ESC European Society of Cardiology

Europace (2017) **19**, 1449–1453
doi:10.1093/europace/euw286

Safe automatic one-lead electrocardiogram analysis in screening for atrial fibrillation

Emma Svennberg[1]*, Martin Stridh[2], Johan Engdahl[3], Faris Al-Khalili[1], Leif Friberg[1], Viveka Frykman[1], and Mårten Rosenqvist[1]

[1]Department of Clinical Sciences, Cardiology Unit, Karolinska Institutet, Danderyd University Hospital, Stockholm SE-182 88, Sweden; [2]Department of Biomedical Engineering, Lund University, Lund and Cardiolund Research AB, Lund, Sweden; and [3]Department of Medicine, Halland Hospital, Halmstad, Sweden

Received 17 June 2016; accepted after revision 13 August 2016; online publish-ahead-of-print 6 October 2016

Aims Screening for atrial fibrillation (AF) using intermittent electrocardiogram (ECG) recordings can identify individuals at risk of AF-related morbidity in particular stroke. We aimed to validate the performance of an AF screening algorithm compared with manual ECG analysis by specially trained nurses and physicians (gold standard) in 30 s intermittent one-lead ECG recordings.

Methods and results The STROKESTOP study is a mass-screening study for AF using intermittent ECG recordings. All individuals in the study without known AF registered a 30-s ECG recording in Lead I two times daily for 2 weeks, and all ECGs were manually interpreted. A computerized algorithm was used to analyse 80 149 ECG recordings in 3209 individuals. The computerized algorithm annotated 87.1% (n = 69 789) of the recordings as sinus rhythm/minor rhythm disturbances. The manual interpretation (gold standard) was that 69 758 ECGs were normal, making the negative predictive value of the algorithm 99.9%. The number of ECGs requiring manual interpretation in order to find one pathological ECG was reduced from 288 to 35. Atrial fibrillation was diagnosed in 84 patients by manual interpretation, in all of whom the algorithm indicated pathology. On an ECG level, 278 ECGs were manually interpreted as AF, and of these the algorithm annotated 272 ECGs as pathological (sensitivity 97.8%).

Conclusion Automatic ECG screening using a computerized algorithm safely identifies normal ECGs in Lead I and reduces the need for manual evaluation of individual ECGs with >85% with 100% sensitivity on an individual basis.

Keywords Atrial fibrillation • Algorithms • Electrocardiography • Mass screening • Stroke

Introduction

Atrial fibrillation (AF) is the most common clinical arrhythmia affecting at least 3% of the adult population,[1] and as the population ages it is estimated to double in prevalence until 2060. The risk of stroke in patients with AF is increased five-fold, and due to the stroke risk the European Society of Cardiology recommend opportunistic screening by pulse for AF in individuals aged 65 and above.[2]

By controlling the pulse, with subsequent ECG if the pulse was irregular, the detection of AF was shown to be 1.6%.[3] In a meta-analyses of prior 12-lead ECG screening for AF, 1.4% of individuals over the age of 65 were found to have newly detected AF.[4] In order to assess if intermittent ECG screening would be a more efficient way to find AF, and to reduce future co-morbidities, the STROKESTOP study was initiated. The STROKESTOP study is a large prospective randomized screening study in an elderly population.[5] Totally, 7173 participants were screened for AF with a one-lead, handheld ECG recorder intermittently during a fortnight. Of the screened population, 3% were found to have new AF, and of those >90% initiated oral anticoagulant (OAC) treatment. The screening generated ~190 000 ECGs. Specially trained registered nurses manually sorted ECGs into two categories—normal ECGs and abnormal ECGs. All abnormal ECG recordings were sent to a physician for interpretation. A large proportion of the study costs was driven by the manual interpretation of normal ECGs by the nurses. A mathematical algorithm that could safely identify normal ECGs could limit the amount of manual work, reduce the cost of AF screening, and hence facilitate population-based AF screening.

