Appendix D: Clinical evidence tables

Table 19: Gandolfo, 2015[79]

Reference	Gandolfo, 2015[79]				
Study type	Observational				
Recruitment	consecutive				
Setting	Stroke unit				
Country	Italy				
Sample size	207				
Sample characteristics	Stroke unit inpatients; 103 women; mean age 77.7 years; 86.5% recent ischaemic CVA/TIA; 13.5% haemorrhagic stroke; within 48 hour window post stroke				
Inclusion criteria	Patients admitted to stroke unit because of recent (<48 hours) TIA/stroke				
Exclusion criteria	Patients with rhythm controlled by pacemakers or defibrillators				
Index test(s), including number of repetitions and duration	Triple blood pressure measurement by the **Microlife AFib device** (total session time 10 minutes) usually on day of admission to SU, and <48hrs. Done by trained SU nurse				
Gold standard	**Standard 12 lead ECG**, interpreted by expert cardiologist (ECG performed by trained SU nurse). Normally done on day of admission to SU, and <48 hours of admission.				
Expertise of index test interpreter	Not stated				
Simultaneous index/gold vs non simultaneous	Not simultaneous; gold standard followed index 'immediately after the end' during a 10 minute evaluation session, and never>48 hours difference between them.				
Results		**Gold standard +ve**	**Gold standard -ve**	**Total**	**Sensitivity**:0.895 (0.760-0.958) **Specificity**:0.988(0.958-0.997) **PPV**: 0.94 **NPV**:0.98
	Index test +ve	34	2	36	

Reference	Gandolfo, 2015[79]				
	Index test -ve	4	167	171	
	Total	38	169	207	
Source of funding	None reported. Italian Association against Stroke provided the Microlife devices free of charge.				
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): None				

Table 20 Kaleschke, 2009[117]

Reference	Kaleschke, 2009[117]
Study type	Observational
Recruitment	consecutive
Setting	Outpatient AF clinic (AFNET centre at University Hospital)
Country	Germany
Sample size	508
Sample characteristics	66% male; mean age 61.4; mean BMI 26.6;
Inclusion criteria	Clinical indication for 12 lead surface ECG; No other details provided.
Exclusion criteria	<18 years; pacemaker or defibrillator
Index test(s), including number of repetitions and duration	Patient-activated 'leadless' ECG device (Omron HeartScan HCG-801-E). Lightweight, handheld ECG recording system with LCD display and digital storing capacity for offline, digital analysis8 (height 121 mm, width 67 mm, depth 24 mm, and weight 130 g). It records 30 s of a single-channel ECG. The ECG is recorded as the potential between two stainless-steel electrodes integrated into the surface of the device. The device is ready to record a few seconds after turning it on. For ECG recording, the lower surface of the device, which contains one electrode, is attached to the chest. The index finger of the right hand holds the device. This finger is in contact with the second electrode. By pressing the start button, the recording is activated for 30 s. Done by patient after instruction – unclear of expertise of instructor and the quality of instruction. Data emailed to centre.
Gold standard	12 lead surface ECG. Analysed by single blinded observer. Expertise of operator and interpreter unclear but likely to be high as lead author of study, who appears to be a cardiologist

Reference	Kaleschke, 2009[117]				
Expertise of index test interpreter	Analysed by single blinded observer.				
Simultaneous index/gold vs non simultaneous	No – 10-15 second delay				
Results		**Gold standard +ve**	**Gold standard -ve**	**Total**	**Sensitivity**:0.99(0.96-1.00) **Specificity**:0.96(0.94-0.98) **PPV**:0.92(0.86-0.96) **NPV**:1.00 (0.98-1.00)(this does not make sense given that sensitivity is not 100%)
	Index test +ve				
	Index test -ve				
	Total	128 (or 143)*	377 (or 362)*	505	
	*Note: discrepancy between number with AF given in text and table. In text number with actual AF is 128, but in table it is 143. Raw data not provided by paper and due to these discrepancies (plus NPV provided not tallying with sensitivity) the raw data has not been calculated				
Source of funding	The study was conducted by AFNET which received financial support for this study in the form of an unrestricted grant by Omron Healthcare.				
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious (population not clearly that defined in protocol – people with cardiovascular risk factors for AF (other than just age)and/or symptoms suggestive of AF)				

Table 21 Kearley, 2014[128]

Reference	Kearley, 2014[128]
Study type	Observational
Recruitment	consecutive
Setting	GP practices
Country	UK
Sample size	1000

Reference	Kearley, 2014[128]
Sample characteristics	Mean age 79.7; 49.3% male; Hx of AF 11%; HF 3.1; hypertension 53%; DM 12.2%; Stroke 3.1%; TIA 6.5%; Patients with AF on AADs 8.7%
Inclusion criteria	Participants aged 75 or over, living at home from 6 General practices in the UK
Exclusion criteria	People with pacemakers and defibrillators; unable to give consent; terminal illness; other reasons why participation is inappropriate at discretion of GP;
Index test(s), including number of repetitions and duration	3 methods tested in the following order, by any of 9 registered nurses working at the practices: 1. Watch BP –modified oscillometric BP monitor which flashes when it detects an irregular pulse during automatic BP measurement 2. Omron auto analysis – using an Omron monitor (model HCG-801) which involved placing one electrode on the bare chest wall 5 cm below the nipple, while the patient held the other electrode with the right index finger. The monitor records a single-lead ECG tracing, and displays a message indicating the presence of possible AF. The device's analysis algorithm includes several cardiac rhythms which could potentially be AF, including fast and irregular, slow and irregular, irregular and those where analysis is impossible. The single-lead recording and text message were recorded and saved for later downloading and analysis 3. Merlin ECG trace the nurse applied a Merlin ECG event recorder (Meditech Ltd, Hungary) which resembles a watch, on participants' left wrist. The participant covered the electrode on the face of the device with the palm of their right hand for 30 s. The recording, with no automated analysis, was saved to a computer for later downloading and analysis. Unlike the Omron, the Merlin monitor does not require removal of any clothing, making it possible for use in public settings, an advantage for participants experiencing an intermittent arrhythmia. The nurse recorded the results of the WatchBP monitor and the Omron automated text message during the initial examination. Each single-lead ECG trace was sent for interpretation to two independent cardiologists after removing all clinical information and patient identification except for date of birth and the text message (Omron only). Details of the algorithms not provided
Gold standard	12 lead ECG, independently interpreted by one of 4 cardiologists, all of which had completed cardiology specialist training of 5-6 years. Performed immediately after the index tests
Expertise of index test interpreter	Automated for Watch BP and Omron / cardiologists for Omron and Merlin
Simultaneous index/gold vs non simultaneous	No – the gold standard followed the index tests on the same time, but interval unclear.

Reference	Kearley, 2014[128]
Results	*If unclear on index test it was always counted as a positive test* Watch BP Sensitivity: 94.9% (87.5 – 98.6) Specificity: 89.7% (87.5-91.6) PPV: 44.1 (36.5-51.9) NPV: 99.5(98.8 – 99.9) TP 75, FN 4, FP 95, TN 825 Omron auto-analysis Sensitivity: 98.7% (93.2 – 100) Specificity: 76.2% (73.3-78.9) PPV: 26.3 (21.3-31.7) NPV: 99.9(99.2 – 100) TP 78, FN 1, FP 219, TN 701 Omron ECG trace interpreted by the 4 cardiologists (pooled results using meta-analysis of the 4 cardiologists results) Sensitivity: 94.4% Specificity: 94.6% Merlin ECG trace interpreted by the 4 cardiologists (pooled results using meta-analysis of the 4 cardiologists results) Sensitivity: 93.9% Specificity: 90.1%
Source of funding	This publication presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0407-10347) and the NIHR School for Primary Care Research.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 22 Kollias, 2018[132]

Reference	Kollias, 2018[132]
Study type	Observational
Recruitment	consecutive
Setting	Hypertension clinic
Country	Greece
Sample size	100
Sample characteristics	Patients attending a hypertension clinic. Age 70.6; BMI 29.1; 52.9% male; 11% stroke; 85% hypertension; 20% DM; 7% CAD; 82% antihypertensive treatment; CHADSVASC score 3.06
Inclusion criteria	Patients attending a hypertension clinic for BP assessment, treated or untreated for hypertension; aged >=65; aged 50-64 with symptoms suggesting arrhythmias or with stroke/AF history; clinical indication for ambulatory blood pressure monitoring
Exclusion criteria	Pacemaker implantation
Index test(s), including number of repetitions and duration	24 hour ambulatory blood pressure monitoring (ABPM), using the validated oscillometric device Microlife WatchBP O3 Afib with measurements programmed at 20-minute intervals for 24 hours. This device has an implemented algorithm for automated AF detection during each BP measurement. The presence of AF is depicted in the ABPM when AF is detected, and the total number of BP readings with AF detection is reported. The AF detector functions as follows: the device measures the last 10 pulse intervals during cuff deflation and calculates the mean and SD of the time intervals. Each of the 10 pulse beat intervals that is 25% longer or 25% shorter than the mean time interval is discarded, to reduce the effect of premature beats. The remaining time intervals are used to calculate the irregularity index, defined as the SD divided by the mean of the time intervals. If the irregularity index exceeds a threshold value of 0.06, an AF symbol is ascribed indicating that the patient has AF. Subjects were instructed to perform their usual daily activities but remain still with their arm extended and relaxed during each BP measurement. Day and night periods were defined according to the individual patients' diaries
Gold standard	24 hour Holter recording using the SpiderView (ELA Medical, Sorin Group) multichannel system recorder which was performed simultaneously with 24-hour ABPM. Time was synchronized in the 2 devices before each application. A cardiologist (one of the 2 lead study authors) assessed the recordings using the EasyScope Multiday ELA Medical software. Artifacts, falsely interpreted as ectopic beats, were subtracted from the ECG report when calculating the number of ectopic beats. Criteria for abnormal 24-hour ECG recording were the following: flutter or AF episode of any duration; supraventricular or ventricular ectopic beats >720/24 hours; supraventricular couplets ≥50/24 hours; supraventricular or ventricular bigeminy ≥50/24 hours; supraventricular runs ≥20/24 hours or ≥10 beats/run; ventricular tachycardia of any duration; sinus pause >3s; and second- or third-degree atrioventricular block. These criteria were selected to include all clinically important and potentially hazardous arrhythmias, as well as arrhythmias that increase the risk of AF and stroke.
Expertise of index test interpreter	Fully automated

Reference	Kollias, 2018[132]
Simultaneous index/gold vs non simultaneous	Simultaneous
Results	Sensitivity: 93% (91% to 94%) Specificity: 87% (86% to 88%) TP 1013, FN 78, FP 78, TN 4609 Note: these are not based on individual patient 'diagnoses' – instead these are based on the entire sample of 6410 valid ABPM readings from the 100 participants over the 24 hours (64 valid readings per patient, based on a reading every 20 minutes for 24 hours [thus 72 possible readings per patient]). Therefore we have considerable increase in the precision of the accuracy, which does not take into account correlation between values derived from the same person.
Source of funding	Microlife, Widnau, Switzerland provided ambulatory blood pressure monitors with atrial fibrillation detector for this study, but was not involved in the study design, analysis, and article preparation.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): None

Table 23 Marazzi, 2012 [162]

Reference	Marazzi, 2012[162]
Study type	Observational
Recruitment	consecutive
Setting	Hypertension Clinic
Country	Italy
Sample size	550
Sample characteristics	Mean age 67 years; 54.3% male; bp 139.8/86.9
Inclusion criteria	Patients referred to hypertension clinic
Exclusion criteria	<18 years; pacemaker; implanted defibrillator
Index test(s), including number	1. Microlife BP A200 Plus – an automated oscillometric bp measurement device. A specially dedicated algorithm adds an extra function of AF detection, via evaluation of pulse rate irregularity. Device measures last 10 pulse intervals during cuff

Reference	Marazzi, 2012[162]
of repetitions and duration	deflation and calculates mean and sd of the intervals. The irregularity index was defined as the sd/mean of the time intervals. After deletion of outliers (+/- 25% of mean) to reduce effect of premature beats, if the irregularity index exceeded 0.06, the rhythm was considered irregular. This was used on one arm. 2. Omron M6 device – an automatic oscillometric device for self-measurement of BP. Also has an additional function of detecting AF. The threshold irregularity index was set at 0.066. This was done simultaneously on the other arm of patients
Gold standard	12 lead ECG interpreted by board-certified cardiologists blinded to the readings of the devices.
Expertise of index test interpreter	NA – both index tests are fully automated
Simultaneous index/gold vs non simultaneous	Yes.
Results	Omron M6 Sensitivity 100%, Specificity 94.2%; TP 101, FN 0, FP 23, TN 379 Microlife BP A200 Plus Sensitivity 92.1%, Specificity 97%; TP 93, FN 8, FP 12, TN 390
Source of funding	None reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): None

Table 24 Koltowski, 2019[133]

Reference	Koltowski, 2019[133]
Study type	Observational
Recruitment	consecutive
Setting	Tertiary cardiovascular centre
Country	Poland
Sample size	100
Sample characteristics	Mean age 68; male 66%; patients at a tertiary cardiovascular care center, admitted for hospital elective and treatment procedures for various cardiac conditions.; body mass 80.7kg; BMI 28; smoking history 43.5%; DM 20.4%; hypertension 68.4%; dyslipidemia 46.4%; CKD 32.7%; thyroid dysfunction 18.4%; COPD 6.12%; Stroke 17.35%; PAD 12.24%; stable angina 47.4%; ACS 15.31%; MI

