Family Practice, 2021, 70–75
doi:10.1093/fampra/cmaa076
Advance Access publication 7 August 2020

OXFORD

Health Service Research

Evaluation of general practitioners' single-lead electrocardiogram interpretation skills: a case-vignette study

Evert P.M. Karregat*, Jelle C.L. Himmelreich, Wim A.M. Lucassen, Wim B. Busschers, Henk C.P.M. van Weert and Ralf E. Harskamp

Amsterdam Public Health & Amsterdam Cardiovascular Sciences Research Institute, Department of General Practice, University of Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands

*Correspondence to E.P.M. Karregat, Department of General Practice, Amsterdam UMC, Location AMC, Meibergdreef 9, Room J2-126, 1105 AZ, Amsterdam, The Netherlands; E-mail: e.p.karregat@amsterdamumc.nl

Abstract

Background: Handheld single-lead electrocardiograms (1L-ECG) present a welcome addition to the diagnostic arsenal of general practitioners (GPs). However, little is known about GPs' 1L-ECG interpretation skills, and thus its reliability in real-world practice.

Objective: To determine the diagnostic accuracy of GPs in diagnosing atrial fibrillation or flutter (AF/Afl) based on 1L-ECGs, with and without the aid of automatic algorithm interpretation, as well as other relevant ECG abnormalities.

Methods: We invited 2239 Dutch GPs for an online case-vignette study. GPs were asked to interpret four 1L-ECGs, randomly drawn from a pool of 80 case-vignettes. These vignettes were obtained from a primary care study that used smartphone-operated 1L-ECG recordings using the AliveCor KardiaMobile. Interpretation of all 1L-ECGs by a panel of cardiologists was used as reference standard.

Results: A total of 457 (20.4%) GPs responded and interpreted a total of 1613 1L-ECGs. Sensitivity and specificity for AF/Afl (prevalence 13%) were 92.5% (95% CI: 82.5–97.0%) and 89.8% (95% CI: 85.5–92.9%), respectively. PPV and NPV for AF/Afl were 45.7% (95% CI: 22.4–70.9%) and 98.8% (95% CI: 97.1–99.5%), respectively. GP interpretation skills did not improve in case-vignettes where the outcome of automatic AF-detection algorithm was provided. In detecting any relevant ECG abnormality (prevalence 22%), sensitivity, specificity, PPV and NPV were 96.3% (95% CI: 92.8–98.2%), 68.8% (95% CI: 62.4–74.6%), 43.9% (95% CI: 27.7–61.5%) and 97.9% (95% CI: 94.9–99.1%), respectively.

Conclusions: GPs can safely rule out cardiac arrhythmias with 1L-ECGs. However, whenever an abnormality is suspected, confirmation by an expert-reader is warranted.

Key words: Arrhythmias, cardiac, atrial fibrillation, computers, handheld, electrocardiography, mobile applications, primary health care

Introduction

Patients frequently consult their general practitioner (GP) with symptoms that may indicate an underlying cardiac arrhythmia (1,2). When a cardiac arrhythmia is suspected, a 12-lead electrocardiogram (12L-ECG) is indicated (2–4). Due to logistical challenges, only a third of ECGs is performed while a patient is experiencing symptoms (5). Obtaining an ECG during symptoms is important to diagnose

Key Messages

- Using 1L-ECGs, GPs can safely rule out cardiac arrhythmias, including AF.
- When GPs suspect an ECG abnormality, they are incorrect in half of the cases.
- GPs should reconsider their diagnosis when different from the algorithm.
- Expert-confirmation is warranted whenever GPs suspect a 1L-ECG abnormality.

an arrhythmia, but also to reassure patients when a normal ECG is found during symptoms (5).

New 'point-of-care' handheld single-lead ECG (1L-ECG) devices overcome some of the logistical challenges of performing an ECG in primary care and therefore likely improve diagnostic gain. Despite their great potential, diagnostic accuracy and implementation of these devices in primary care, has not been well investigated (6). Moreover, little is known about the added value of 1L-ECG devices' automatic atrial fibrillation (AF) detection algorithm to a GP's visual assessment of the 1L-ECG recording. This is important as a number of agencies have already endorsed the use of these devices (7).