* Corresponding author. Tel: +46 8 739584822. *E-mail address*: emma.svennberg@ds.se

What's new?

- Systematic screening for atrial fibrillation (AF) using intermittent 1-lead electrocardiograms (ECGs) is a powerful way to detect AF, but is cumbersome as it generates a vast number of ECGs that require manual interpretation.
- A computerized algorithm can be used to rule out pathology in >85% of ECGs, reducing the need for manual interpretation eight-fold.
- The computerized algorithm was 100% sensitive for the AF detection on an individual basis.

Automatic detection of AF has been accomplished in several other monitoring systems, including implantable loop recorders and pacemaker systems.[6,7]

In a prior pilot study[8] using 2837 one-lead ECGs, an algorithm was created in order to distinguish normal from pathological ECGs.

The aim of the present study is to validate if a mathematical algorithm could be used to safely distinguish normal from pathological ECGs, hence reducing the cost of manual ECG interpretation and facilitating AF screening.

Methods

The STROKESTOP study

The method, and initial results from the STROKESTOP study has been described in a prior publication.[9] In summary, all individuals born in 1936 or 1937 and living in Stockholm county ($n = 23\,888$) or in the rural region of Halland ($n = 4880$) at the end of 2011 was randomized in a 1:1 fashion to be invited by mail to participate in a screening programme for AF, or to enter a control group. Individuals without known AF were equipped with a handheld one-lead device (www.zenicor.com) for intermittent ECG recordings during a 2-week period, and instructed to register ECGs using their thumbs two times a day. The device has an integrated mobile transmitter that sends 30 s ECG strip data to a database. The AF was defined as at least one 30-s[10] recording with irregular rhythm without p-waves, or a minimum of two similar episodes lasting 10–29 s during 2 weeks of intermittent recording. All 189 715 ECG recordings were manually interpreted by specially trained research nurses, and all abnormal ECGs were referred to the investigating cardiologist. All individual ECGs were annotated as AF, other arrhythmia or not annotated (meaning sinus rhythm). In addition, short episodes of AF (<30 s) could be annotated as AF, although not fulfilling AF criteria, leading to the possibility of ECGs being annotated as AF, but the patient not getting an AF diagnosis. The ECG interpretation skills of the nurses were verified by way of random controls. All individuals with newly detected AF, and AF patients not treated with OACs, were offered structured follow-up by a cardiologist to ensure adequate treatment according to current European guidelines.[10] All participants in the STROKESTOP study gave written informed consent, and the study was approved by the local ethics committee (DNR 2011-1363-31/3).

In this study, we aim to validate the performance of an AF detection algorithm compared with manual ECG analysis by specially trained nurses and physicians (gold standard) in intermittent ECG recordings. For this purpose, we retrospectively analysed a dataset containing 80 149 ECGs from the first consecutive 3209 participants.

The algorithm

The algorithm consists of five analysis stages: pre-processing for disturbance reduction, beat detection and classification, rhythm pattern analysis and categorization, average beat analysis, and p-wave analysis, *Figure 1*.

Quality assessment is an important part in all five analysis stages. The rhythm pattern analysis investigates the entire rhythm pattern, taking both the exact timing and morphology of each individual beat into account. Irregularities are measured at individual beat level as well as with increasing temporal scales up to the entire signal length. In particular, during the rhythm pattern analysis stage the algorithm identifies and counts all shorter or longer R–R intervals (that cannot be explained by ectopic beats/AV block II/bigemini or trigemini rhythms) using several different time scales (sliding windows of 3, 5, 10 s). The algorithm allows a few spread-out irregular beats to exist without the need for manual inspection, but if several irregularities exist within these short-time windows the algorithm cannot