Reference	Koltowski, 2019[133]
	25.5%; PCI/CABG 27.6%; other cardiac surgery 3.1%; HF 43.9%; LVEF 49%; AF 34.7%; CIED implanted 34.7%; pacemaker 24.5%; ablation 6.1%
Inclusion criteria	Undergoing regular 12-lead ECG due to standard diagnosis on admission in stable state
Exclusion criteria	Need for urgent medical care
Index test(s), including number of repetitions and duration	Kardia mobile ECG. Kardia Mobile (KM) (AliveCor Inc., San Francisco, CA, USA) is a portable, mobile, connected electrocardiogram (ECG) device available to iOS and Android platform smartphone owners. It consists of a small device with two conducting plates that wirelessly connect with a smartphone, and an application installed on user smartphones. It enables one-lead ECG recording e.g. in cases of the onset of unsettling symptoms (palpitations, chest pain, dyspnea, and others). KM was designed to detect periods of atrial fibrillation (AF), which, if confirmed by the FDA-approved algorithm, can then be reported to the physician responsible for the follow-up of a given patient.
Gold standard	12 lead ECG, carried out first. Two technicians responsible for 12 lead ECG measurement. Analysed by 3 independent teams comprising 2 cardiologists each.
Expertise of index test interpreter	A physician recorded KM ECGs. Analysed by 3 independent teams comprising 2 cardiologists each. ECG quality (good, acceptable, poor), rhythm (sinus rhythm, AF, atrial flutter [AFl] or pacemaker rhythm), presence of pathological Q wave as well as PQ, RR and QT measurements were assessed.
Simultaneous index/gold vs non simultaneous	No – index test carried out immediately after 12 lead ECG.
Results	No raw diagnostic data, or data from which the diagnostic data could be calculated, were provided in the paper. Sensitivity: 92.8% Specificity 100%
Source of funding	The research was performed within the statutory fund of the First Chair and Department of Cardiology of the Medical University of Warsaw and received no external funding.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 25 Kristensen, 2016[138]

Reference	Kristensen, 2016[138]
Study type	Observational
Recruitment	Selective case-control
Setting	GP clinic
Country	Denmark
Sample size	93
Sample characteristics	54% male; age 67; IHD 11%; hypertension 54%; DM 21%; known AF diagnosis 36%; Medication affecting heart rhythm 47%
Inclusion criteria	Patients from a GP clinic in Aalborg, Denmark, who performed a routine 12-lead ECG were invited to participate. The invited patients either had known paroxysmal AF or were invited among patients who came for an annual routine health check. The aim was to include 30–50% with a diagnosis of AF and 50–70% without AF. Thus this was not a consecutive sample.
Exclusion criteria	Patients with severe dementia, mental illness or poor ECG readings
Index test(s), including number of repetitions and duration	A 30 s three-lead recording using a PEM device (Portable ECG Monitor, Beijing Choice Electronic Technology Co., Ltd., Beijing, China) The PEM is capable of storing the data/ECG. The ECGs were transferred from the PEM to a personal computer and were evaluated after printing. The PEM recordings were analysed by two GPs who were blinded for the results of the ECG recordings as well as for the patients' characteristics except for gender and age.
Gold standard	Standard 10 second 12 lead ECG. Blinded to the PEM registrations the ECG recordings were evaluated by a senior GP and a cardiologist specialized in Electrophysiology (SR). Another cardiologist settled any disagreement over evaluation. We defined AF as irregular supraventricular arrhythmia without p-waves at the baseline.
Expertise of index test interpreter	Expertise of 2 GPs not described
Simultaneous index/gold vs non simultaneous	Yes, simultaneous
Results	Sensitivity: 86.7% Specificity: 98.6% PPV: 86.7% NPV: 97.3% TP 13, FN 2, FP 1, TN 73
Source of funding	The PEM device was financed by the Research Unit for General Practice in the North Denmark Region, but otherwise the project received no external funding.

Reference	Kristensen, 2016[138]
Limitations	Risk of bias (QUADAS 2 – risk of bias): No Serious risk Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 26 Wiesel, 2014 [279]

Reference	Wiesel, 2014[279]
Study type	Observational
Recruitment	consecutive
Setting	Outpatient cardiology clinics
Country	USA
Sample size	183
Sample characteristics	Age 74; male 59%; ethnicity: white/Black/Asian/Hispanic 71%/16%/4%/9%; hypertension 92%; DM 25%; CHF 17%; Stroke 6%; CAD 41%; Hx AF 27%; ACEs 33%; ARBs 17%; diuretics 26%; beta blockers 62%; calcium blockers 33%; digoxin 9%; anticoagulant 23%; AADs 3%
Inclusion criteria	All patients aged >50 attending 2 outpatient cardiology clinics
Exclusion criteria	Patients with pacemakers or defibrillators
Index test(s), including number of repetitions and duration	Omron M6 Comfort – 1 reading only used Microlife BP A 200 – 3 sequential readings used (combined to give a single reading based on the majority rule in which the final reading is considered positive for AF if at least 2 of 3 individual readings are positive for AF).
Gold standard	12 lead ECG done by technician, prior to index tests. Interpreted by a board certified cardiologist who was blinded to the results of the BPM readings
Expertise of index test interpreter	Unclear, but likely to be automated
Simultaneous index/gold vs non simultaneous	No, 12 lead EGC done before index tests (interval not reported)
Results	Omron Sensitivity: 30% (15.4 to 49.1)

Reference	Wiesel, 2014[279]
	Specificity 97% (92.5 to 99.2) TP 10, FN 20, FP 5, TN 148 Microlife (majority rule after 3 readings) Sensitivity: 100% (85.9 to 100) Specificity 92% (86.2 to 95.7) TP 30, FN 0, FP 12, TN 141
Source of funding	This study was funded by Microlife Corporation, Taipei, Taiwan.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): None

Table 27 Wiesel, 2013[278]

Reference	Wiesel, 2013[278]
Study type	Observational
Recruitment	consecutive
Setting	General Internist Clinics
Country	USA
Sample size	160
Sample characteristics	Age 67; male 37%; white 71%; black 5%, Hispanic 5%; Asian 4%; hypertension 85%; DM 12%; CHF 6%; stroke 3%; AF 12%; CHADS2 1.4; ACEI 27%; ARB 16%; Ca channel blocker 15%; beta blocker 27%; diuretic 28%; warfarin 10%
Inclusion criteria	Patients attending general internists offices; more than or equal to 1 of the following criteria: Age >=65; hypertension, DM, CHF, stroke; patients allowed to have AF
Exclusion criteria	Pacemakers; defibrillators
Index test(s), including number of repetitions and duration	AF-BP monitor device, to take home and use daily for 30 days, charting results on a log. If AF event detected automatically, subject had to take 2 additional sequential readings. Using the majority rule, if either 2 or all 3 indicated AF, the subject was to wait 1 hour and obtain a fourth reading. If this last reading indicated AF, the subject was to record another ECG on the gold standard device and transmit that as well (in addition to the routine gold standard ECGs being sent prior to AF-BP monitor readings).

Reference	Wiesel, 2013[278]
Gold standard	Electrocardiographic event monitor (Heartrak 2)[assumed equivalent to Holter] was also provided to patients to obtain 60 s CG recordings before all the AF-BP readings. Patients transmitted the ECG read-outs to the monitoring centre daily. Readings reviewed by board-certified cardiologist, blinded to the results of the AF-BP monitor readings
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	Not simultaneous. ECG done first, a short time before BP measures.
Results	117 patients were fully compliant, with multiple readings taken daily on both index and gold standard devices. These had: Sensitivity 100, specificity 92.6; TP 8, FN 0, FP 8, TN 101 But this leads to best case results because non-compliant subjects excluded. Logistic regression analysis estimated: Sensitivity 100, specificity 90.4; TP 14, FN 0, FP 13, TN 112 There was a total of 3,316 days with AF-BP monitor readings and electrocardiographic readings. On the basis of the initial daily AF-BP monitor readings, the AF-BP monitor demonstrated sensitivity of 99.2% (93.7 to 100) and specificity of 92.9% (92.3 to 93.4) for detecting AF.
Source of funding	This study was funded by Microlife Corporation, Florida.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): None

Table 28 Wiesel, 2009[280]

Reference	Wiesel, 2009[280]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology outpatients
Country	USA

Reference	Wiesel, 2009[280]
Sample size	405
Sample characteristics	Mean age 73; male 51%; white 82%; black 8%; other 10%; CHF 6.7%; Hypertension 51.6%; DM 14.8%; CAD 37.3%
Inclusion criteria	Unselected general cardiology outpatients seen for scheduled visits in 2 cardiology centres in NY
Exclusion criteria	Pacemakers; defibrillators
Index test(s), including number of repetitions and duration	An oscillometric automatic blood pressure monitor (model BP3MQ1-2D; Microlife USA, Dunedin, FL) with an irregular heartbeat detection feature was modified such that the irregular heartbeat icon flashes when AF was detected. The device measures the last 10 pulse intervals during cuff deflation and calculates the mean and standard deviation of the intervals. An irregularity index is defined as the standard deviation divided by the mean of the time intervals. In order to reduce the effect of premature beats on the irregularity index, a cutoff value of 25% was chosen so that each of the ten pulse beat intervals that is 25% greater than or 25% less than the mean time interval is deleted. The remaining time intervals are used to calculate the irregularity index. If the irregularity index exceeds a threshold value of 0.06, the rhythm is considered irregular. The number of beats analyzed, and the irregularity index threshold value of 0.06 were chosen to maximize sensitivity for detecting AF. 3x readings taken by a trained technician. No interpretation as automated. For the three-sequential readings, the final reading was considered to be irregular if two or three of the individual readings were irregular. If none or only one of the three readings was irregular, the combined three-sequential reading was considered regular.
Gold standard	Standard 12 lead ECG taken by a trained technician, usually within 2 mins of the index test but at worst within the same 15 minute slot as the index test reading. Interpreted by a board certified cardiologist who was blinded to the index test results and other information.
Expertise of index test interpreter	NA as automated
Simultaneous index/gold vs non simultaneous	Not simultaneous – within a few minutes of each other
Results	<u>Single readings of microlife (n=3 x 405 readings)</u> Sensitivity: 95.3(92.8 to 97.6), Specificity 86.4 (84.3 to 88.7); TP 266, FN 13, FP 127, TN 809 <u>3 readings (majority rule) of microlife (n=405)</u> Sensitivity: 96.8(91 to 99), Specificity 88.8 (85 to 92); TP 90, FN 3, FP 35, TN 277
Source of funding	This study was supported by a grant from: Microlife USA, Inc., Dunedin, FL.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious

Reference	Wiesel, 2009[280]
	Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 29 Vaes, 2014[265]

Reference	Vaes, 2014[265]
Study type	Observational
Recruitment	Selective case/control
Setting	General Practices
Country	Belgium
Sample size	191
Sample characteristics	Age 74.2; male 52.4%; BMI 26.6; CHADSVASC 3; DM 21.5%; hypertension 81.7%; CAD 13.1%; TIA/CVA 11%; PAD 4.2%; AF 53.9%; warfarin 51.8%; DOACs 20.9%; antiplatelets 15.7%
Inclusion criteria	Participating general practitioners were asked to invite patients with known, paroxysmal or chronic atrial fibrillation to participate in the study. Furthermore, this convenience sample was added up with subjects without a history of atrial fibrillation.
Exclusion criteria	Pacemaker in active mode
Index test(s), including number of repetitions and duration	Each participant was tested with the MyDiagnostick (Applied Biomedical Systems BV, Maastricht, The Netherlands) by a single researcher who was not blinded for the medical history of the patient. This device has the form of a rod with a metal handle on both ends. In these handles electrodes make it possible to record a single-lead ECG that is analyzed automatically. The patient was asked to grasp the device by both handles. After one minute the ECG lead was analysed and LED indicators gave a red or green signal that could be interpreted as the presence or absence of atrial fibrillation. Three consecutive recordings with the MyDiagnostick with a 1 – 2 minute interval were done. The overall three measurements on the MyDiagnostick were viewed for each patient. A green light three times was interpreted as a negative result and a red light three times as a positive result. The non-uniform results of the MyDiagnostick were interpreted in favour of the most common outcome (i.e. 2x red and 1x green was interpreted as a positive result, while 1x red and 2x green was interpreted as a negative result. The method of detection of AF in the MyDiagnostick device is based on the measurement of R-R interval irregularity. Prior AF detection, the acquired ECG (1 minute) is pre-processed and R-waves are detected. From all detected R-wave annotations, R-R intervals are computed and used as an input for AF detection. The AF algorithm calculates an overall AF score based on a base rhythm-, periodicity- and variability score. The base rhythm score is based on a normal sinus rhythm state-machine chaining normal R-R intervals, including occasional premature intervals and short runs of tachycardia. Creation of long chains reflects a fit of the sinus rhythm state-model, lowering the probability of AF. The periodicity and variability scores are based on the R-R autocorrelation

Reference	Vaes, 2014[265]
	function. Periodicity of R-R interval patterns will generate multiple correlation peaks, whereas R-R interval irregularity will lower correlation at only a small shift. The overall AF score is obtained by linear combination of all scores and compared to a threshold, producing a dichotomous result (AF/no AF).
Gold standard	Afterwards a 12-lead electrocardiogram (ECG) (gold standard) was carried out once by the same researcher. The ECGs were done using digital machines and the data were immediately printed. The ECGs were analyzed off-line on the basis of the Minnesota Code Classification System for Electrocardiographic Findings by an experienced cardiologist, blinded for the software interpretation and the results from the MyDiagnostick.
Expertise of index test interpreter	NA as fully automated
Simultaneous index/gold vs non simultaneous	Not simultaneous
Results	TP 90, FN 6, FP 6, TN 79 Sensitivity 94% (87-98) Specificity 93% (85-97) Based on an expected prevalence of 6% in the population: PPV: 45% (24-68) NPV 99% (99-100)
Source of funding	No funding reported but equipment from industry
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious (population not that defined in protocol – not all people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF)