Therefore, we evaluated the 1L-ECG interpretation skills of GPs with regards to AF, atrial flutter (Afl) and other relevant ECG abnormalities. Secondly, we assessed whether the availability of the automatic algorithm interpretation for AF altered GPs accuracy of rhythm assessment. Furthermore, we assessed GP interpretation skills for other clinically relevant ECG abnormalities.

Methods

We reported this study in accordance with the Standards for Reporting Diagnostic Accuracy (STARD) 2015 statement (see Supplementary Appendix) (8).

Design and setting

This is an online study using case vignettes among GPs in order to assess their ability to correctly interpret 1L-ECGs for AF and/or Afl (AF/Afl) and other relevant ECG abnormalities.

Participants

We contacted GPs by e-mail through their affiliation with one of four participating universities in The Netherlands (Amsterdam UMC, location AMC; Erasmus MC, Rotterdam; UMC Groningen; Maastricht UMC). We sent non-responding GPs reminders after 2 and 4 weeks, providing a non-responder questionnaire asking for a short rationale why they had not participated. Invited GPs received access to an online secured electronic survey (Limesurvey, Amsterdam UMC, The Netherlands), using an physician-personalized link to prevent both multiple use of the invitation and unnecessary e-mail reminders to responders. After the 6 week study period, we removed all traceable information, making the survey completely anonymous. Before ECG assessment, we obtained a self-assessment of participants' ECG knowledge (on a 1 to 10 scale).

Collection of 1L-ECGs

We used 1L-ECGs previously collected as part of the VESTA study (9). The VESTA study validated the KardiaMobile 1L-ECG (AliveCor, Mountain View, CA) in consecutive primary care patients who were assigned to standard 12L-ECG for any non-acute indication as ordered by the local GP. Each patient's indication for undergoing 12L-ECG in the VESTA study was classified as either symptom driven (for recent non-acute cardiac symptoms such as palpitations) or protocol driven (as part of a protocolized care programme—i.e. primary or secondary prevention of cardiovascular disease, at the discretion of the GP).

The KardiaMobile is a smartphone-operated device, able to easily record 1L-ECGs (10,11). The device consists of two small electrodes which a patient holds in both hands for 30 s. The smartphone application enables a direct visual interpretation of the 1L-ECG signal. In addition, the application has an in-built computer algorithm for automatic detection of AF.

Selection of ECG vignettes and batches

We selected 80 cases from the VESTA for our case-vignette study. Table 1 shows the prevalence of ECG abnormalities in both the overall VESTA cohort and the 80 cases selected for the current analysis (9). We enclosed a case vignette for each 1L-ECG consisting of patient characteristics and indication for performing the ECG. We limited the number of ECG interpretation assignments to four per GP. We therefore created 20 'batches' with four ECGs each. In these batches, we aimed for an even distribution of symptom and protocol-driven ECGs. Also, half of the ECGs in each batch included an automatic algorithm interpretation for suspicion of AF. The 1L-ECGs and GPs were randomly assigned to these batches using an online randomization tool (12).

Study procedure

Figure 1 depicts the flow chart of the study procedure. Participating GPs viewed the ECGs and reported their findings in the online survey. The following options were offered, of which GPs could select one or more per ECG: (1) normal ECG; (2) atrial fibrillation or flutter; (3) small complex tachycardia (excl. sinus tachycardia); (4) bradycardia (<45 bpm); (5) high-grade AV block; (6) bundle branch block; (7) other findings (e.g. premature atrial complex, inversed T wave, etc.) providing an open text field to answer.

Finally, we asked GPs whether they assessed these 1L-ECG device rhythm strips as a useful tool for their daily practice.