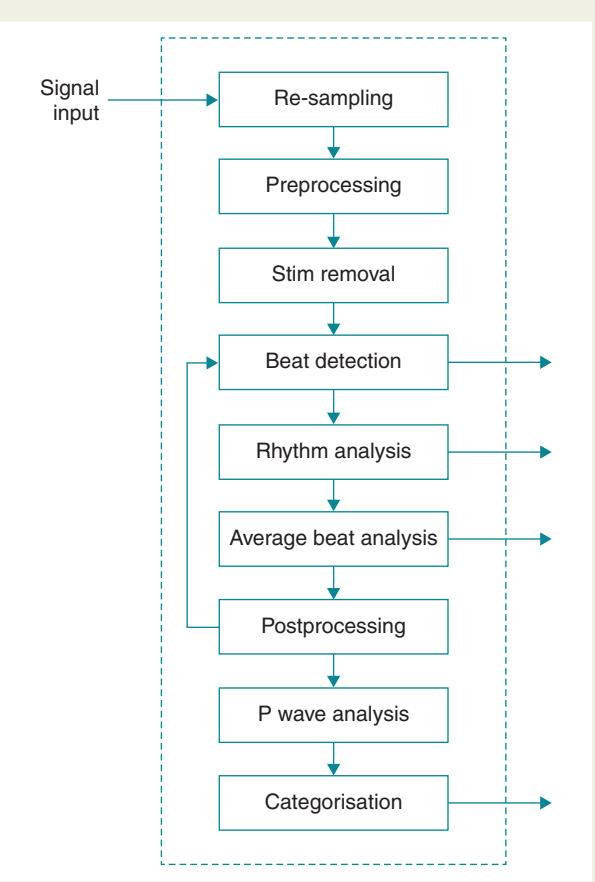

Figure 1 The five steps of processing of signals performed by the algorithm.

exclude a possible AF episode, and thus will mark it as irregular sequence, for manual inspection. Average beat analysis is performed for all different beat classes with their individual morphologies. The p-wave analysis is performed using a robust waveform clustering approach.

Electrocardiograms from the handheld devices were sorted by the algorithm into four categories: (0) poor quality, (1) only minor rhythm deviation or sinus rhythm, (2) irregular rhythm requiring manual interpretation (possible AF), and (3) other pathologies. Group 0 contains signals with obvious quality issues. Group 1 consists of signals with sinus rhythm or only very minor rhythm deviations. Group 2 contains all cases of irregular rhythms, hence recordings which the system cannot reject as possible AF. Group 3 includes ECGs with possible indications for pacemakers, fast regular rhythms, bigemini and trigemini rhythms and registrations with over five supraventricular/ventricular ectopic beats.

The arrhythmia analysis has been tuned in a pilot study[8] and by analysing Lead I in 12-lead ECG databases during sinus rhythm and AF (www.cardiolund.com).

An additional quality control was done by manual evaluation of 2961 randomly selected ECGs categorized by the algorithm as only minor rhythm deviations or sinus rhythm, that were sent for quality control to Heidelberg Medical Consultancy Pvt. Ltd, where three cardiotechnicians blinded to the results of the algorithm reviewed the ECGs.

Statistics

All categorical variables are expressed as numbers and percentages. Manual interpretation of ECGs was used as gold standard. Sensitivity, specificity, negative predictive value, and positive predictive value were calculated; and 95% confidence intervals (CIs) for sensitivity and specificity were calculated using the Clopper–Pearson method.

Results

Sinus rhythm/minor rhythm deviations

We analysed 80 149 ECG recordings of 30 s duration in 3209 individuals. Of these, 69 789 (87.1%) ECGs were classified by the algorithm as containing only minor rhythm deviations or sinus rhythm (Group 1). Into the irregular rhythm category (Group 2) 6092 (7.6%) ECGs were sorted requiring manual interpretation. The algorithm classified 3475 (4.3%) as other pathology (Group 3) and 793 (0.99%) were deemed poor quality (Group 0; *Figure 2*).