Table 30 Somerville, 2000[240]

Reference	Somerville, 2000[240]
Study type	Observational
Recruitment	Selective case/control

Reference	Somerville, 2000[240]
Setting	One GP surgery
Country	UK
Sample size	86
Sample characteristics	30% with AF; no other details provided
Inclusion criteria	The study patients were all recruited from a single practice. Patients aged 65 years or over with a diagnosis of atrial fibrillation were identified by searching computerised records using the Read Codes for atrial fibrillation and digoxin prescription. An equal number of patients aged 65 years or over, without either code in their computer records, was sampled. All patients were invited to attend the surgery by appointment.
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	One nurse (Nurse A) saw all the patients and had no prior knowledge of their medical history. Her background was in both community and accident and emergency (A&E) nursing, and she had experience of taking and interpreting electrocardiograms. • She palpated the pulse and recorded the result as 'regular' or 'irregular'. • She then recorded Bipolar ECGs, labelling them with an identifying number only. Bipolar ECGs depend on limb leads only, do not require removal of clothing, and therefore are a simpler, quicker procedure. • She also recorded 12 lead ECG (see gold standard below) At a later date bipolar and 12 lead ECG were interpreted independently by the nurse and one of the GP partners in the practice. They were unaware of the results of the pulse palpation. Other nurses with different previous experience of pulse palpation and ECG interpretation reviewed a random sample of the patients (this is why n for each person differs). Nurse B was a practice nurse with no additional ECG training. Nurse C was also a practice nurse but formerly worked on a coronary care unit and had been trained there to interpret ECGs.
Gold standard	The 12-lead electrocardiogram was taken by Nurse A, but interpreted by the consultant cardiologist.
Expertise of index test interpreter	Expertise at the tests not described.
Simultaneous index/gold vs non simultaneous	Unclear – not reported so assumption that it was not simultaneous
Results	Nurse A pulse: TP 26, FN 0, TN 46, FP 14; sensitivity 100(87-100); specificity 77(66-87) Nurse B pulse: TP 12, FN 1, TN 21, FP 4; sensitivity 92(64-100); specificity 84(64-96)

Reference	Somerville, 2000[240]
	Nurse A bipolar ECG: TP 24, FN 2, TN 53, FP 7; sensitivity 92(75-99); specificity 88(80-97) Nurse B bipolar ECG: TP 12, FN 1, TN 23, FP 2; sensitivity 92(64-100); specificity 92(74-99) Nurse C bipolar ECG: TP 13, FN 0, TN 35, FP 0; sensitivity 100(75-100); specificity 100(90-100) GP bipolar ECG: TP 25, FN 1, TN 59, FP 1; sensitivity 96(80-100); specificity 98(91-100) Nurse A 12 lead ECG: TP 25, FN 1, TN 56, FP 4; sensitivity 96(80-100); specificity 93(84-98) Nurse B 12 lead ECG: TP 13, FN 0, TN 19, FP 6; sensitivity 100(75-100); specificity 76(59-93) GP 12 lead ECG: TP 26, FN 0, TN 59, FP 1; sensitivity 100(87-100); specificity 98(91-100)
Source of funding	An initial pilot study was funded by a Small Projects Grant from the West MidlandsbRegional Health Authority. This led to the full study, which was supported by the North Staffordshire Health Authority.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious (population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF)

Table 31 Wiesel, 2004[281]

Reference	Wiesel, 2004[281]
Study type	Observational
Recruitment	consecutive
Setting	Outpatients followed by a cardiology practice
Country	USA
Sample size	450 people contributing to 464 office visits (14 attended twice)
Sample characteristics	59% men; mean age 69; most common associated medical conditions were hypertension, CAD and DM
Inclusion criteria	Unselected outpatients followed by an urban cardiology practice who had an ECG performed during scheduled office visits.
Exclusion criteria	None reported

Reference	Wiesel, 2004[281]
Index test(s), including number of repetitions and duration	Omron 712C automatic sphygmomanometer, modified to analyse the time interval between beats during deflation of the cuff. Irregularity index calculated via software on laptop and compared to threshold of 0.066. This test carried out twice (ideally) within 5 minutes of the 12 lead ECG. In total 446 paired readings were analysed
Gold standard	12 lead ECG performed during scheduled office visits. Expertise of interpreter unclear, though likely to be a cardiologist given that it was measured in a cardiology practice.
Expertise of index test interpreter	Not reported, though partially automated and defined by calculation rather than trace interpretation, so probably NA
Simultaneous index/gold vs non simultaneous	Not simultaneous – within 5 minutes
Results	Sensitivity 100%; Specificity 91%; TP 54, FN 0, FP 36 , TN 360
Source of funding	None reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 32 Mant, 2007[161]

Reference	Mant, 2007[161]
Study type	Observational
Recruitment	consecutive
Setting	25 General Practice surgeries in UK
Country	UK
Sample size	A random sample of 9866 people aged 65 or over was taken. A random half of these were invited for an ECG, and the remaining half were invited if opportunistic screening had previously identified them as having an irregular pulse. This led to 2595 12 lead ECGs being recorded, including 238 from opportunistic screening in 2001-3.
Sample characteristics	Patients taken from 25 General practices in central England. 1 GP and 1 practice nurse involved in the study. All practitioners had 1 hour training on AF detection.
Inclusion criteria	See above
Exclusion criteria	None reported

Reference	Mant, 2007[161]
Index test(s), including number of repetitions and duration	• 12 lead interpretive software • 12 lead interpreted by GP • Limb lead ECG interpreted by GP • Chest lead ECG interpreted by GP • 12 lead interpreted by practice nurse • Limb lead ECG interpreted by practice nurse • Chest lead ECG interpreted by practice nurse • 12 lead interpretive software combined with GP interpretation (positive if either or both is positive) All practitioners blinded to patients' identities, the diagnoses made by the specialists, and the diagnoses generated by the interpretative software
Gold standard	Two consultant cardiologists, blinded to the software interpretation and that of the primary care practitioners, read all the 12 lead electrocardiograms independently of each other. If the cardiologists disagreed, then a third consultant cardiologist arbitrated.
Expertise of index test interpreter	All nurses and GPs received one hour's training
Simultaneous index/gold vs non simultaneous	All readings taken simultaneously.
Results	The only ECGs taken were 12 lead ECGs. However a random third of the 2553 valid ECGs were printed out as single thoracic lead ECGs (the trace that would have been seen if only a single thoracic lead had been used) and a random third as limb lead ECGs (ditto). The other third printed out in full as 12 lead ECGs. These were then assembled into 25 batches of 100 ECGs, comprising a third each of 12 lead, thoracic lead and limb lead traces. These were then sent to 49 practices (one dropped out) one unique batch being duplicated and distributed to 2 practices. These results below denote the accuracy of the different personnel/ECG traces relative to gold standard of cardiologist 12 lead ECG. For uncertain results these have been taken as no AF (this is what authors of paper did). <u>12 lead interpretive software (Biolog interpretive software)</u> Sensitivity: 83.3(78.3-88.2) Specificity:99.1(98.7-99.5) TP 179, FN 36, FP 21, TN 2320 <u>12 lead interpreted by GP</u> Sensitivity:79.8(70.5-87.2) Specificity:91.6(90.1-93.1) TP 79, FN 22, FP 114, TN 1241

Reference	Mant, 2007[161]
	Limb lead ECG interpreted by GP Sensitivity:82.5(74.8-88.7) Specificity:88.5(84.6-88.3) TP 104, FN 22, FP 156, TN 1202 Chest lead ECG interpreted by GP Sensitivity:84.8(78.7-91) Specificity:86.4(84.6-88.3) TP 112, FN 20, FP 180, TN 1145 12 lead interpreted by practice nurse Sensitivity:77.1(67.4-85) Specificity:85.1(83-86.9) TP 74, FN 22, FP 198, TN 1132 Limb lead ECG interpreted by practice nurse Sensitivity:72.0(63.9-80.1) Specificity:83.4(81.4-85.4) TP 85, FN 33, FP 220, TN 1107 Chest lead ECG interpreted by practice nurse Sensitivity:68.7(60.1-76.4) Specificity:82.8(80.7-84.8) TP 92, FN 42, FP 22, TN 1066 12 lead interpretive software combined with GP interpretation (positive if either or both is positive) Sensitivity:91.9(86.6-97.3) Specificity:91.1(89.6-92.6) TP 91, FN 8, FP 121, TN 1234
Source of funding	The work was funded by the Health Technology Assessment Programme. The authors are independent from the funders of the research. The views expressed in this publication are those of the authors and not necessarily those of the funders or the Department of Health.
Limitations	Risk of bias (QUADAS 2 – risk of bias): No Serious risk Indirectness (QUADAS 2 - applicability): none

Table 33 Lown, 2018[156]

Reference	Lown, 2018[156]
Study type	Observational
Recruitment	Selective case/control
Setting	3 General Practices in the UK
Country	UK
Sample size	418
Sample characteristics	Individuals from 3 general practices aged >65 both with and without a coded diagnosis of AF in their medical records were invited to attend a Single screening visit at their local general practice. Mean age 73.9; 79 found to have AF
Inclusion criteria	>=65; from the 3 designated general practices
Exclusion criteria	Participants were excluded if they, had a pacemaker, were deemed unsuitable by their named General Practitioner (GP) (e.g., terminally ill and bedridden), lacked capacity, or had a previous moderate or severe skin reaction to electrode gel.
Index test(s), including number of repetitions and duration	Participants were screened for AF by study nurses using 4 devices (WatchBP, AliveCor, PH7, and BG2) in a random sequence. WatchBP detects pulse intervals (during 3 consecutive blood pressure [BP] measurement cycles) and uses an algorithm to indicate AF via an AFicon on the display. AliveCor senses limb-lead ECG data when the participant's thumbs are placed on electrodes. It can detect AF during a single measurement period. An accompanying application displays the corresponding ECG trace and subsequent diagnostic algorithm result. The AliveCor algorithm used in the trial (Kardia version 4.7.0) produces 4 results: suspected AF, normal, unreadable, and unclassified (if the ECG was not classified in the previous categories with a normal heart rate). Normal and unclassified results were thus inferred as non-AF results and unreadable recordings as no result. PH7 can detect AF during a single measurement period. The results for PH7 are displayed immediately after the measurement period on the screen of the tablet running the corresponding application. The Polar HY (PH7) can also detect AF during a single measurement period. It is a commercially available heart rate sensor used by recreational and professional athletes. Firstbeat Bodyguard 2 (BG2) is a reliable R-R interval recording device. The results for the BG2 device were calculated off-line. RNs blinded to gold standard results

Reference	Lown, 2018[156]
Gold standard	12 lead ECG interpreted by 2 cardiologists, with a third cardiologist adjudicating disagreements. ECG done in same session but not reported to be at the same exact time as the other tests. Blinded to index test results
Expertise of index test interpreter	NA as automated for AliveCor, WatchBP and PH7. Unclear how BG2 was interpreted.
Simultaneous index/gold vs non simultaneous	Unclear, but unlikely
Results	Alive Cor Sensitivity 87.8(78.71-93.99); Specificity 98.81(96.98-99.67); TP 72, FN 10; TN 332; FP 4 Watch BP Sensitivity 96.34(89.68-99.24); Specificity 93.45(90.25-95.85); TP 79, FN 3; TN 314; FP 22 PH7 Sensitivity 96.34(89.68-99.24); Specificity 98.21(96.17-99.34); TP 79, FN 3; TN 330; FP 6 BG2 Sensitivity 96.34(89.68-99.24); Specificity 98.51(95-99.52); TP 79, FN 3; TN 331; FP 5
Source of funding	This paper presents independent research funded by the National Institute of Health Research School for Primary Care Research (NIHRSPCR) FR11:ProjectNo:318.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 34 Yan, 2018 [288]

Reference	Yan, 2018[288]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology inpatients
Country	Hong Kong

Reference	Yan, 2018[288]
Sample size	233
Sample characteristics	Mean age 70.3; 71.4% men; AF present in 34.6% at time of study; BMI 24.6; CHADSVASC 3.6; history of AF 53.9%; DM 35%; vascular disease 50.7%; TIA/stroke 18.9%; CHF 31.8%; pacemaker 3.2%; hypertension 5.9%; no antithrombotic treatment 51.2%; DOACS 13.4%; VKAs 15.7%
Inclusion criteria	Patients admitted to the cardiology ward of the hospital for clinical reasons
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	Two iPhone 6S units (Apple Inc, Cupertino, CA) installed with the Cardiio Rhythm application were used for simultaneous facial and fingertip photoplethysmographic detection. Cardiio Rhythm application is a novel smartphone application that measures the rhythm of the heart through recording pulsatile photoplethysmographic signal from either the finger-tip or the face without physical contact. The camera detects subtle beat-to-beat variations of skin colour on the basis of the amount of reflected light that changes, according to the arterial blood volume pulsations. Photoplethysmographic waveforms were sampled at 30 Hz, and each measurement recorded 512 samples (17 seconds). Detection of AF was based on an irregularly irregular pattern in the photoplethysmographic waveform attributable to AF. Three successive 20-second (total, 60 seconds) recordings were acquired per patient and analyzed for heart rate regularity by Cardiio Rhythm (Cardiio Inc, Cambridge, MA) smartphone application. Pulse irregularity in ≥1 photoplethysmographic readings or 3 uninterpretable photoplethysmographic readings were considered a positive AF screening result.
Gold standard	12 lead ECG was performed after the photoplethysmographic measurements. Interpreted by a cardiologist blinded to index test results
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	Not simultaneous, though in same session
Results	Fingertip plethysmography Sensitivity 94.7(87.1-97.9); Specificity 93(87.5-96.1); TP 71, FN 4, TN 132, FP 10 Facial photoplethysmography Sensitivity 94.7(87.1-97.9); Specificity 95.8(91.1-98.1); TP 71, FN 4; TN 136, FP 6
Source of funding	Hong Kong Research Grants Council—General Research Fund (reference no. 14118314). Cardiio Inc provided the iPhones for study purposes.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious

Reference	Yan, 2018[288]
	Indirectness (QUADAS 2 - applicability): Serious (population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF)

Table 35 Tieleman, 2014[258]

Reference	Tieleman, 2014[258]
Study type	Observational
Recruitment	consecutive
Setting	Outpatients/GP practice
Country	Netherlands
Sample size	Part 1: 192, part 2: 676
Sample characteristics	Part 1:Age 69.4 years; 48.4% male Part 2: Age 74 years
Inclusion criteria	Part 1: Patients visiting the outpatient cardiology clinic Part 2: Patients attending 2 GP clinics for influenza vaccination
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	The MyDiagnostick (www.mydiagnostick.com, MyDiagnostick Medical BV) is intended to discriminate AF from a normal cardiac rhythm (normal sinus rhythm, NSR) based on the ECG. This is achieved by an easy accessible device that can be used by both care providers like general practitioners, nurses, cardiologists and patients. The device has the shape of a stick (length 26 cm, diameter 2 cm) with metallic electrodes at both ends as shown in Figure 1. The MyDiagnostick does not depend on any infrastructure or communication channels and can be used anytime, anywhere by simply holding the device in both hands for 60 s until the result is revealed. While holding the device, it will flash on the rhythm of the detected heartbeat. After 1 min, the MyDiagnostick either turns green, indicating a normal cardiac rhythm, or red in the case of AF. The algorithm is designed in such a way that it will diagnose AF in case the arrhythmia is present during at least 75% (45 s) of the1 min ECG recording. The MyDiagnostick will store up to 140 1 min ECG Lead I strips. A priority storage scheme is implemented in the MyDiagnostick aiming at storage of the most recent AF episodes. When more than 140 recordings are made, only the non-AF ECGs are overwritten, unless all non-AF strips are replaced by AF recordings. This allows for long-term autonomous use of the device without the burden of losing relevant ECG data. MyDiagnostik held for 1 minute by the patient.
Gold standard	12 lead ECG, performed immediately after index test. Assessed by a cardiologist blinded for the MyDiagnostik AF outcome.
Expertise of index test interpreter	NA as fully automated

Reference	Tieleman, 2014[258]
Simultaneous index/gold vs non simultaneous	No, but ECG followed immediately after index test
Results	Part 1: Sensitivity: 100 (93-100); Specificity 95.9 (91.3-98.1); TP 53, FN 0, FP 6, TN 133 Part 2: Sensitivity: 100; specificity 99; TP 55, FN 0, FP 6, TN 615 Combined (not in paper but no reason why not): Sensitivity 100, specificity 98.4; TP 108, FN 0, FP 12, TN 748
Source of funding	The work was supported by MyDiagnostick Medical BV, Maastricht, The Netherlands. Funding to pay the Open Access publication charges for this article was provided by MyDiagnostick Medical BV, Maastricht, The Netherlands.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): serious (includes healthy population)

Table 36 Stergiou, 2009[243]

Reference	Stergiou, 2009[243]
Study type	Observational
Recruitment	Selective case/control
Setting	Outpatients hypertension clinic
Country	Greece
Sample size	73
Sample characteristics	Age 70.5; 65.8% male; BMI 27; smokers 5.5%; CVD 39.7%; DM 15.1%; hypertension 63%; systolic bp 138; diastolic bp 80; AF 37%
Inclusion criteria	Subjects with known sustained AF, or other non-AF arrhythmias, and controls with sinus rhythm were recruited among those attending an Outpatients Hypertension Clinic, patients admitted in a University Department of Medicine wards and healthy volunteers.
Exclusion criteria	Exclusion criteria were age <35 years, presence of a pacemaker, and/or an implanted defibrillator and refusal to participate.

Reference	Stergiou, 2009[243]
Index test(s), including number of repetitions and duration	An automated oscillometric device for self-home BP monitoring, which has been validated earlier for BP measurement accuracy, and an additional function, which allows AF detection during routine BP measurement, has been developed (Microlife BPA100 Plus, Microlife, Heerbrugg, Switzerland). Atrial fibrillation is detected during the usual BP recording by the application of an in-built algorithm, which analyses the irregularity of the pulse rate. The average time interval of the last 10 beats, during deflation, is calculated and intervals that are 25% shorter or longer than that of the average are discarded. The mean of the remaining intervals is calculated with its s.d., and an AF diagnosis is made, if the s.d. per mean ratio is >0.06. Four devices were donated by the manufacturer for carrying out this study. 3 measures of BP were taken from each person (with at least 5 mins rest in the lying position and with at least 30s between measurements), and the accuracy of 1,2 and 3 measurements was taken.
Gold standard	12 lead ECG, interpreted by one of the study authors and an expert cardiologist.
Expertise of index test interpreter	NA as automated
Simultaneous index/gold vs non simultaneous	Yes, the ECG was recorded during the deflation phase of each BP measurement, which is when the AF detector in the BP device works.
Results	<u>Using just the first reading per patient (thus modelling the accuracy if just one BP measure is done):</u> Sensitivity:0.93 (0.74-0.99); specificity 0.89(0.76-0.96); TP 25, FN 2, FP 5, TN 40 <u>Using just the first 2 readings per patient (thus modelling the accuracy if just 2 BP measures are done) [AF diagnosis if just one is positive]:</u> Sensitivity:1.00 (0.84-1); specificity 0.76(0.60-0.87); TP 27, FN 0, FP 11, TN 34 <u>Using all 3 readings per patient (thus modelling the accuracy if 3 BP measures are done) [AF diagnosis if just one is positive]:</u> Sensitivity:1.00 (0.84-1); specificity 0.69(0.53-0.81); TP 27, FN 0, FP 14, TN 31 <u>Using all 3 readings per patient (thus modelling the accuracy if 3 BP measures are done) [AF diagnosis if 2/3 are positive – 'MAJORITY RULE']:</u> Sensitivity:1.00 (0.84-1); specificity 0.89(0.75-0.96); TP 27, FN 0, FP 5, TN 40
Source of funding	This study was funded by the Hypertension Center, Third University Department of Medicine, Athens.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 37 Bumgarner, 2018[26]

Reference	Bumgarner, 2018[26]
Study type	Observational case control.
Recruitment	Selective case/control
Setting	Patients attending for cardioversion
Country	USA
Sample size	100
Sample characteristics	Age 68.2; female 17%; warfarin 32%; DOACs 68%; CV performed 85%
Inclusion criteria	Consecutive patients with a diagnosis of AF who presented for scheduled elective CV with or without a planned transesophageal echo-cardiogram were screened for enrolment. Inclusion criteria included all adult patients age 18 to 90 years who were able to provide informed consent and willing to wear the KB before and after cardioversion
Exclusion criteria	Implanted pacemaker; defibrillator
Index test(s), including number of repetitions and duration	In November 2017, the Kardia Band (KB) (AliveCor) was introduced as the first U.S Apple Watch accessory that allows a patient to record a rhythm strip equivalent to lead I for 30 s. The KB is coupled with an application that provides an instantaneous and automatic rhythm adjudication algorithm for the diagnosis of AF. This algorithm measures rhythm irregularity and P-wave absence in real time to classify the rhythm strip as “possible AF.” If the criteria for AF is not met, the KB algorithm classifies regular rhythms with P waves as “normal” if the rate is between 50 and 100 beats/min or “unclassified” for those rhythms with rates <50 or >100 beats/min or if the recording is noisy or shorter than 30 s. The application can inform the patient when AF is detected and transmit these results to the physician instantaneously. If a cardioversion was performed (done in 85% of participants) then another ECG and KB recording was made. Automated readings and physician-reviewed readings both evaluated.
Gold standard	12 lead ECG, interpreted by 2 blinded electrophysiologists, with a third electrophysiologist used for adjudication if there was no agreement.
Expertise of index test interpreter	Automated so NA. But also interpreted by 2 blinded electrophysiologists, with a third electrophysiologist used for adjudication if there was no agreement.
Simultaneous index/gold vs non simultaneous	Author states they considered it simultaneous, but the ECG reading preceded the KB recording

Reference	Bumgarner, 2018[26]
Results	KB algorithm automatic reading (this is the most relevant as this will be the most likely way it is used clinically) *Ignoring missing values:* Sensitivity 93(86-99); Specificity 84(73-95); TP 63, FN 5, TN 37, FP 7 *Designating unclear vales as –ve readings:* Sensitivity 69.2; specificity 91.0; TP 63, FN 28, TN 71, FP 7 KB algorithm reading interpreted by electrophysiologists *Ignoring missing values:* Sensitivity 99(96-100); Specificity 83(74-92); TP 80, FN 1, TN 55, FP 11 *Designating unclear vales as –ve readings:* Sensitivity 87.9; specificity 85.9; TP 80, FN 11, TN 67, FP 11
Source of funding	AliveCor provided the Kardia Band monitors that were connected to an Apple Watch and paired via Bluetooth to a smartphone device for utilization in the study. AliveCor was not involved in the design, implementation, data analysis, or manuscript preparation of the study.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 38 Caldwell, 2012[29]

Reference	Caldwell, 2012[29]
Study type	Case-control observational
Recruitment	Selective case/control
Setting	Anticoagulation outpatient clinic
Country	UK
Sample size	157
Sample characteristics	Not reported

Reference	Caldwell, 2012[29]
Inclusion criteria	Consecutive patients with chronic AF attending the anticoagulation clinic, and consecutive patients with no prior diagnosis of AF attending for a routine ECG
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	• 5s 6 lead ECG from conventionally positioned limb electrodes (4 limb-leads) • 5s Supine 4-electrode 6-lead frontal plane ECG recording in supine using the prototype recorder placed on the lower thorax/abdomen • 5s Seated 4-electrode 6-lead frontal plane ECG prototype recording with loosened clothing only interpreted by 1 semi-blinded (Observer A) and 2 blinded cardiologists (observers B and C) Prototype recorder had 4 copper electrodes mounted on a plastic frame, and colour-coded to represent the four ECG frontal-plane limb electrodes. The red right arm ECG electrode cable was connected to the right upper prototype electrode, the yellow to the left upper, the green to the left lower, and the black to the right lower. The upper and lower electrodes were mounted 8 cm apart, and the upper pair and lower pair were 16 cm apart.
Gold standard	Conventional 10 second 12 lead ECG, in supine undressed position, interpreted by 2 blinded and 1 semi-blinded cardiologists. Where there was a disagreement between observers, the 'gold standard' result was assumed to be the most prevalent response from the three observers.
Expertise of index test interpreter	Consultant cardiologists
Simultaneous index/gold vs non simultaneous	Not simultaneous – but all done within the same session.
Results	<u>5s 6 lead ECG from conventionally positioned limb electrodes (4 limb-leads)</u> Observer 1: sensitivity 0.97(0.91-1); specificity 1.0(0.95-1); TP 76, FN 2, FP 0, TN 79 Observer 2:sensitivity: 0.94(0.86-0.98); specificity 0.97(0.91-1); TP 73, FN 5, FP 2, TN 77 Observer 3: sensitivity: 0.99(0.93-1); specificity 0.94(0.86-0.98); TP 77, FN 1, FP 5, TN 74 <u>5s Supine 4-electrode 6-lead frontal plane ECG recording in supine using the prototype recorder placed on the lower thorax/abdomen</u> Observer 1: sensitivity 0.97(0.91-1); specificity 1.0(0.95-1) ; TP 76, FN 2, FP 0, TN 79 Observer 2:sensitivity: 0.94(0.86-0.98); specificity 0.96(0.89-0.99) ; TP 73, FN 5, FP 3, TN 76 Observer 3: sensitivity: 0.96(0.86-0.99); specificity 0.95(0.88-0.99) ; TP 75, FN 3, FP 4, TN 75 <u>5s Seated 4-electrode 6-lead frontal plane ECG prototype recording with loosened clothing only</u> Observer 1: sensitivity 0.97(0.91-1); specificity 1.0(0.95-1) ; TP 76, FN 2, FP 0, TN 79

Reference	Caldwell, 2012[29]
	Observer 2:sensitivity: 0.90(0.81-0.95); specificity 0.96(0.89-0.99) ; TP 70, FN 8, FP 3, TN 76 Observer 3: sensitivity: 0.97(0.91-1); specificity 0.96(0.89-0.99) ; TP 76, FN 2, FP 3, TN 76
Source of funding	This work has been funded by a TrusTECH Pathfinder Proof of Concept Grant.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 39 Fan, 2019[77]

Reference	Fan, 2019[77]
Study type	Observational
Recruitment	consecutive
Setting	General Hospital
Country	China
Sample size	112
Sample characteristics	Mean age 58; female 46%; BMI 24.44; HF 4%; hypertension 52%; DM 27%; stroke/TIA/SE 7%; CAD 45%; vascular disease 55%; COPD 2%; renal dysfunction 4%; hepatic dysfunction 0%; sleep apnea 4%; hyperthyroidism 2%; current smoking 29%; median CHADSVASC 2; median HAS-BLED 1; OAC 18%; antiplatelets 27%;
Inclusion criteria	Aged 18 or over
Exclusion criteria	Patients unable to use mobile phones and smart bands, with mental or memory problems, or with a pacemaker or implantable cardioverter defibrillator.
Index test(s), including number of repetitions and duration	Huawei mate 9 mobile phone – for 3 minutes Huawei Honor 7x mobile phone – for 3 minutes Smart band – Huawei band 2 – for 3 minutes Participants were simultaneously tested with mobile phones (HUAWEI Mate 9, HUAWEI Honor 7X), smart bands (HUAWEI Band 2), and 12-lead ECG for 3 minutes. Participants were advised to lie down in a supine position and breathe spontaneously. A HUAWEI Mate 9 (mobile phone 1) was positioned on the left-hand finger (either the index or middle finger) with the camera lens and LED light placed on the fingertip of the participant. Similarly, a HUAWEI Honor 7X (mobile phone 2) was positioned on the finger of the right hand. PPG measurements were performed by using the Heartbeats mobile phone app. Pulse waveform recordings were performed by the participants under the supervision of trained study personnel. A dedicated data collection app, Heartbeats (Preventicus