Outcomes of interest

Primary outcome measures were sensitivity, specificity, positive and negative predictive values (PPV and NPV, respectively) of the 1L-ECG as assessed by GPs in detection of (i) AF/Afl, both with and without the presence of the automatic algorithm interpretation for AF and (ii) 'any relevant 1L-ECG abnormality'. The latter was defined as AF/Afl and/or bradycardia (<45/min), high-grade AV-block, bundle branch block, small complex tachycardia, repolarisation-disorders and other relevant findings (e.g. suspicion of WPW syndrome or prolonged QT time).

Other outcomes of interest included GPs' diagnostic performance stratified by ECG indication (protocol versus symptom driven) as well as the previously mentioned relevant ECG abnormalities. Furthermore, we performed analyses based upon the various GP characteristics (e.g. experience level, self-rated ECG knowledge, etc.). Finally, we analysed GPs' assessment of the 1L-ECG as a viable tool in their daily practice.

Table 1. Prevalence of 1L-ECG abnormalities in this online case vignette versus natural prevalence of 1L-ECG abnormalities in primary care (VESTA)

Abnormalities	Current case-vignette study			VESTA study natural prevalence[c]
	Symptom driven (n = 38)	Protocol driven (n = 42)	Total percentage (n = 80)[b]	
AF/Afl[a]	7	3	12.5%	10.7%
BBB[a]	0	5	6.3%	6.1%
SVT[a]	0	1[d]	1.3%	1.4%
Negative T[a]	0	1	1.3%	0.5%
PVC	3	5	10.0%	10.7%
PAC	0	2	2.5%	4.2%

Prevalence of rhythm abnormalities in symptom- and protocol-driven cases are shown as numbers.
[a]Variable included in the outcome 'any relevant 1L-ECG abnormality'.
[b]Some ECGs contained multiple abnormalities.
[c]Based upon cardiologists' interpretation of 1L-ECG.
[d]Regrettably the wrong case vignette was provided making this a protocol-driven (asymptomatic) case.

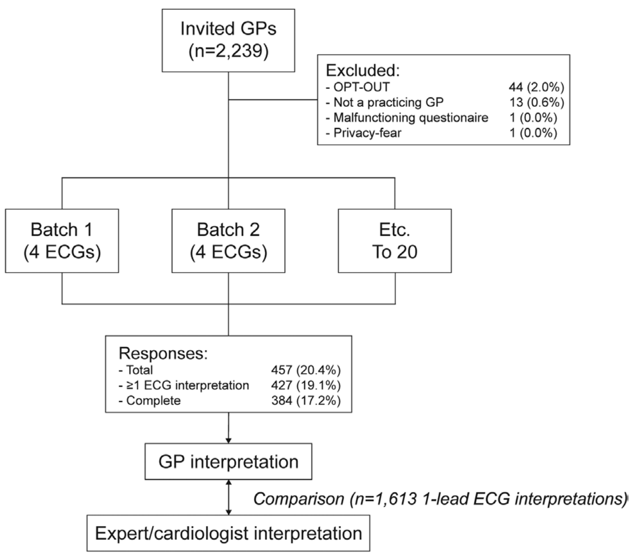

Figure 1. Flow chart of online case-vignette study design among GPs, evaluating their 1L-ECG interpretation skills.

Reference standard

Two independent cardiologists interpreted all 1L-ECGs, with a third cardiologist as a referee in case of disagreement (9). No clinical information was available to these assessors.

Sample size

A formal sample size calculation was not deemed suitable both given the paucity on data to make accurate predictions on GPs' diagnostic accuracy and due to the used procedure with fixed batches and randomization at beforehand, whereby data are not independent (13). Therefore, we chose to pragmatically approach all GPs of whom e-mail addresses were available.

Statistical methods

Summary statistics for discrete variables are presented as numbers and percentages. We show non-normally distributed numerical data as median with IQR.

We used logistic mixed models to estimate sensitivity and specificity where random effects for potential dependencies at the level of the responding GP, batch and ECG level were included in the model. Based on Akaike's Information Criterion the most parsimonious model with respect to the random structure was used for estimation of the characteristics of interest.

We estimated the PPV and NPV by repeated sampling from the 80 ECGs. Here we used logistic models and bootstrapping techniques to also obtain 95% confidence intervals. All results were based on 500 runs.