In the group classified by the algorithm as only minor rhythm disturbance or sinus rhythm, there was agreement between the system and the manual interpretation in 69 758 of 69 789 ECGs. The algorithm had 99.9% negative predictive value (*Table 1*).

In the additional blinded quality check by Heidelberg Medical Group 2961 ECGs sorted by the algorithm as only minor rhythm disturbance or sinus rhythm. The manual quality check labelled 2635 ECGs as no rhythm deviation/normal sinus rhythm, 235 bradycardias with a heart rate below 60 bpm, and 74 as sinus arrhythmia, which gives in total 2944 ECGs (99.4%). Very few (<20) were bradycardia below 50 bpm. The rest of the signals were manually annotated as possible AF (10), AV block II (1), fast rhythm (6), and fast/slow episode (2). Neither of the 10 cases manually annotated as

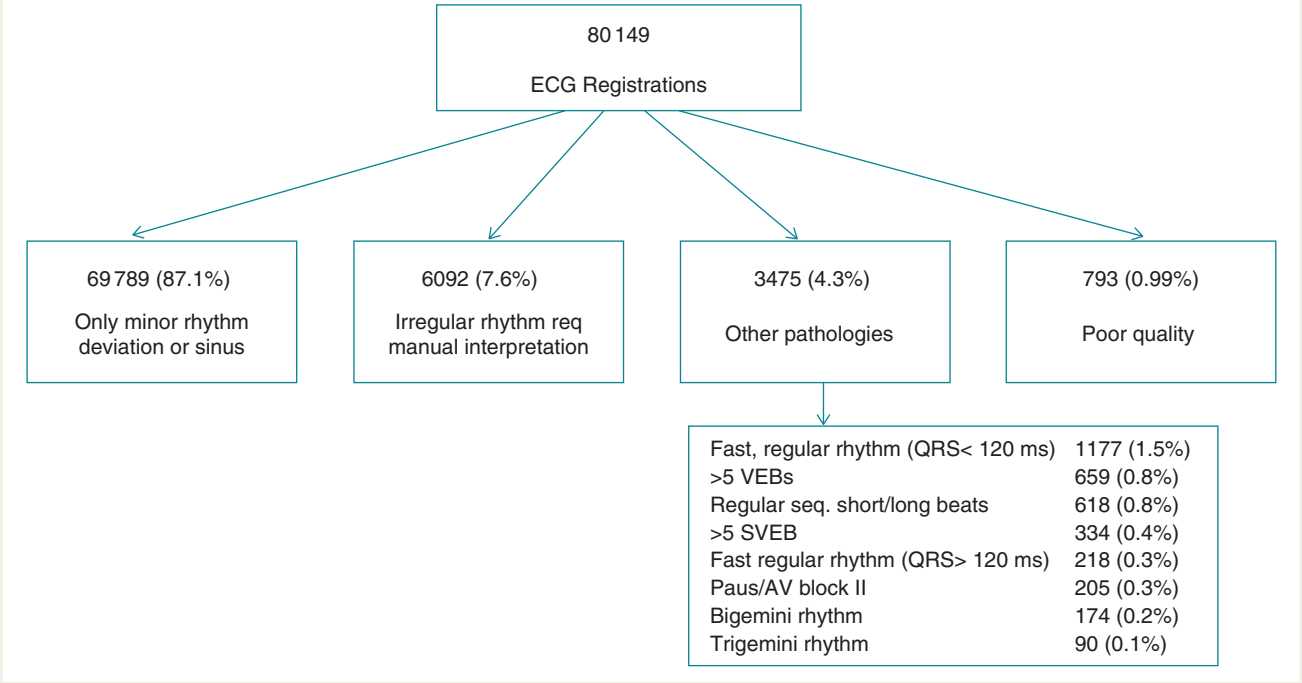

Figure 2 Sub classification of ECGs by the algorithm. Seq, sequences; SVEB, supraventricular ectopic beats; PPM, pacemaker; VEB, ventricular ectopic beats; Bigemini rhythm, ectopic beat every other beat; Trigemini, ectopic beat every third beat.