Reference	Fan, 2019[77]
	GmbH, Jena, Germany), was responsible for the pulse waveform signal acquisition and was installed in the HUAWEI mobile phones. Then all 3-minute pulse waveform recordings using the smart devices were uploaded to the online cloud center and analyzed by a realizable algorithm (PRO AF PPG) provided by Preventicus (Preventicus GmbH, Jena, Germany).
Gold standard	12 lead ECG, for 3 minutes. Interpreted by 2 independent cardiologists blinded to the baseline information of participants
Expertise of index test interpreter	Unclear but appears that the algorithm used in the index devices (PRO AF PFG) was automated
Simultaneous index/gold vs non simultaneous	Yes
Results	Does not appear to be analysed by person but instead by segments of trace. The paper states that 1 minute sections were used but insufficient other information given for mobile phones. Thus raw data not possible to calculate for mobile phones. For smart phones stated that 280 AF traces and 334 SR traces on ECG, so possible to calculate raw values. Huawei mate 9 mobile phone – sensitivity 94.4 (88.9-97.4); specificity 100 (97.2-100); raw data not calculable Huawei Honor 7x mobile phone – sensitivity 95.6 (90.2-98.2); specificity 99.4 (96.2-100); raw data not calculable Smart band – sensitivity 95.4 (92-97.4); specificity 99.7 (98.1-100) TP 267, FN 23, FP 1, TN 333
Source of funding	This research project was funded by the Chinese PLA Healthcare Foundation (17BJ208) and National Natural Science Foundation of China (H2501). HUAWEI (Huawei Technologies Co, Ltd, Shenzhen, China) provided the mobile phones (Mate 9, Honor 7X) and smart bands (Band 2) for study purposes. Preventicus (Preventicus GmbH, Jena, Germany) provided the Heartbeats mobile phone app and the PRO AF PPG algorithm.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 40 Arevalo-Manso, 2016[7]

Reference	Arevalo-Manso, 2016[7]
Study type	Consecutive, observational
Recruitment	consecutive

Reference	Arevalo-Manso, 2016[7]
Setting	Stroke Unit
Country	Spain
Sample size	76
Sample characteristics	Patients referred to a stroke centre which provides expertise to a population of about one million people, and has a dedicated SU with continuous bedside ECG monitoring for six patients. Patients are admitted to the SU from the emergency room or the TIA clinic within the first 48 h from the onset of symptoms and remain there for at least 24 h. During their stay in the SU, patients are continuously evaluated by the same specialised stroke team and a nurse continually assesses the patient's ECG, vital signs, and neurological state. After the acute phase, patients are transferred from the SU to the neurology ward until discharge or transfer to a rehabilitation centre or a care facility. There were two samples in this study. "Study" group (n=17) were age 72.6; 47.1% men; 70.6% hypertension; 35.3% DM; 64.7% dyslipidaemia; 23.5% smokers; 35.3% CAD; 11.8% PAD; 0% TIA; 100% brain infarction; antiplatelets 52.9%; OACs 5.9%. These were assigned to one bed in the SU that was equipped with the AF-RS monitor "Control" group (n=59) were 71.9 yrs; 62.7% men; 69.55 hypertension; 25.4% DM; 61% dyslipidaemia; 20.3% smokers; 15.3% CAD; 5.1% PAD; 11.9% TIA; 88.1% brain infarction; antiplatelets 39%; OACs 3.4%. These were assigned to 5 beds in the SU that were equipped with a standard monitor Patients assigned non-randomly to these groups on basis of availability of the bed and the criteria of the neurologists on call, who were unaware of the study.
Inclusion criteria	Age>18 years and having been admitted to the SU for an acute TIA or ischaemic stroke.
Exclusion criteria	History of AF
Index test(s), including number of repetitions and duration	1. Study group only: From November 2011–January 2012, a monitor equipped with **AF-RS** (DASH 5000, General Electric Healthcare, Milwaukee, Wisconsin, USA) was temporarily assigned by the manufacturer's technical service to our SU, replacing another damaged unit. This monitor included the GE-EK Pro arrhythmia algorithm v.11 (General Electric Healthcare), which uses **four simultaneous leads** for analysis and sounds a specific alarm when an AF event is detected. When the AF device sounded an alarm, the ECG traces were automatically saved to digital memory and were subsequently examined by a neurologist. In addition, following the AF alarm, the nurse conducted a 12-lead ECG for confirmation. Median duration 2 days 2. During the study period, the other five beds in the SU were equipped with standard ECG monitoring devices without AF-RS. Three of the devices were from the same manufacturer as the new device (DASH 2500, General Electric Healthcare). The two remaining monitors belonged to another manufacturer (Mod. 90369, Spacelabs Healthcare, Issaquah, Washington, USA). The five standard monitors included the following set of automatic alarm signals: (a) ventricular fibrillation; (b) upper and lower heart rate limits (usually set to 120 and 50 beats per min, respectively); and (c) cardiac asystole. When the SU nurse suspected AF from the ECG traces on the monitor display, the nurse took a 12-lead ECG for confirmation, which was subsequently reviewed by the neurologist on call. Median duration 2 days

Reference	Arevalo-Manso, 2016[7]
Gold standard	A 12-lead ECG is performed upon admission to the emergency room; a daily 12-lead ECG (Page Writer 100, Hewlett Packard, Palo Alto, California, USA) is performed on all patients during their stay in the SU, and another 12-lead ECG is performed if AF is suspected; a 24 h Holter ECG is scheduled for selected patients when AF has not previously been identified by another method. The definitive (gold standard) AF diagnoses were established by the neurologist/cardiologist based on the results of the 12-lead ECG and the 24 h Holter ECG. AF was defined as absolutely irregular intervals between two R waves, in the absence of P-waves or in the presence of fibrillatory waves with an atrial cycle length variable and <200 ms, lasting at least 30 s.
Expertise of index test interpreter	Throughout the study, the observation of the ECG monitoring was performed by the same nurses who all had received the same standardised training in the detection of AF and other alterations in cardiac rhythm.
Simultaneous index/gold vs non simultaneous	No
Results	**AF-RS monitor** Sensitivity 57.1(25-84.2); Specificity 100(72.2-100); TP 4, FN 3; FP 0, TN 10 **Standard monitor** Sensitivity 7.7(1.4-33.3); Specificity 100(92.3-100); TP 1, FN 12; FP 0, TN 46
Source of funding	IdiPAZ Health Research Institute.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 41 Desteghe, 2017[58]

Reference	Desteghe, 2017[58]
Study type	Observational
Recruitment	consecutive
Setting	Cardiac inpatients
Country	Belgium
Sample size	344
Sample characteristics	Patients admitted to cardiac wards in a tertiary hospital in Belgium. Patients with an implanted device comprised 17.2% of the cardiology population: 60% was actively paced, 7.3% was intermittently paced, and 32.7% was not being paced during the

Reference	Desteghe, 2017[58]
	recordings. Based on chart review, 35.6% of the screened study population was known with AF. At the moment of the study, 11.9% showed AF on their 12-lead ECG. Of the entire AF population, the majority had paroxysmal AF (54.4%) while those in AF at the time of screening were mostly permanently in AF.
Inclusion criteria	Patients admitted to cardiac wards in a tertiary hospital in Belgium; able to give informed consent
Exclusion criteria	Age <18 years, patients in isolation, and those who were unable to hold both devices properly.
Index test(s), including number of repetitions and duration	Each patient was asked by a single researcher to consecutively hold two handheld ECG devices: the MyDiagnostick (Applied Biomedical Systems BV, The Netherlands) and the AliveCor (AliveCor Inc., USA). To record a single-lead ECG with the MyDiagnostick, the patient has to hold the rod-like device with both hands for 1 min. For this study, the device was programmed in screening mode, meaning that all ECG recordings are stored together with a recording time, date, and automated algorithm diagnosis. During the screening, the recording time and the patient's identification data were noted by the operator. After a screening session, the ECG recordings were uploaded to a computer and linked to the patients' identification by means of the accompanying software. The algorithm of the MyDiagnostick will indicate AF based on an irregular R–R interval which is present during at least 75% of the 1-min recording. The AliveCor is coupled with an iPhone and allows a noise-filtered lead I ECG recording by means of the corresponding AliveECG app. After each 30 s recording, identification data are directly entered and stored in the app. Together with the automated rhythm diagnosis, these data are wirelessly transferred to a web-based software platform. The automated algorithm of the AliveCor is based on the criteria of P-wave absence and R–R interval irregularity to diagnose AF.
Gold standard	At the cardiology department, a full 10-s 12-lead ECG recording was performed by a trained nurse immediately before recording with the two handheld devices. **At the department of geriatrics, a 6-lead limb ECG was taken (30 s duration), so these results are not reported below**. Every recording was later reviewed randomly and independently by two electrophysiologists who were blinded for the automated analysis of the devices.
Expertise of index test interpreter	Automatic detection by algorithm. But there was also manual detection of the traces from both index tests by the same 2 electrophysicists who interpreted the 12 lead ECG
Simultaneous index/gold vs non simultaneous	No – 12 lead done immediately before index tests
Results	**Cardiology (ref standard 12 lead)** **My Diagnostik** Automated with implanted device [PM/ICD] patients included Sensitivity: 60.5% Specificity: 93.3%

Reference	Desteghe, 2017[58]
	(TP 23, FN 15, FP 19, TN 263) Automated with PM/ICD patients excluded Sensitivity: 81.8% Specificity: 94.2% (TP 18, FN 4, FP 14, TN 229) Electrophysiologist 1 with PM/ICD patients included Sensitivity: 68.4% Specificity: 91.1% (TP 26, FN 8, FP 16, TN 257) 13 illegible – these are taken into account when calculating accuracy Electrophysiologist 1 with PM/ICD patients excluded Sensitivity: 77.3% Specificity: 93% (TP 17, FN 3, FP 11, TN 226). 8 illegible – these are taken into account when calculating accuracy Electrophysiologist 2 with PM/ICD patients included Sensitivity: 55.3% Specificity: 94.3% (TP 21, FN 14, FP 7, TN 266). 12 illegible – these are taken into account when calculating accuracy Electrophysiologist 2 with PM/ICD patients excluded Sensitivity: 72.7% Specificity: 95.9% (TP 16, FN 4, FP 4, TN 233). 8 illegible – these are taken into account when calculating accuracy **AliveCor** Automated with PM/ICD patients included Sensitivity: 36.8% Specificity: 96.1% (TP 14, FN 24, FP 11, TN 271) Automated with PM/ICD patients excluded Sensitivity: 54.5% Specificity: 97.5% (TP 12, FN 10, FP 6, TN 237)

Reference	Desteghe, 2017[58]
	Electrophysiologist 1 with PM/ICD patients included Sensitivity: 68.4% Specificity: 92.6% (TP 26, FN 8, FP 8, TN 261) 17 illegible – these are taken into account when calculating accuracy Electrophysiologist 1 with PM/ICD patients excluded Sensitivity: 90.9% Specificity: 94.7% (TP 20, FN 0, FP 5, TN 230) 10 illegible – these are taken into account when calculating accuracy Electrophysiologist 2 with PM/ICD patients included Sensitivity: 63.2% Specificity: 95.7% (TP 24, FN 14, FP 4, TN 270) 8 illegible – these are taken into account when calculating accuracy Electrophysiologist 2 with PM/ICD patients excluded Sensitivity: 90.9% Specificity: 96.3% (TP 20, FN 2, FP 3, TN 234) 6 illegible – these are taken into account when calculating accuracy
Source of funding	This study is part of the Limburg Clinical Research Program (LCRP) UHasselt-ZOL-Jessa, supported by the foundation Limburg Sterk Merk, province of Limburg, Flemish government, Hasselt University, Ziekenhuis Oost-Limburg, and Jessa Hospital. Applied Biomedical Systems BV and AliveCor, Inc., provided the devices for this study for free but were not involved in any aspect of the trial.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 42 Haverkamp, 2019[96]

Reference	Haverkamp, 2019[96]
Study type	Observational
Recruitment	consecutive

Reference	Haverkamp, 2019[96]
Setting	Cardiac inpatients
Country	Norway
Sample size	94
Sample characteristics	37% female; mean age 58;
Inclusion criteria	People having ongoing scECG cardiac surveillance who were admitted to the cardiac ward at a university hospital.
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	ECG Check, an FDA-approved mobile heart monitor manufactured by Cardiac Designs. By putting two fingers on the ECG Check, it registers a 30-s, one-lead ECG and stores it on a device (smart-phone, tablet) via Bluetooth. The application's algorithm then classifies the spECGs as either “Normal” or “Abnormal”, and it also estimates the frequency using the RR interval. The participants performed the recording as independently as possible, supervised by study investigators and with assistance if needed.
Gold standard	Standard 12 lead ECG. Shortly after acquiring the index ECG, 12 lead ECG reports were extracted for comparison. However unclear when the 12 lead ECG was actually recorded. Expertise of 12 lead ECG interpreters not described.
Expertise of index test interpreter	The subjects were given basic instructions on how to use the index ECG device and send the result to an email address created for the purpose.
Simultaneous index/gold vs non simultaneous	Unclear – seems very unlikely
Results	Sensitivity 100%, specificity 94%; TP 11, FN 0, FP 5, TN 78
Source of funding	Reported no funding from any source
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 43 McManus, 2016[164]