We assessed statistical significance at the 0.05 level in all analyses. We performed statistical analyses using R statistical software version 3.5.1 (14), with the lme4 package version 1.1–19 (15).

Results

A total of 457 GPs (20.4%) responded to our invitation. The responders subsequently interpreted 1613 ECGs, as depicted in Figure 1. Participants generally reflected the characteristics of GPs in The Netherlands, with a median age of 47 and median 14 years of clinical experience as a GP (Table 2). However, in our sample women (43% versus 51%) as well as those with a part time and unsteady (for hire) job situation (5% versus 17%) were underrepresented compared with the national average (16). Self-rated ECG knowledge scored a median of 6, with a quarter of participants reporting to never interpret ECGs (27%).

Accuracy of GP interpretation of 1L-ECGs

Prevalence of AF/Afl and 'any relevant 1L-ECG abnormality' was 13% and 22%, respectively. For our primary outcome, diagnostic accuracy of GPs in detecting AF/Afl, we found a sensitivity of 92.5% and specificity of 89.8% with a PPV and NPV of 45.7% and 98.8%, respectively (Table 3).

Sensitivity and specificity of the automatic AF detection algorithm were 100% when the algorithm outcomes were dichotomized into 'possible AF' versus a clustering of all other algorithm outcomes. However, GP performance in detecting AF/Afl did not improve when the outcome of the automatic AF algorithm was provided in the case vignette.

Sensitivity, specificity, PPV and NPV of GPs in detecting 'any relevant 1L-ECG abnormality' were 96.3%, 68.8%, 43.9% and 97.9%, respectively.

Sensitivity and specificity for AF/Afl showed similar results for symptom and protocol-driven ECGs while the PPV was higher in symptom-driven cases (Supplementary Table 1). Sensitivity was higher in symptom-driven cases compared with protocol-driven cases for 'any relevant 1L-ECG abnormality'. Supplementary Table 2 shows the stratified analysis on GP-specific features and their association on GPs' performance. In sum, age, experience years as a GP and the factor 'doubt' were statistically associated with correct 1L-ECG interpretation by GPs. Furthermore, incomplete responders had a tendency to perform worse compared with complete responders. GPs' self-rated ECG knowledge and the number of ECGs a GP interprets per month were not significantly associated with 1L-ECG interpretation.

Supplementary Table 3 displays the diagnostic accuracy analysis on the specified outcomes SVT, BBB and repolarization disorders. All showed high specificity and low sensitivity.

Table 2. Characteristics of participating GPs in this online case-vignette study evaluating GPs' 1L-ECG interpretation skills

	Number or median	Percentage or IQR
GP characteristics (n = 384)		
Sex (female)	165	43.0%
Age (years)	47	39.00–56.00
Experience as a GP (years)	14.0	6.75–22.00
ECG knowledge (self-rated)[a]		
Complete responders (n = 384)	6.0	5.00–7.00
Incomplete responders (n = 73)	6.0	4.00–7.00
ECG interpretations per month (n = 384)		
Null	102	26.6%
1 to 5	198	51.6%
6 to 10	65	16.9%
More than 10	19	4.9%
Practice adherence area (n = 383)		
Rural	93	24.3%
Urbanized	166	43.3%
Strongly urbanized	121	31.6%
Diverse	3	0.8%
Type of employment (n = 384)		
Holder of own GP practice	292	76.0%
GP with permanent basis	74	19.3%
GP without permanent basis	18	4.7%

[a]Scale 1–10; difference between complete and incomplete responders: P = 0.049.

Clinical relevance according to participants

Of those GPs who responded to the question on clinical relevance (n = 384), 226 (59%) were positive about a 1L-ECG point-of-care tool for use in clinical practice. Most responders found it useful in patients presenting with palpitations (96%) and screening for AF (84%) (Table 4).

Non-responders

In total, 229 non-responders answered the non-responder questionnaire. Most non-responders declined participation for reasons not related to the questionnaire. A smaller group did not participate because they did not feel affinity with reading ECGs (18%) or felt their ECG reading skills were not sufficient (17%) (Supplementary Table 4).