Table 1 AF = ECGs manually annotated as AF. Sensitivity 97.8%, specificity 88.2%, PPV 2.8%, and NPV 99.99%

	AF diagnosed by manual interpretation	No AF by manual interpretation	Total
Algorithm indicated pathology[a]	272	9295	9567
No pathology indicated by algorithm[b]	6	69 783	69 789
Total	278	79 078	79 356

[a]Positive finding indicates that ECG deemed as either possible AF or other pathology by the algorithm.
[b]Negative finding indicates that ECG deemed as no rhythm deviation by the algorithm.

possible AF cases were annotated as AF by the investigator. The AV block II case was solely an annotation mistake as it was also commented as normal sinus rhythm, and the patient had sinus rhythm. The manual interpretation estimated the heart rate slightly differently than the automated algorithm explaining that some ECGs get manually interpreted as fast rhythm (>100 bpm).

AF detection

Of 278 signals manually interpreted as AF, the algorithm agreed that 272 ECGs were pathological. Hence sensitivity for AF detection by the algorithm was 97.8% (95% CI 95.4–99.1). The six ECGs that differed belonged to five patients, two of whom did not get an AF diagnosis due to not fulfilling diagnostic criteria (episodes lasting <30 s). The three remaining patients did get a diagnosis of AF, and all had other registrations marked as Category 2 'irregular rhythm requiring manual interpretation', hence no patient with AF was missed by the algorithm.

In Group 2 classified by the algorithm as 'Irregular rhythm requiring manual interpretation' there were 6092 ECGs, and 250 of these (4.1%) were manually interpreted as AF. In total the positive predictive value of the algorithm was 2.8%. The total specificity of the algorithm was 88.2% (95% CI 88.0–88.4). A randomized control of 100 ECGs from group two showed that the classification into the group with irregular rhythm was mainly due to supraventricular ectopic beats or false beat detections due to disturbances.

Without the algorithm 288 ECGs needed to be manually interpreted in order to find one ECG with AF. Using the algorithm 35 ECGs need to be manually interpreted in order to find AF or other rhythm deviations.

Using the algorithm would per 1000 individuals participating in screening (on average producing 26 ECGs), result in 22 620 ECGs fewer needing manual interpretation. This could, based on a 20-s review per ECG save 125.7 h, and the cost reduction would (depending on salary) amount to ~12 570 Euros/1000 screened.

The group 'Other pathology' included 3475 ECGs (4.3%), mainly consisting of 1177 (1.5% of all ECGs) ECGs with fast, regular rhythm over 100 bpm, followed by 659 (0.8%) ECGs with over 5 ventricular ectopic beats, and 618 (0.8%) ECGs with sequences of repeated short/long beats (*Figure 2*). ECGs with high-grade AV block and sinus arrests >2.2 s constituted 205 ECGs (0.3%). Of the ECGs manually interpreted as AF 22 were found in the group 'Other Pathology', with the majority in the fast regular rhythm group. These mainly represented ECGs with atrial flutter, as the system only allowed one type of manual annotation for AF and atrial flutter, so an ECG manually annotated as AF could indeed be AF or atrial flutter, and several of these ECGs came from one individual.

In total 9567 out of 80 149 (11.9%) ECGs were marked as either possible AF or other pathology, requiring manual interpretation.

Results per individual

In total ECGs from 3209 individuals were collected. Eighty-four individuals were manually annotated as having one or several episodes of AF and were classified in the study as individuals with newly discovered AF. Of these, the system classified 80 individuals as irregular rhythm requiring manual interpretation and 4 individuals as having other ECG pathologies. Hence all individuals with newly discovered AF would have been highlighted by the system for manual interpretation.