Reference	McManus, 2016[164]
Study type	Observational

Reference	McManus, 2016[164]
Recruitment	People before and after a cardioversion – thus very much a case-control situation
Settings	Cardiac inpatients
Country	USA
Sample size	128
Sample characteristics	Age 66.2yrs ; non-white 7%; 18% women; hypertension 75.7%; DM 28.2%; CAD 25%; CHF 32.8%; stroke 13.3%
Inclusion criteria	The original PULSESMART cohort included 76 participants with AF scheduled to undergo elective cardioversion at the University of Massachusetts Medical Center (UMMC). For the present study, the sample were enriched with an additional 55 participants (22 adults with AF, 15 with PACs, and 15 with PVCs) to create a cohort comprised of a more representative array of benign (PAC and PVC) and malignant (AF) causes of an irregular pulse. Patients with frequent PACs or PVCs were identified from a roster of inpatients on the cardiac telemetry unit at the UMMC. Study staff performed a review of hospital telemetry recordings on a daily basis to identify patients with frequent ectopy.
Exclusion criteria	Not reported
Index test(s), including number of repetitions and duration	Original study participants had 2 minute pulse waveforms recorded before and after elective cardioversion by study staff using a labelled study iPhone 4S. Participants were asked to hold the iPhone 4S in their hand, with their right first or second finger on the standard camera and lamp for 2 minutes, during which time the pulse waveform was recorded. Pulse recordings were obtained with patients in the supine position. A video of user's fingertip blood flow intensity at 640×480 pixel resolution was sampled at a rate of 30 frames/sec for 2 minutes. An average of the intensity values from the green band from the RGB video is analyzed. All iPhone pulse recordings were downloaded using a de-identified study number to enable post-processing and analysis, using threshold values of RMSDD .1093, ShE=0.4890, Poincare plot=0.2.
Gold standard	Contemporaneous 12 lead ECG-telemetry data was recorded and used as a gold-standard for rhythm determination. Trained physicians reviewed all ECG and/or telemetry data to determine heart rhythm using standard criteria. In cases where reviewers disagreed about the rhythm diagnosis, a "tie-breaker" reader was consulted.
Expertise of index test interpreter	probably automated
Simultaneous index/gold vs non simultaneous	Yes
Results	Sensitivity 0.97, specificity 0.935 for the detection of an irregular pulse from AF when compared to the gold-standard diagnosis of AF by 12-lead ECG TP 95, FN 3, FP 6, TN 85
Source of funding	This work was funded in part by NIH grant 1R15HL121761, as well as the office of Naval Research work unit N00014-12-1-0171. DDM's time was funded by NIH grant KL2RR031981. Dr. Saczynski was supported in part by funding from the National Institute on

Reference	McManus, 2016[164]
	Aging (K01AG33643). Drs. McManus and Saczynski were supported in part by funding from the National Heart Lung and Blood Institute (U01HL105268). Dr. Boyer was supported by 1K24DA037109.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 44 Muller, 2009[172]

Reference	Muller, 2009[172]
Study type	Observational
Recruitment	24 with AF and 24 without – thus appears to be case control but described as consecutive
Setting	Internal Medicine Clinic
Country	Germany
Sample size	48
Sample characteristics	Mean age 62; 29/48 male; 24 with AF; consecutive patients at an internal medicine department.
Inclusion criteria	Presence of an indication for 24 hr Holter ECG
Exclusion criteria	Antibradycardic pacemakers; implantable cardioverters and defibrillators
Index test(s), including number of repetitions and duration	Vitaphone 3100 BT external loop recorder. Portable external device weighing 85kg and 8 x10 x 1.4 cm in size. Recorded event ECGs manually when triggered by the patient or automatically when there was AF, bradycardia, tachycardia or pauses. The automatic detection of fibrillation was based on recognition of arrhythmia in the QRS complex. The loop recorder could record events for up to 40 mins. Codes designating the type of event (ie AF) were transmitted making it an automated device.
Gold standard	24 hours 3 channel ECG (Holter). Connected to same points on skin as index test. Expertise of the interpreter unclear but likely to be the physician
Expertise of index test interpreter	Automated, but also appeared to be additionally evaluated by a physician
Simultaneous index/gold vs non simultaneous	Yes: The Holter was constantly recording. The index loop recorder was on intermittently, triggered by events, and so likely to be simultaneous
Results	Sensitivity 100, specificity 50; TP 24, FN 0, FP 12, TN 12

Reference	Muller, 2009[172]
Source of funding	None reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 45 Park, 2015[186]

Reference	Park, 2015[186]
Study type	Observational
Recruitment	consecutive
Setting	Unclear
Country	South Korea
Sample size	17
Sample characteristics	Patients c/o palpitations. No other details given.
Inclusion criteria	Patients with palpitations
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	The mobile ECG device ER-2000s is a pocket-sized (64 x 95 x 22mm3), battery-powered device that weighs 106g.. There are two different modes for recording an ECG rhythm strip with the ER-2000s. Mode 1 uses three ECG electrodes that are attached to the anterior chest wall and mode 2 uses the side chest channel and finger channel. The data obtained can be transmitted by USB cable, micro SD, or Bluetooth. The ER-2000s can record a real-time continuous cardiac rhythm strip for up to 2500 h. In this study, patients were instructed to push the record button when they believed they were experiencing a cardiac symptom.
Gold standard	12 lead ECG. The standard 12-lead ECG data were recorded on a piece of paper at a speed of 25 mm/s simultaneously, and compared with that recorded by the ER-2000s. The rhythm strips obtained from the 12-lead ECG were read in random order by two independent investigators who were blinded to patients' medical history and clinical characteristics, and rhythm status was compared. From the 12-lead ECG data, one lead with the most similar QRS vector and amplitude was chosen to compare the detailed morphologies of P, QRS, and T-wave with those obtained by mode 1 of ER-2000s.

Reference	Park, 2015[186]
Expertise of index test interpreter	The rhythm strips obtained from the ER-2000s were read in random order by two independent investigators who were blinded to patients' medical history and clinical characteristics
Simultaneous index/gold vs non simultaneous	Simultaneous
Results	Sensitivity 100%, specificity 100% This is derived from: 'The accuracy of rhythm diagnosis obtained by the two different modes of ER-2000s was accurate compared to that obtained by the 12-lead ECG in all patients, except in patient 3 in whom ER-2000s showed one atrial premature beat while 12-lead ECG showed sinus rhythm'. Since AF was differentiated from atrial premature beats in this study, specificity must still have been 100.
Source of funding	This study was supported by a research grant from Boryung Soo & Soo Ltd.,Seoul, Korea.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 46 Roten, 2012[219]

Reference	Roten, 2012[219]
Study type	Observational
Recruitment	consecutive
Setting	Cardiac outpatients clinic
Country	Switzerland
Sample size	88 (12 patients undergoing ablation included twice, before and after ablation) – therefore 100 datasets
Sample characteristics	Patients attending clinic for assessment of AF burden prior to ablation, and attending for screening post ablation; age 62.4; male 73%; hypertension 58%; DM 8%; IHD 18%; LVEF 60; LV diam 49mm; pre-ablation 15%; post ablation 52%; no ablation 46%
Inclusion criteria	Patients attending clinic for assessment of AF burden prior to ablation, and attending for screening post ablation; patients with known or suspected paroxysmal AF;
Exclusion criteria	Patients with persistent AF; patients unable to handle the devices independently.
Index test(s), including number of repetitions and duration	7 day triggered ECG (R.Test Evolution 3). This system monitors and displays the heart rate and summarises the number of atrial and ventricular premature beats as well as supraventricular and ventricular tachycardias during up to 8 days, but without recording a continuous ECG. It can store one ECG channel for a total duration of 20 min. Triggers for recording an ECG stripe can be programmed individually as well as the recording window before and after each trigger and the maximum possible number of

Reference	Roten, 2012[219]
	recordings for each trigger. Once the maximum number of recordings for a trigger is attained, only events better fulfilling triggering criteria than already recorded events (eg, longer pauses) will be recorded and replace less severe recordings. For this study, the triggers for recording an ECG by the tECG were programmed as absolute pauses (>2 sec), premature beats (<mean [RR –(25% x mean RR)]), bursts (>= 6 premature beats < mean RR – 25% x mean RR), or manual trigger. Two electrodes were applied to each patient, one on the upper part of the sternum and one on the left anterior axillary line at the lower left border of the ribcage. The ECG was derived from between the two electrodes. With the software RTSoft (Novacor) all recorded events as well as the 7-day heart rate histogram and arrhythmia summary were printed for analysis. The heart rate histogram in this device is only displayed at times when signal quality is suitable for automatic signal analysis, otherwise gaps are displayed. The duration of effective monitoring was calculated from the heart rate histogram and represents the total time with monitoring of heart rate (ie, signal suitable for automatic rhythm analysis). Heart rhythm of all recorded events was diagnosed. In case of a recording triggered by an artefact and showing sinus rhythm, the recorded event was labelled an artefact.
Gold standard	7 day continuous Holter (Lifecard CF). This system allows continuous recording of two ECG channels for 7 days. Three ECG electrodes were applied to each patient: one right to the upper border of the sternum (electrode 1); one on the right mid-clavicular line at the lower right border of the ribcage (electrode 2); and one on the left anterior axillary line at the lower left border of the ribcage (electrode 3). ECGs were derived from between electrodes 1 to 3 and 2 to 3. Event was arrhythmias (AF, atrial flutter or atrial tachycardia) of >=30 seconds duration. Interpreted by 2 experienced electrophysiologists.
Expertise of index test interpreter	2 experienced electrophysiologists
Simultaneous index/gold vs non simultaneous	Yes – both devices were simultaneously worn by every patient for 7 days. They could be removed occasionally (ie when showering) but they were asked not to selectively wear one device.
Results	Sensitivity 88%, specificity 100%; TP 37, FN 5, FP 0 TN 58 Note that the 5 FNs were due to no recording or no monitoring at these points – however it is right to deem these as FNs as such omissions are an intrinsic drawback of a non-continuous device.
Source of funding	Dr Tanner was supported by a grant from the Swiss Foundation for Pacemaker and Electrophysiology.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 47 Proesmans, 2019[197]

Reference	Proesmans, 2019[197]
Study type	Observational
Recruitment	Selective case/control
Setting	GP centres
Country	Belgium
Sample size	223
Sample characteristics	Age 77; male 46.6%; median (IQR) CHADSVASC 4(3-6); CHF 28.7%; DM 20.2%; stroke or TIA 22.4%; OACs 55.6%; mobile phone ownership 16.1%. From 17 GP centres.
Inclusion criteria	Known paroxysmal or persistent AF; aged >=65; other subjects without a history of AF.
Exclusion criteria	Active pacemakers
Index test(s), including number of repetitions and duration	**FibriCheck app**. a PPG signal was acquired with the rear camera of an iPhone 5S (Apple Inc). PPG is a technique whereby a volumetric measurement is optically obtained. A classic application of the PPG technique is the pulse oximeter, which illuminates the skin and measures changes in light intensity with blood volume pulse variation in the local arterioles and uses this information to determine arterial oxygen saturation and pulse frequency. The same principle can be applied by using the camera of a mobile phone and measuring the amount of reflected light. In this way, each heartbeat is recorded, and the rhythm can be determined on the basis of the intervals between heartbeats (ie, RR-intervals). The FibriCheck app provides software to obtain and analyze such measurements with most common mobile phones. To obtain a high-quality PPG signal, subjects were asked to adopt a sitting position with both arms resting on a table, holding the iPhone 5S in a vertical position with their right hand. Subsequently, they were asked to cover the flashlight and the rear camera horizontally with their left index finger. The measurement time to acquire the PPG signal with the FibriCheck app is 1 min, visualized by a countdown clock on the mobile phone screen. To minimalize motion artefacts, subjects were instructed not to speak or move during the registration process. Subjects were asked to independently perform 3 consecutive measurements. To avoid evoking a reaction following the result of a measurement, researchers and participants were blinded for the PPG signal during the measurements and the automated interpretations after the measurements. Simultaneously with the PPG measurement, a synchronized **single-lead ECG** was obtained using the ECG-bone (Interuniversity Micro-Electronics Centre, IMEC). This module was attached with a patch on the left side of the subject's chest above ribs 2 and 3 and was wirelessly connected to the iPhone 5S with the help of the FibriCheck app. This procedure was performed by the same researcher who helped with the operation of the FibriCheck app.
Gold standard	The same researcher obtained a 12-lead ECG (gold standard). The ECGs were taken using digital machines CardiMax FCP-7101 (Fukuda Denshi), CP 50 (Welch Allyn), Universal ECG (QRS Diagnostic), and ECG-1150 (Nihon Kohden Corporation) and the data were immediately printed. All 12-lead ECGs were analyzed offline on the basis of the Minnesota Code Classification System for Electrocardiographic Findings (code 8-3-1) by 2 experienced, independent cardiologists blinded to all other data. In case of a disagreement, a third cardiologist was consulted to interpret the rhythm.

