivity	96%		
Specificity	97%		

In the NPL, the electrophysiologists displayed a sensitivity of 96%, a specificity of 97% and a κ of 0.92. One patient with atrial flutter was misidentified as being in sinus rhythm. Two patients in sinus rhythm were described as being in AF. Of these, one was in sinus rhythm with continuous ventricular stimulation and one patient displayed a LBBB.

4. App algorithm diagnosis

A cross-tabulation of the app algorithm diagnoses of the respective ACK recordings with the gold-standard 12-lead ECG diagnoses is shown in Table 2. As mentioned above, a goal of the present study was to exclude as few patients as possible from a sensitivity and specificity analysis of the algorithm. We therefore calculated sensitivity and specificity for the app algorithm without excluding the groups of patients receiving a diagnosis of "unclassified" or "no analysis" by the algorithm.

The app algorithm thus had a sensitivity of 70% and a specificity of 69% in the lead I recordings with a κ of 0.53. In the NPL, the algorithm had a sensitivity of 55% and a specificity of 60% with a κ of 0.40. All of the patients with ECG recordings that the app algorithm defined as "normal" were indeed in sinus rhythm, thus leading to a negative predictive value of 100% in the present patient population for both the lead I and the NPL.

Table 2
App algorithm diagnoses of the ACK ECG recordings compared to the gold-standard 12-lead ECG diagnosis. AF = atrial fibrillation, Afl = atrial flutter.

Table 2a		12 lead ECG sinus rhythm (n = 65)	AF/Afl (n = 27)
App algorithm diagnosis of **lead I**	Sinus rhythm	45	0
	Possible AF	8	19
	unclassified	11	4
	no analysis	1	4

Table 2b		12-lead ECG sinus rhythm (n = 65)	AF/Afl (n = 27)
App algorithm diagnosis of **novel parasternal lead**	sinus rhythm	39	0
	Possible AF	3	15
	unclassified	21	7
	no analysis	2	5

4.1. Atrial flutter

Five patients exhibited atrial flutter at the time of ECG recording. The electrophysiologists displayed an accuracy of 100% (5 of 5 ECGs correctly interpreted) in detecting atrial flutter on an ACK lead I recording. In the NPL recording, the electrophysiologists showed a 90% accuracy: 4 of 5 ECGs were correctly interpreted, while there was disagreement in 1 recording (electrophysiologist 1: atrial flutter, electrophysiologist 2: sinus rhythm). The app algorithm diagnosed 4 of 10 ECG recordings with "possible AF" while diagnosing a further 4 of 10 ECGs with "unclassified". 2 ECGs received a connotation of "no analysis". The app algorithm therefore displayed a sensitivity of 40% in the diagnosis of atrial flutter.

4.2. Pacemaker stimulation

Out of 99 patients included in the present study, 12 patients had continuous ventricular pacemaker stimulation at the time of ECG recording. Six of these 12 patients were previously excluded from the statistical analysis because of the difficulty of diagnosing AF when regular ventricular pacemaker stimulation is present. In these 6 patients with AF and ventricular stimulation on 12-lead ECG, the electrophysiologists and the app algorithm misidentified all ACK ECG recordings as sinus rhythm. In the remaining 6 patients with sinus rhythm and ventricular stimulation, the electrophysiologists had an accuracy of 75% in the lead I recordings and of 83% in the NPL. The app algorithm had an accuracy of 50% in the lead I and an accuracy of 17% in the NPL.

4.3. The novel parasternal lead as a secondary recording vector

To evaluate a potential benefit of the addition of the NPL to the recording of a lead I ACK ECG, we decided to separately analyze the NPL recordings of those patients whose lead I ACK recordings were interpreted by the app algorithm as either "unclassified" or "no analysis". Table 3 displays the electrophysiologists' and the app algorithm diagnosis of the NPL recordings in these 20 patients.

The electrophysiologists' sensitivity in these NPL recordings was 100% and the specificity 92% with an accuracy of $\kappa = 0.90$. The app algorithm diagnoses displayed a sensitivity of 38% and a specificity of 17% with a κ of 0.20.

4.4. Predictors of correct diagnoses

To analyze for possible predictors of correct rhythm diagnoses, we

Table 3
Electrophysiologists' and app algorithm diagnoses of the novel parasternal lead recordings of 20 patients in which the app algorithm diagnosis of the lead I recording was either "unclassified" or "no diagnosis". EP = electrophysiologists, AF = atrial fibrillation, Afl = atrial flutter.

Table 3a		12-lead ECG sinus rhythm (n = 12)	AF/Afl(n = 8)
EP diagnosis of novel parasternal lead	Sinus rhythm	11	0
	Possible AF	1	8

Table 3b		12-lead ECG Sinus rhythm (n = 12)	AF/Afl (n = 8)
App algorithm diagnosis of novel parasternal lead	Sinus rhythm	2	0
	Possible AF	0	3
	Unclassified	9	1
	No analysis	1	4

compared the mean age, gender and other demographic variables. There were no statistically significant predictors identifiable (all $p > 0.05$).

5. Discussion

The present study is the first investigator-initiated, unsponsored report to analyze the diagnostic value of multiple recording vectors of the AliveCor Kardia ECG recorder in a large "all-comer" inpatient cohort. We were able to demonstrate that the ACK is able to record high quality one-lead ECG whose interpretation by experienced electrophysiologists can reliably determine underlying rhythm in both the manufacturer-recommended lead I and a novel parasternal lead.