In the algorithm 966 (30%) individuals were classified as free of ECG pathologies, and the remainder had 1 or more ECGs to be manually interpreted.

Discussion

In this study, we show that an automatic sorting algorithm could safely discern one-lead ECGs that had been manually interpreted as normal, with a negative predictive value of 99.9 and 100% sensitivity on an individual level for AF detection. Usage of the automatic sorting algorithm reduces the number of ECGs that would need manual inspection eight-fold.

Mass-screening for AF in an elderly population using intermittent ECG recordings has in a prior study shown cost-efficacy.[11] Using the automated algorithm has the potential of reducing the costs for intermittent ECG screening, and furthermore increasing the simplicity of the procedure further enabling population screening for AF.

In an automated system, it is crucial not to miss any AF episodes as AF might have devastating yet preventable consequences. The sensitivity of an automated system must be very high to be acceptable. In this system sensitivity was effective at a rate of 97.8% on an ECG level and 100% on an individual level. The cost of this high level of sensitivity is a reduced specificity.

The positive predictive value of the interpretation of possible AF is indeed low, and all ECGs marked as possible AF require manual interpretation. Of the 6092 ECGs marked as 'irregular rhythm requiring manual annotation' only 250 (4.1%) were manually interpreted as AF. However, no individual with AF was missed by the algorithm using this approach. In the majority of individuals there will be some degree of manual interpretation. In order to ensure safety, the system also annotates other pathologies of interest such as candidates for possible pacemaker treatment.

Prior studies have shown that atrial flutter is difficult to diagnose with intermittent one-lead ECGs[12] and in order not to miss any individuals with atrial flutter—hence indication for OAC treatment—

all individuals with fast, regular rhythm (>100 bpm), regardless of QRS duration, need to be manually interpreted.

A special group of interest are individuals with an increased frequency of supraventricular ectopic beats (SVEB). In one study over 30 SVEBs/h showed an increased risk of stroke[10] and predisposition for the development of AF. In individuals with cryptogenic stroke, SVEBs were also a significant predictor for future AF.[13] These individuals could be identified using this algorithm and more extensive screening could be initiated.

Limitations

The intermittent ECGs in this study only use Lead I from a handheld ECG device, which has been shown to have a high sensitivity and specificity for AF detection.[12] Handheld measurements are more susceptible to noise and rapid baseline wander, which may cause false beat detections. In our study ~1% of the ECGs were deemed too unfit to interpret. This study was performed in an elderly population who are more likely to be afflicted with diseases causing tremor, which would affect quality. As atrial activity, depicted by p-waves, is less prominent in Lead I, the algorithm can only utilize the presence of p-waves when they are clearly discernible in the background noise. These two aspects, the higher level of noise and the absence of p-waves, are probably the keys to the limited specificity of the system, as episodes classified as irregular rhythm by the system were mainly due to several SVEB without discernible p-waves or false beat detections due to the noise level. In future studies, a hip lead might be used to record ECGs in Lead II, which might improve ECG quality further, or other long-term ECG recording devices using several leads might be required in this group.

Manual interpretation was used as the gold standard. Even if nurses and physicians were well trained with regard to arrhythmia detection, episodes of AF could have been missed, and episodes manually interpreted as AF could have been wrong.

The differences observed between manual and automatic interpretation is driven mainly by the large amount of false positive ECGs by the automatic algorithm. This is a safety aspect to make sure no individuals with AF are missed by the automatic screening.

Conclusion

To conclude, an automatic screening algorithm safely identifies normal ECGs and reduces the need for manual interpretation of individual ECGs with >85%. The sensitivity of the system in identifying AF is 100% on an individual level.