Reference	Proesmans, 2019[197]
Expertise of index test interpreter	Researcher so likely to have high expertise
Simultaneous index/gold vs non simultaneous	Unclear – no mention of synchronicity
Results	**PPG** Sensitivity 95.6% (89.1-98.8); specificity 96.6%(91.4%-99.1%) when excluding the 16/223 index test results of 'insufficient quality' TP 87, FN 4, FP 4, TN 112 Sensitivity 87% (78.8-92.9); specificity 96.8 (91.9-99.1)% when including the 16/223 index test results of 'insufficient quality' as sinus rhythm TP 87, FN 13, FP 4, TN 119 Sensitivity 96% (90.1-98.9); specificity 91.1% (84.6-95.5) when including the 16/223 index test results of 'insufficient quality' as AF TP 87, FN 4, FP 11, TN 112 **1 lead ECG** Sensitivity 94.7% (88.1-98.3); specificity 96.6%(91.3%-99.0%) when excluding the 13/223 index test results of 'insufficient quality' TP 86, FN 5, FP 4, TN 106 Sensitivity 90% (82.4-95.1); specificity 96.8%(91.9-99.1) when including the 13/223 index test results of 'insufficient quality' as sinus rhythm Unclear raw data Sensitivity 95% (88.7-98.4); specificity 91.1% (83.6-94.9) when including the 13/223 index test results of 'insufficient quality' as AF Unclear raw data
Source of funding	Qompium (Hasselt, Belgium) provided the mobile phone and free use of the FibriCheck app. IMEC (Leuven, Heverlee, Belgium) offered the ECG-bone device without cost. Both companies had the opportunity to check the final version of the manuscript and to make recommendations but were not involved in the data collection, analysis, or decision to submit the report for publication.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 48 Rozen, 2018[220]

Reference	Rozen, 2018[220]
Study type	Observational – case control
Recruitment	Selective case/control
Setting	Cardioversion patients
Country	USA
Sample size	99 (but each patient contributed two sets of data – pre-cardioversion and post-cardioversion).
Sample characteristics	Patients with paroxysmal AF referred for Holter monitoring for arrhythmia detection. 73 men/24 women; age 67.7; 91.8% white; 1% Hispanic/Latino; 1% Black; 1% Asian
Inclusion criteria	Consecutive patients with a diagnosis of AF who were scheduled for elective direct current cardioversion (DCCV) at MGH
Exclusion criteria	<18 years
Index test(s), including number of repetitions and duration	Cardio Rhythm Mobile Application (CRMA). CRMA recordings done before and after CV. The CRMA was installed and used on an iPhone to obtain readings for all patients before and after CV. This application was developed to be used a supervised machine learning technique known as a support vector machine to classify PPG waveforms. The underlying feature extraction algorithm analyses the degree of self-similarity of a PPG waveform over time to find repeating patterns instead of simply assessing beat-to-beat changes in the PPG waveform. Each patient placed his or her index finger against the camera of the iPhone and the application was turned on to record a reading. Twenty-second finger pulse recordings were obtained for each patient 3 times before and 3 times after the CV procedure. The CRMA recordings were labelled as AF if at least 2 of the 3 recordings were sufficiently irregular; otherwise, the CRMA recordings were labelled as non-AF.
Gold standard	12 lead ECG, done before and after CV. A12-lead ECG, obtained as part of the standard CV procedure, was used as the gold standard for rhythm classification. In the rare cases in which a 12-lead ECG was not available, single-lead rhythm strips obtained concurrently with the Cardiio Rhythm Mobile Application recordings were used. Two board-certified cardiologists (AR1 and AR2) interpreted the 12-lead ECGs or, in rare cases, the single-lead rhythm strips. Both readers were blinded to the CRMA results and to each other's interpretation of the ECGs. In case of a discrepancy between the readings by the 2 cardiologists, a senior electrophysiologist with more than 40 years of clinical experience (JNR) interpreted the ECG and his conclusion was used as the final diagnosis.
Expertise of index test interpreter	Unclear if automated or not; no reporting of who would have interpreted it

Reference	Rozen, 2018[220]
Simultaneous index/gold vs non simultaneous	Unclear
Results	Sensitivity 93.1(86.9-97.2); specificity 90.9%(82.9-96); TP 94, FN 7, FP 8, TN 80 Based on 97 sets of data for pre-CV and 92 sets of fata for post CV [5 missing from post-CV measurements because of normal sinus rhythm at baseline (n=1), contraindication to procedure (n=3), drop-out (n=1)]
Source of funding	No funding reported. Drs.Yukkee and Ming-Zher Poh are employees of Cardiio, Inc. and have an ownership stake in the company. Dr Ming- Zher Poh has a patent for the AF detection algorithm described here. There are no other potential conflicts of interest relevant to this study.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 49 Sejr, 2019[233]

Reference	Sejr, 2019[233]
Study type	Observational
Recruitment	consecutive
Setting	Neurology inpatients
Country	Denmark
Sample size	1412
Sample characteristics	56% male; age 72.8; TIA 39.8%; Ischaemic stroke 60.2%; hypertension 58.4%; LVEF <40% 1.4%; DM 14.3%; current smoker 24.6%; OACs 0.78%;
Inclusion criteria	Acute ischaemic stroke or transient ischaemic attack (TIA) with first symptoms within 1 week, age ≥60 years, no AF on 12-lead admission ECG, no prior AF according to International Classification of Diseases codes (ICD-10) from outpatient clinic visits, hospitalisations or review of medical records, no active cancer, no implanted pacemaker, no expected low compliance or precedent participation in this study and written informed consent.
Exclusion criteria	See above
Index test(s), including number	R.Test Evolution 4 (NorDiaTech, Paris, France) was device used as External loop recording (ELR). This device is non-invasive and records heart rhythm using two skin electrodes attached over sternum and cardiac apex. ELR recorders were attached by nurses

Reference	Sejr, 2019[233]
of repetitions and duration	after manufacturer's recommendations. The ELR analyses segments of 64 consecutive RR intervals (intervals between R waves), when at least two-thirds of these intervals are irregular, categorises heart rhythm as AF and stores a recording of AF episode in memory. Depending on heart rate, the ELR is able to categorise AF episodes lasting from approximately 25 s, thereby suitable for detecting AF exceeding 30 s. Storing capacity is 60 min, and if this is exceeded, only the most characteristic AF episodes are kept. AF episodes with fastest heart rates are kept in memory. We adjusted ELR according to manufacturer's recommendations. We saved 1 min recording per AF episode, allowing for a maximum of 54 AF recordings per patient, while 6 min were spared for storage of episodes of other arrhythmia. Analysis of ELR findings was blinded for continuous ECG recording results.
Gold standard	Continuous ECG monitoring for 48 hours. The continuous ECG recorder used was Life Card CF digital ECG recorder from Spacelabs Healthcare Diagnostic Cardiology (Washington, USA). Nurses trained and experienced in analysing continuous ECG recordings reviewed recordings. Episodes classified as AF were verified by the three members of the research team. Analysis was blinded to ELR results. AF was defined according to current guidelines, as an atrial arrhythmia with irregular intervals between R waves, without detectable normal P waves and lasting more than 30 s
Expertise of index test interpreter	Three experienced members of the research team (MHS, OM and JCN) each reviewed and classified as AF or non-AF all recordings automatically classified as AF by the ELR. In case of ambiguity, agreement was reached by consensus.
Simultaneous index/gold vs non simultaneous	Yes
Results	**Automated ELR** Sensitivity 92(79-98); specificity 87(85-88); TP 35, FN 3, FP 179, TN 1195 **Cardiologist-verified ELR** Sensitivity 84(69-94); specificity 98(97-99); TP 32, FN 6, FP 27, TN 1347
Source of funding	This work was supported by Health Research Fund of Central Denmark Region (1-31-72-15-14), Danish Heart Foundation (14-R97-A5075-22884/17-R115-A7606-22069) and Aase and Ejnar Danielsen Foundation (10-001847). Novo Nordisk Foundation (NNF16OC0018658) and an institutional unrestricted grant from Abbott, Denmark, supported JCN.
Limitations	Risk of bias (QUADAS 2 – risk of bias): No serious risk Indirectness (QUADAS 2 - applicability): none

Table 50 Mulder, 2012[171]

Reference	Mulder, 2012[171]
Study type	Observational
Recruitment	consecutive

Reference	Mulder, 2012[171]
Setting	Cardiac outpatients
Country	Netherlands
Sample size	96
Sample characteristics	Patients who had undergone PVI 12 months previously for paroxysmal AF; 25% female; 39% hypertension; 7% LVEF <55%; 13% mitral regurgitation grade 2; age 59; duration of AF 7 years
Inclusion criteria	Patients who had undergone PVI 12 months previously for paroxysmal AF
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	Holter for 1,2,3,4,5,6 days
Gold standard	A 7-day Holter was performed in all patients and evaluated for arrhythmia episodes and the duration of each episode. A documented arrhythmia recurrence was defined as an LA arrhythmia comprising AF/flutter/tachycardia lasting more than 30 seconds. Expertise of interpreters not given
Expertise of index test interpreter	Not reported
Simultaneous index/gold vs non simultaneous	Yes – not directly reported but can be inferred
Results	Because > 1 measurement made on each person the data were clustered o this has been adjusted for in the analysis. For calculating sensitivity and NPV in the clustered data (e.g., seven parts of 1 day within a 7-day Holter), first the intraclass correlation coefficient (ICC), or ρ, was calculated as a measure of the relation of clustered data. Value of ρ range from 0 (no clustering, people within a cluster are just the same as people in the other clusters) to 1 (people in the same cluster are more similar to each other than to people in other clusters). If ρ = 0, the binomial estimator was used for the sensitivity and NPV, between 0.2 and 0.4, the ratio estimator, within-cluster correlation estimator or weighted estimator, when 0.6 the weighted estimator was used. No false negatives so specificity 100% for all time points. Raw data not really calculable because of adjustments, but raw data have bene calculated below on basis of AF=21, no AF=75 on 7 day Holter 1 day: sensitivity 53%; specificity 100%; TP 11; FN 10, FP 0, TN 75 2 days sensitivity 68%; specificity 100%; TP 14; FN 7, FP 0, TN 75 3 days sensitivity 80%; specificity 100%; TP 17; FN 4, FP 0, TN 75 4 days sensitivity 88%; specificity 100%; TP 18; FN 3, FP 0, TN 75 5 days sensitivity 94%; specificity 100%; TP 19; FN 2, FP 0, TN 75

Reference	Mulder, 2012[171]
	6 days sensitivity 98%; specificity 100%; TP 20; FN 1, FP 0, TN 75
Source of funding	The Cardiology Department has received grant support for research from Ablation Frontiers, Inc.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 51 Kao, 2018[123]

Reference	Kao, 2018[123]
Study type	Unclear but likely to be case-control
Recruitment	Unclear – likely to be case/control
Setting	Emergency department
Country	Taiwan
Sample size	63 (1 excluded as not fulfilling inclusion criteria)
Sample characteristics	Recruited from emergency department; age 67; 56% male; AF 29/62
Inclusion criteria	Aged >20 years; either with AF or no AF (diagnosed by 12 lead ECG).
Exclusion criteria	People exposed to high frequency surgical equipment during testing' people with cardiac pacemakers or implantable defibrillators; pregnant women
Index test(s), including number of repetitions and duration	The Heart Spectrum Blood Pressure Monitor. Human blood pressure and heart rate were measured using the oscillometric method. Each heartbeat causes the heart to emit blood, and then the sensor of the Heart Spectrum Blood Pressure Monitor on the arm detects the blood pressure and depicts the time-domain pressure wave. The time-domain pressure wave is converted to an energy-domain frequency wave via Fast Fourier Transform (FFT). There are primary frequency peaks when the wave is converted via FFT. When observing abnormal frequency, the frequency peaks other than the primary frequency peaks are considered heart noises, and can be quantified as the heart index, as described below. We defined the first frequency region as the first heart rate frequency ± 0.5 frequency interval, the second frequency region as the second heart rate frequency ± 0.5 frequency interval, and the third frequency region as the third heart rate frequency ± 0.5 frequency interval. For example, if the first heart rate frequency is 60 beats per minute (1 Hz), then the first frequency region is 30 to 90 beats per min, the second frequency region is 90 to 150 beats per min, and the third frequency region is 150 to 210 beats per min, wherein the heart index I1 is the sum of noise in the first frequency region, the heart index I2 is the sum of noise in the second frequency region, and the heart index I3 is the sum of noise in the third

Reference	Kao, 2018[123]
	frequency region. The heart index = I1 + I2 + I3. We defined the heart noise as the number of other spikes above 1/20 for each region. The scale factor of 1/20 was determined by removing the background noise from clinical pre-test results. AF analysis: Measurements were obtained from each subject consecutively three times using method 1 (M1), method 2 (M2), and method 3 (M3). M1 involved the following: standard blood pressure measurement was used to determine the heart index and was compared with the 12-lead ECG synchronously. M2 involved the following: from M1, the systolic and diastolic pressures were obtained and the mean arterial pressure (MAP) was calculated. MAP was then used as the constant pressure measurement to determine the heart index and was compared with the 12-lead ECG results at the same time. M3 involved the following: a constant pressure (60 mmHg) was used to analyze the heart index and to compare it with the simultaneous 12-lead ECG results.
Gold standard	12 lead ECG. 'Interpreted by the examining physician'
Expertise of index test interpreter	Physician
Simultaneous index/gold vs non simultaneous	Yes, simultaneous
Results	Method 1: sensitivity 97%, specificity 97%; TP 28, FN 1, FP 1, TN 32 Method 2: sensitivity 90%, specificity 100%; TP 26, FN 3, FP 0, TN 33 Method 3: sensitivity 100%, specificity 94%; TP 29, FN 0, FP 2, TN 31
Source of funding	This study was supported by the Medical and Pharmaceutical Industry Technology and Development Center. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 52 McManus, 2013[165]

Reference	McManus, 2013[165]
Study type	Observational
Recruitment	Selective case/control (paired)
Setting	Cardioversion patients