We introduced the NPL by placing the ACK in a left parasternal position, as one in five of all lead I ACK recordings contained a critical amount of artifact. The novel parasternal lead was universally applicable, as a recording could be obtained in all included patients. Our results show that the NPL yields comparable results to the standard ACK vector. Furthermore, the NPL maintains its diagnostic accuracy when interpreted by an electrophysiologist in cases where the app algorithm diagnosis of the lead I is ambiguous (see Table 3a). A possible everyday clinical approach might thus include instructing a patient to 1) always record a lead I ACK ECG during symptomatic episodes, 2) record a NPL ECG when the immediately available app algorithm diagnosis of the lead I ECG recorded in step 1 shows a diagnosis of "unclassified" or "no analysis" and 3) making both recordings available to the treating physician for interpretation.

We detected a significantly lower sensitivity (55–70%) and specificity (60–69%) of the ACK app algorithm for the diagnosis of AF or flutter than previous working groups [8,10,11]. This can probably be largely attributed to the absence of exclusion criteria and our decision to include patients with "unclassified" and "no analysis" diagnoses by the algorithm. While William et al. reported a sensitivity of 97% and a specificity of 94% for the app algorithm in 52 patients, this was mainly due to the authors' decision to exclude 28% of all ACK recordings, which were annotated as "unclassified" [8]. In our opinion, excluding a large amount of patients who use a population-based screening tool misses its target. In addition, other previous studies applying the ACK to larger patient populations [10,11] did not systematically compare the recordings to the gold-standard 12-lead ECG, thereby possibly distorting their analyses.

In contrast to previous publications, our data systematically comparing the ACK to the 12-lead ECG indicate that the main utility of the app algorithm itself is its high negative predictive value. In addition, when the app algorithm displays a diagnosis of "unclassified" or "no analysis", further analysis by an experienced medical professional should be sought.

The analysis by electrophysiologists had a very high sensitivity (96–100%) and specificity (94–97%) for the presence of AF or atrial flutter with an excellent agreement with the respective 12-lead ECG diagnoses. In contrast to the ACK app algorithm, the electrophysiologists were able to detect atrial flutter and ventricular stimulation with sinus rhythm with a high accuracy. No patient-specific predictors of inaccurate diagnoses or a large amount of artifact could be found.

6. Limitations

The present study was conducted as a single center study with its inherent limitations. Although patients were recruited from an inpatient unit at a large tertiary center, the prevalence of AF recordings remained relatively low. However, we believe that a prevalence of arrhythmias of 29% represent a sufficient number to draw conclusions from with regard to diagnostic accuracy. Apart from that, our patient group is the largest with a systematic and blinded ECG analysis of ACK recordings compared to the gold-standard 12-lead ECG.

7. Conclusion

The AliveCor Kardia is a smartphone-based one-lead ECG recorder allowing for an accurate rhythm diagnosis by an interpreting cardiologist both in the manufacturer-recommended lead I and a novel parasternal lead. The diagnostic app algorithm has a very high negative predictive value. In cases where the app does not show an interpretation of "normal rhythm", the underlying recordings should be interpreted by a cardiologist.

Decleration of Competing Interest

None.

References

[1] Lin HJ, Wolf PA, Kelly-Hayes M, et al. Stroke severity in atrial fibrillation. The Framingham Study. Stroke 1996;27(10):1760–4. PubMed PMID: 8841325.

[2] Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the global burden of disease study 2010. Lancet 2012;380(9859):2095–128. https://doi.org/10.1016/S0140-6736(12)61728-0. PubMed PMID:23245604.

[3] Kirchhof P, Benussi S, Kotecha D, et al. ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016;37(38):2893–962. https://doi.org/10.1093/eurheartj/ehw210. PubMed PMID:27567408.

[4] Turakhia MP, Desai M, Hedlin H, et al. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: the Apple Heart Study. Am Heart J 2019;207:66–75. https://doi.org/10.1016/j.ahj.2018.09.002. PubMed PMID:30392584.

[5] Freedman B, Camm J, Calkins H, et al. Screening for atrial fibrillation: a report of the AF-screen international collaboration. Circulation 2017;135(19):1851–67. https://doi.org/10.1161/CIRCULATIONAHA.116.026693. PubMed PMID:28483832.

[6] Mairesse GH, Moran P, Van Gelder IC, 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 Estimulacion Cardiaca y Electrofisiologia (SOLAECE). Europace 2017;19(10):1589–623. https://doi.org/10.1093/europace/eux177. PubMed PMID:29048522.

[7] US Food and Drug Administration. 510(k) Premarket Notification 2014 [updated 29.04.201906.05.2019]. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K140933.

[8] William AD, Kanbour M, Callahan T, et al. Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: the iREAD study. Heart Rhythm 2018;15(10):1561–5. https://doi.org/10.1016/j.hrthm.2018.06.037. PubMed PMID:30143448.

[9] Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37(5):360–3. PubMed PMID:15883903.

[10] Lowres N, Neubeck L, Salkeld G, et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies. The SEARCH-AF study. Thromb Haemost 2014;111(6):1167–76. https://doi.org/10.1160/TH14-03-0231. PubMed PMID:24687081.

[11] Chan PH, Wong CK, Pun L, et al. Head-to-Head comparison of the AliveCor heart monitor and microlife watchbp office AFIB for atrial fibrillation screening in a primary care setting. Circulation 2017;135(1):110–2. https://doi.org/10.1161/CIRCULATIONAHA.116.024439. PubMed PMID:28028066.