Funding

E.S. has received lecture fees from Merck Sharp & Dohme, Bristol-Myers Squibb-Pfizer, Boehringer-Ingelheim, Bayer, and Sanofi, and a research grant from Boehringer-Ingelheim. M.S. has no disclosures. J.E. has received consultancy fees from Sanofi and Pfizer, lecture fees from Astra Zeneca, Boehringer-Ingelheim, Medtronic, Bayer and Bristol-Myers Squibb and an unrestricted research grant from Boehringer-Ingelheim. F.A.-K. has received lecture fees from Boehringer-Ingelheim, Bayer, and Pfizer. L.F. has received grants and/or given lectures sponsored by Bayer, Boehringer-Ingelheim, Bristol-Myers Squibb, Pfizer, Sanofi, and St Jude. V.F. lecture fees/grants from Merck-Sharp and Dohme, Medtronic, Bayer, Boehringer-Ingelheim, Laerdahl and research collaboration with St Jude Medical and Medtronic. M.R. has received unrestricted research grants/lecture fees/consultant fees from Bayer, Boehringer-Ingelheim, Britsol-Myers Squibb, Medtronic, Pfizer, St Jude, and Zenicor.

Conflict of interest: none declared.

References

1. Friberg L, Bergfeldt L. Atrial fibrillation prevalence revisited. *J Intern Med* 2013;**274**:461–8.
2. Camm AJ, Lip GY, De Caterina R, Savelieva I, Atar D, Hohnloser SH *et al.* 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. *Europace* 2012;**14**:1385–41.
3. Fitzmaurice DA, Hobbs FD, Jowett S, Mant J, Murray ET, Holder R *et al.* Screening versus routine practice in detection of atrial fibrillation in patients aged 65 or over: cluster randomised controlled trial. *BMJ* 2007;**335**:383.
4. Lowres N, Neubeck L, Redfern J, Freedman SB. Screening to identify unknown atrial fibrillation. A systematic review. *Thromb Haemost* 2013;**110**:213–22.
5. Friberg L, Engdahl J, Frykman V, Svennberg E, Levin LA, Rosenqvist M. Population screening of 75- and 76-year-old men and women for silent atrial fibrillation (STROKESTOP). *Europace* 2013;**15**:135–40.
6. Passman RS, Weinberg KM, Freher M, Denes P, Schaechter A, Goldberger JJ *et al.* Accuracy of mode switch algorithms for detection of atrial tachyarrhythmias. *J Cardiovasc Electrophysiol* 2004;**15**:773–7.
7. Brignole M, Bellardine Black CL, Thomsen PE, Sutton R, Moya A, Stadler RW *et al.* Improved arrhythmia detection in implantable loop recorders. *J Cardiovasc Electrophysiol* 2008;**19**:928–34.
8. Stridh M, Rosenqvist M. Automatic screening of atrial fibrillation in thumb-ECG recordings. *Comput Cardiol* 2012;**39**:193–6.
9. Svennberg E, Engdahl J, Al-Khalili F, Friberg L, Frykman V, Rosenqvist M. Mass screening for untreated atrial fibrillation: the STROKESTOP Study. *Circulation* 2015. doi:10.1161/CIRCULATIONAHA.114.014343.
10. Binici Z, Intzilakis T, Nielsen OW, Kober L, Sajadieh A. Excessive supraventricular ectopic activity and increased risk of atrial fibrillation and stroke. *Circulation* 2010;**121**:1904–11.
11. Aronsson M, Svennberg E, Rosenqvist M, Engdahl J, Al-Khalili F, Friberg L *et al.* Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ECG recording. *Europace* 2015;**17**:1023–9.
12. Doliwa PS, Frykman V, Rosenqvist M. Short-term ECG for out of hospital detection of silent atrial fibrillation episodes. *Scand Cardiovasc J* 2009;**43**:163–8.
13. Kochhauser S, Dechering DG, Dittrich R, Reinke F, Ritter MA, Ramtin S *et al.* Supraventricular premature beats and short atrial runs predict atrial fibrillation in continuously monitored patients with cryptogenic stroke. *Stroke* 2014;**45**:884–6.