Discussion

GPs were able to safely exclude AF/Afl and other relevant ECG abnormalities on a 1L-ECG. When GPs suspected AF/Afl and other ECG abnormalities; however, their assessment was inaccurate in half of such cases. Providing GPs with the automatic algorithm interpretation for AF did not improve their performance.

Previous work

Two studies investigated 1L-ECG interpretation by GPs for AF. In 2007, Mant *et al.* assessed diagnostic accuracy for AF detection in GPs when using merely a single chest lead. They found a sensitivity and specificity of 84.8% and 86.4%, respectively. Interpretation of limb-lead or 12L-ECGs by GPs gave similar results, with a PPV of 40.9% to diagnose AF (17). We found a higher sensitivity and specificity for AF, which may be due to the longer duration of the KardiaMobile rhythm-strip compared with the single chest-lead and our presentation in a case-vignette format. A comparably low PPV was found. More recently, Koshy *et al.* found a diagnostic accuracy of 85% for two GPs interpreting 408 KardiaMobile 1L-ECGs for the presence of AF/Afl (18).

Furthermore, Mant *et al.* showed an increase in sensitivity when combining GP interpretation with automatic algorithm software on 12L-ECGs (17), contrasting with our findings. However, they regarded AF as present when either of them was positive, whilst in our study it was only the GP who decided whether AF was present

Table 3. Diagnostic accuracy of GPs' interpretation of 1L-ECGs

Diagnostic accuracy	AF/Afl								Any relevant abnormality	
	Total		Without algorithm*		With algorithm*		[a]Odds ratio	95% CI		
Prevalence	210/1613 (13%)		84/775 (11%)		126/838 (15%)				352/1613 (22%)	
Sensitivity	92.5	(82.5–97.0)	91.2	(70.2–97.8)	93.4	(80.1–98.0)	1.372	(0.201–9.385)	96.3	(92.8–98.2)
Specificity	89.8	(85.5–92.9)	90.4	(94.2–84.5)	89.2	(83.2–93.2)	0.877	(0.430–1.787)	68.8	(62.4–74.6)
PPV	45.7	(22.4–70.9)	42.7	(14.0–77.4)	50.4	(19.7–80.7)			43.9	(27.7–61.5)
NPV	98.8	(97.1–99.5)	99.4	(99.0–99.7)	99.1	(98.1–99.6)			97.9	(94.9–99.1)

Accuracy measures are presented as percentages with their corresponding 95% CI.
[a]The odds ratio measures the association between availability of the algorithm and the probability of a GP interpreting a 1L-ECG correctly.

Table 4. GPs' views regarding an added value of 1L-ECGs for GPs' clinical practice

Added value (*n* = 384)	Number	Percentage
Yes	226	58.9
No	58	15.1
I don't know	100	26.0
If yes, for what indication?[a]		
Palpitations	217	96.0
Screening for AF	189	83.6
Collapse	105	46.5
Dizziness	85	37.6
Dyspnoea	68	30.1
Chest pain	37	16.4

[a]More than one answer possible.

or not, taking into account the automatic algorithm interpretation whenever available.

Strengths and limitations

A number of strengths deserve to be mentioned. Firstly, we invited as many GPs as possible to try and minimize individual GP effects on outcomes. By using e-mail recruitment, we were able to easily distribute our questionnaire to a large number of GPs. Secondly, the use of case vignettes combined with a case-mix that resembles real-life prevalence of ECG abnormalities contributed to the generalizability of our results. Thirdly, through stratified analysis we were able to show that the availability of the automatic algorithm interpretation did not contribute to a more accurate assessment for AF by GPs. Finally, we were able to perform analysis stratified by indication.

There were a number of limitations in our study. Firstly, selection bias may have been introduced by both our participant selection, all being affiliated with a university medical centre, and the suboptimal response rate. Furthermore, the diagnostic accuracy of incomplete responders tended to be lower compared with complete responders. This may have had a positive effect on our outcomes. Secondly, we presented the 1L-ECGs rhythm strips to the GPs as 30-s overview files. This may differ from the user experience in a smartphone app where only snippets of a few seconds are shown, and one has to 'swipe' through the rest of the recording. Such a swipe functionality was technically impossible to implement in our questionnaire software. In clinical practice, however, GPs are able to compute an overview file as PDF. Thirdly, we forced respondents to choose from a select number of ECG abnormalities. However, we did give GPs a free text box to enter additional information. Because of this, after study completion, we recoded the open text fields into variables for 'repolarization disorders', 'other relevant findings' and 'doubt'. Finally, since 'any relevant 1L-ECG abnormality' is a composite dichotomous outcome, it is possible that GPs and cardiologists have judged a particular 1L-ECG strip as abnormal for different reasons. In such cases, the answer would have been counted as correct, while the underlying interpretation was incorrect.

Clinical relevance

GPs are often confronted with symptoms that may be due to cardiac arrhythmias (1,2). Previously, a 1L-ECG proved a reliable tool for rhythm assessment in primary care when interpreted by cardiologists (9). However, in primary care, it is the GP who is the first to interpret the 1L-ECG strip. Based on the GP's judgement of the 1L-ECG a treatment is started or withheld, or additional investigations are performed. Therefore, reliable interpretation of 1L-ECGs by GPs is a prerequisite for safe and effective implementation of these devices. However, to date, this has not yet been investigated extensively (17,18).

This diagnostic accuracy study shows that GPs can safely rule out AF/Afl and other clinically relevant ECG abnormalities in 1L-ECGs representative of a primary care population. However, the low PPV for AF/Afl as well as for other relevant ECG abnormalities suggests that caution is warranted whenever a GP suspects an abnormality. For these cases, we suggest to either consult a cardiologist for final diagnosis or to perform additional cardiologic investigations.

Furthermore, we showed that providing GPs with an automatic algorithm interpretation for AF does not improve results, despite the algorithm's 100% diagnostic accuracy for AF/Afl in this pool of ECGs. Obviously, GPs did not follow the results of the provided algorithm. It is unclear what factors contribute to this finding. However, it does underscore that GPs should reconsider their diagnosis when their interpretation differs from the algorithm.

In clinical practice, a GP should always take into account all patient information (i.e. symptoms, previous heart diseases, etc.) whenever issuing and interpreting an ECG. Different clinical circumstances give different a priori chances of abnormalities. In our study, stratified analysis for ECG indication showed that the PPV in symptom-driven ECGs was higher for AF/Afl compared with the protocol-driven cases.

The number of ECG-interpretations per month and self-rated ECG knowledge of a GP were not statistically associated with better performance. This implies that our recommendations apply to both GPs with experience in ECG interpretations as well as GPs feeling less comfortable in reading ECGs.

We think that the use of a 1L-ECG device connected with a smartphone application can be a useful tool for GPs in daily practice, with the internal algorithm offering a stronghold for its interpretation. Furthermore, the connected smartphone application offers the opportunity to both easily save the ECG in a patients' record and instantly share the ECG with for example a cardiologist. GPs should, however, be cautious with the generated data in order to protect the privacy of their patients.

Future work

Studies are required not only to assess whether the use of a 1L-ECG alters GPs' medical management but also to address physician and patient satisfaction using a 1L-ECG. Additionally, the learning effect when using a 1L-ECG and its automatic algorithm for AF more often should be subject to further study. Finally, post-implementation studies should clarify which GPs will use these devices, for which patients and with what results.

Conclusion

GPs can safely rule out cardiac arrhythmias, including AF, using single-lead ECGs. However, when an ECG abnormality is suspected, the GP is incorrect in half of the cases. An automatic ECG interpretation algorithm for AF did not improve GPs' diagnostic accuracy. As such, whenever the GP or the algorithm suspects an abnormality, we recommend a low threshold for consulting an ECG expert for confirmation of this abnormality.

Supplementary material

Supplementary material is available at *Family Practice* online.

Acknowledgements

We thank the general practitioners and the personnel of the participating Primary Care Academic Networks for their support and effort. In alphabetical order, these were the Academic Medical Centre in Amsterdam, Erasmus Medical Centre in Rotterdam, Maastricht University Medical Centre and University Medical Centre Groningen. Finally, we would like to express thanks to the team of LimeSurvey software and Marleen Inge, the Data Protection Officer of the Academic Medical Centre, for the insightful comments in the design phase of the study.

Declarations

Ethical approval: Since this was an unobtrusive, anonymous questionnaire, ethical approval was not required.

Funding: This manuscript was supported by university funds. REH received salary support through a Rubicon grant from the Netherlands Organization of Scientific Research (NWO). The authors report no ties to the manufacturer of the investigated device and had full autonomy in design, conduct and reporting of the manuscript.

Conflict of interest: None.

Prior presentation(s): Not applicable.

References

1. Zwietering P, Knottnerus A, Gorgels T, Rinkens P. Occurrence of arrhythmias in general practice. *Scand J Prim Health Care* 1996; **14**(4): 244–50.
2. Raviele A, Giada F, Bergfeldt L *et al.*; European Heart Rhythm Association. Management of patients with palpitations: a position paper from the European Heart Rhythm Association. *Europace* 2011; **13**(7): 920–34.
3. NHG-werkgroep Atriumfibrilleren. NHG-Standaard Atriumfibrilleren. *Huisarts & Wetenschap.* 2017; **60**(9): 460–70.
4. Zimetbaum P, Josephson ME. Evaluation of patients with palpitations. *N Engl J Med* 1998; **338**(19): 1369–73.
5. Zwietering PJ, Knottnerus JA, Rinkens PE, Kleijne MA, Gorgels AP. Arrhythmias in general practice: diagnostic value of patient characteristics, medical history and symptoms. *Fam Pract* 1998; **15**(4): 343–53.
6. Bansal A, Joshi R. Portable out-of-hospital electrocardiography: a review of current technologies. *J Arrhythm* 2018; **34**(2): 129–38.
7. *Economic Impact Evaluation Case Study: AliveCor Kardia Mobile*. Heslington, York, UK: York Health Economics Consortium; February 2018.
8. Bossuyt PM, Reitsma JB, Bruns DE, *et al.* STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Br Med J* 2015; **351**:h5527.
9. Himmelreich JCL, Karregat EPM, Lucassen WAM *et al.* Diagnostic accuracy of a smartphone-operated, single-lead electrocardiography device for detection of rhythm and conduction abnormalities in primary care. *Ann Fam Med* 2019; **17**(5): 403–11.
10. Haberman ZC, Jahn RT, Bose R *et al.* Wireless smartphone ECG enables large-scale screening in diverse populations. *J Cardiovasc Electrophysiol* 2015; **26**(5): 520–6.
11. Desteghe L, Raymaekers Z, Lutin M *et al.* Performance of handheld electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. *Europace* 2017; **19**(1): 29–39.
12. Haahr M. *RANDOM.ORG: True Random Number Service* https://www.random.org (accessed on 10 July 2020).
13. Hayat MJ, Hedlin H. Modern statistical modeling approaches for analyzing repeated-measures data. *Nurs Res* 2012; **61**(3): 188–94.
14. R Core Team. *R: A Language and Environment for Statistical Computing [computer program]*. Vienna, Austria: R Foundation for Statistical Computing; 2018.
15. Bates DMM, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw* 2015; **67**(1): 1–48.
16. Van der Velden LF, Kasteleijn A, Kenens RJ. *Cijfers uit de registratie van huisartsen: peiling 2016*. Utrecht: NIVEL; 2017.
17. Mant J, Fitzmaurice DA, Hobbs FD *et al.* Accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: analysis of data from screening for atrial fibrillation in the elderly (SAFE) trial. *Br Med J* 2007; **335**(7616): 380.
18. Koshy AN, Sajeev JK, Negishi K *et al.* Accuracy of blinded clinician interpretation of single-lead smartphone electrocardiograms and a proposed clinical workflow. *Am Heart J* 2018; **205**: 149–53.