Reference	McManus, 2013[165]
Country	USA
Sample size	76 (undergoing cardioversion for AF; those in AF on 12 lead ECG at pre-CV, and those in sinus rhythm on 12 lead ECG at post-CV measured with iphone device).
Sample characteristics	Age 65.3; male 77%; white 96%; hypertension 71%; hyperlipidaemia 62%; current smoking 8%; DM 28%; CAD 29%; CHF 21%; sleep apnea 16%; 11% CABG; prior cardioversion 27%; stroke 12%
Inclusion criteria	Patients with persistent AF on a roster of patients scheduled to have elective cardioversion for AF
Exclusion criteria	Not reported
Index test(s), including number of repetitions and duration	iPhone 4S camera. Placed directly on right index or second finger for 2 minutes while AF detection application was run. Pulse signal recordings were obtained with patients while they were in a supine position and breathing spontaneously. The application acquired pulsatile signals by illuminating the fingertip using the standard iPhone lamp and recording video signal (30 frames/s) for 2 minutes. The signal was processed by averaging 50 x 50 green band pixels per frame. Researchers interpolated the pulsatile signal to 30 Hz using a cubic spline algorithm followed by peak detection. Normalised RMSSD (root mean square of successive difference) and ShE (Shannon entropy) measured and automatically compared to threshold values of 0.115 and 0.55 respectively (both had to be > threshold).
Gold standard	12 lead ECG done pre- and post-CV. Interpreted by 2 'trained physicians'. In cases where there was disagreement a third expert adjudicator used.
Expertise of index test interpreter	Trained physicians
Simultaneous index/gold vs non simultaneous	Does not appear to be simultaneous. Likely to be the same day at least but average interval unclear
Results	**Using both RMSSD and Shannon entropy (DEFAULT method used automatically in application)** Sensitivity 96.19%; specificity 97.52%; **Using just the RMSSD threshold** Sensitivity 98.18%; specificity 91.5% **Using just Shannon entropy** Sensitivity 97.5%; specificity 82.18%

Reference	McManus, 2013[165]
	Raw data not possible to calculate as paper did not specify numbers of patient-readings with gold standard AF and no AF (cannot assume that all 76 were successfully cardioverted)
Source of funding	This work was funded in part by the Office of Naval Research work unit N00014-12-1-0171. Dr McManus's time was funded by National Institutes of Health through grants 1U01HL105268-01and KL2RR031981.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 53 Williams, 2015[284]

Reference	Williams, 2015[284]
Study type	Observational
Recruitment	Selective case/control but unclear
Setting	Outpatient AF clinic
Country	UK
Sample size	99
Sample characteristics	29 with AF on ECG; other details not reported
Inclusion criteria	Patients attending regular AF clinic at the North west heart centre in University hospital in Manchester; Other patients attending for 12 lead ECG for reasons other than AF
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	Alive-Cor device. 30 second reading taken using application on phone. No further details provided.
Gold standard	12 lead ECG, interpreted blinded by a cardiac physiologist and a GP with special interest in cardiology. Unclear how disagreements were adjudicated.
Expertise of index test interpreter	The same interpreters as for gold standard. Interpreted as AF or no AF.

Reference	Williams, 2015[284]
Simultaneous index/gold vs non simultaneous	12 lead ECG was recorded and printed 'at the same time'.
Results	**Alive Cor using cardiac physiologist as interpreter** Sensitivity 90, specificity 86; TP 26, FN 3, FP 9, TN 57 **Alive Cor using GP as interpreter** Sensitivity 93, specificity 76; TP 27, FN 2, FP 16, TN 50
Source of funding	Reported that no funding received.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 54 Brito, 2018[23]

Reference	Brito, 2018[23]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology inpatients
Country	Switzerland
Sample size	127
Sample characteristics	Age 62; males 64.6%; MI 22.8%; CABG 6.3%; CorAngio 33.9%; valvular Sx 7.9%; sinus at baseline 85%
Inclusion criteria	Consecutive patients admitted to the cardiology ward of Geneva University Hospital for coronarography 17.3%, electrophysiology procedure 26%, pacemaker implantation 3.9%, cardiac failure 3.9%, other 52%.
Exclusion criteria	Patients with pacemaker or cardioverter defibrillator
Index test(s), including number of repetitions and duration	Beurer ME90 device – a handheld ECG recorder. 30 secs recording by 1) holding the device between the index fingers [lead I], and then 2) against the chest corresponding to lead mV4. Handheld recordings and also the automatic interpretation by device downloaded to computer for visualisation by software. Interpretation blinded to gold standard results

Reference	Brito, 2018[23]
Gold standard	12-lead ECG, interpreted by a qualified electrophysiologist
Expertise of index test interpreter	Non-automated handheld device readings also interpreted by the same electrophysiologist.
Simultaneous index/gold vs non simultaneous	No – index tests done shortly after the 12 lead ECG.
Results	Results for detection of AF/flutter **Lead I (automatic) n=123** Sensitivity 88.9(65.3-98.6), specificity 61.9(51.9-71.2); TP 16, FN 2, FP 40, TN 65 **mV4 (automatic) n=119** Sensitivity 94.1(71.3-99.9), specificity 77.2(67-84.3); TP 16, FN 1, FP 24, TN 78 **Lead I and mV4 combined* (automatic) n=119; *only positive if both scores positive** Sensitivity 88.2(63.6-98.5), specificity 84.3(75.8-90.8); TP 15, FN 2, FP 16, TN 86 **Manual analysis by electrophysiologist lead 1 n=126** Sensitivity 84.2(60.4-96.6), specificity 100 (96.6-100); TP 16, FN 3, FP 0, TN 107 **Manual analysis by electrophysiologist mV4 n=126** Sensitivity 84.2(60.4-96.6), specificity 100 (96.6-100); TP 16, FN 3, FP 0, TN 107
Source of funding	Reported no funding received
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 55 Doliwa, 2009[63]

Reference	Doliwa, 2009[63]
Study type	Observational

Reference	Doliwa, 2009[63]
Recruitment	consecutive
Setting	Cardiology outpatient clinic
Country	Sweden
Sample size	100 (the part of the study concerned with diagnostic accuracy)
Sample characteristics	Patients with atrial fibrillation, atrial flutter or sinus rhythm recruited from cardiology department.
Inclusion criteria	As above
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	Thumb ECG device: Zenicor ECG, with 2 thumb sensors (providing a bipolar lead I ECG) on front display panel of 110c 80 x 15 mm box. Thumbs applied for 10 seconds. Automated transmission to interpreter (cardiologist) who interpreted it at a later date.
Gold standard	12 lead ECG, interpreted by cardiologist and blinded to index results
Expertise of index test interpreter	Interpreted by same cardiologist who was blinded to gold standard results
Simultaneous index/gold vs non simultaneous	No – 12 lead done immediately prior to index test
Results	Sensitivity 96, specificity 92; Descriptions of raw data do not tally with these figures. The description suggests: TP 47, FN 4, FP 2, TN 47, which would give sensitivity of 92 and specificity of 96. However, if the final accuracy data are correct, likely there was an error in description, so raw data are: TP 47, FN 2, FP 4, TN 47
Source of funding	Swedish Innovation Agency and Stockholm County Council
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 56 Nigolian, 2018[177]

Reference	Nigolian, 2018[177]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology inpatients
Country	Switzerland
Sample size	52
Sample characteristics	Age 69; male 58%; pacemaker 10%; hypertension 60%; DM 21%; COPD 8%; AF on 12 lead ECG 31%; OACs 40%
Inclusion criteria	Consecutive patients admitted to the cardiology department at a University Hospital
Exclusion criteria	<18 years; inability or unwilling to consent
Index test(s), including number of repetitions and duration	Beurer ME 80 device – a pocket sized (reconstructing 9 lead) ECG device that had electrodes mounted on each end. Can be used 1) between fingers of each hand or 1) against the chest. For this study, tracings corresponding to the bipolar limb leads (I,II and II) and 6 precordial leads (V1-6) were recorded in a successive order. Lead I was obtained by placing the right index on the cathode, and left index on the anode; lead II by placing the right index on the cathode and applying the anode to the left thigh; lead III by placing the left index on the cathode, and applying the anode on the left thigh. Leads V1-6 were obtained by applying directly the anode on the chest in the corresponding locations, while holding the cathode in the right index. A 9 lead ECG was reconstituted for each patient by assembling 5-second sequential sequences from the different recordings of the handheld device. Recordings transmitted to computer for later viewing. Blinded.
Gold standard	Standard 12 lead ECG recorded at 0.05-150Hz using a Schiller Cardiovit AT-170 ECG. Interpreted by a certified cardiologist. Blinded.
Expertise of index test interpreter	Also interpreted by a certified cardiologist and also by a fellow in internal medicine.
Simultaneous index/gold vs non simultaneous	Not simultaneous – 12 lead ECGs reported to be 'followed by' the index test
Results	**With index test interpreted by cardiologist** Sensitivity 100(79-100), specificity 94(81-99); TP 16, FN 0, FP 2, TN 34 **With index test interpreted by fellow in internal medicine** Sensitivity 75(48-93), specificity 89(74-97); TP 12, FN 4, FP 4, TN 32
Source of funding	Paper reports that no funding was received
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious

Reference	Nigolian, 2018[177]
	Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 57 Winkler, 2011[286]

Reference	Winkler, 2011[286]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology inpatients
Country	Germany
Sample size	60
Sample characteristics	Not reported
Inclusion criteria	Patients admitted to the cardiology department
Exclusion criteria	Not reported
Index test(s), including number of repetitions and duration	Handheld ECG device with dry electrodes that records 3 lead ECG (Einthiven I, II and III leads). Records over patient chest for 120 secs. Works by analysis of the irregularity of R-R intervals. Based on the R-R differences the AF index was calculated. AF index calculated in overlapping 52 beat windows from the histogram of R-R interval differences. The index is calculated from the ratio of histogram width to height, the position of the histogram peak, and the number of premature ventricular beats according to the formula: index=[(HistW/HistH)x20] – HistM –PVC%], where HistW = width of histogram of R-R differences, HistH is the height of the histogram of R-R differences, HistM is the position of the histogram peak and PVC% is the % of premature ventricular beats in the 52 beat window. ROC analysis showed AF Index threshold value of 25 was ideal and this was used as the threshold in the study. 52 beat window required for calculation of AF index.
Gold standard	12 lead ECG. Recorded by nurse and interpreted by cardiologist.
Expertise of index test interpreter	Done by automated algorithm
Simultaneous index/gold vs non simultaneous	No – index done just before 12 lead ECG
Results	Sensitivity 92.9, specificity 90.9; raw data difficult to ascertain as description of raw data is flawed by different numbers having the index and gold standard – thus not possible to calculate raw values.

Reference	Winkler, 2011[286]
Source of funding	No conflicts reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 58 William, 2018[283]

Reference	William, 2018[283]
Study type	Observational
Recruitment	Consecutive, but paired analysis in that each patient was medically CV or not
Setting	Cardiac inpatients
Country	USA
Sample size	52 participants with 225 sets of measurements
Sample characteristics	Age 68.1; 67.3% male; PAF 21.2%; persistent AF 78.8%; palpitations 42.3%; SOB 65.4%; lightheadedness 17.3%; chest pain 5.8%; fatigue 51.9%
Inclusion criteria	Patients with a diagnosis of AF admitted for AAD therapy; aged 35-85; history of PAF or persistent AF; baseline corrected QT interval <470 or 500 if QRS duration >120ms
Exclusion criteria	Patients with pacemakers; patients with defibrillators
Index test(s), including number of repetitions and duration	Kardia Mobile Cardiac Monitor (provided by Alive-Cor, with a wi-fi enabled smart ipod device). This is a handheld device. Used immediately after the ECG – patients had to do a 30 second reading (equivalent to a lead I ECG) by placing at least 1 finger from each hand on the electrodes. Rhythm strip automatically analysed using the algorithm. Details of the algorithm not provided. The strips also downloaded for review by blinded electrophysiologist.
Gold standard	12 lead ECG, done 2 hours after each of the 6 twice daily AAD doses during the period of admission (patients in AF after 4th dose given electrical CV). Interpreted by blinded electrophysiologist
Expertise of index test interpreter	Electrophysiologist for non-automatic; NA for automated
Simultaneous index/gold vs non simultaneous	Not quite – index test done ‘immediately’ after ECG

Reference	William, 2018[283]
Results	**Note that of the 225 recording sets, there were 2 non-interpretable 12 lead ECGs, and 62 non-interpretable index test recordings.** **KMCM automated (with uninterpretable index readings not included)** Sensitivity 96.6, specificity 94.1; TP 57, FN 2, FP 6, TN 96 **KMCM physician interpreted (with uninterpretable index readings not included)** Sensitivity 100, specificity 89.2; TP 75, FN 0, FP 15, TN 124 **KMCM automated (with uninterpretable index readings included as negative) NOT IN PAPER** Sensitivity 71.25, specificity 67.1; TP 57, FN 23, FP 47, TN 96 **KMCM physician interpreted (with uninterpretable index readings included as negative) NOT IN PAPER** Sensitivity 93.75, specificity 86.71; TP 75, FN 5, FP 19, TN 124
Source of funding	Dr Varma serves on advisory board of and as a consultant to Medtronic and Abbott and on speakers bureau for Biotronik. Dr Trakji serves on the advisory board of Medtronic and AliveCor
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 59 Velthuis, 2013[268]

Reference	Velthuis, 2013[268]
Study type	Observational
Recruitment	consecutive
Setting	Stroke Unit
Country	Netherlands
Sample size	153

Reference	Velthuis, 2013[268]
Sample characteristics	Age 67; HT 59.5%; DM 19%; COPD 5.9%; TIA 10.5%; iCVA 7.8%; CAD 6.5%; HF 1.3%; Valve disease 6.5%; Bradytachy syndrome 0.7%; other arrhythmia 0.7%
Inclusion criteria	Consecutive patients aged >18 years admitted with a provisional diagnosis of acute ischaemic stroke
Exclusion criteria	Patients with known history of AF
Index test(s), including number of repetitions and duration	24 hour external loop recorder (single channel device 3100 BT, Vitaphone, Mannheim), using automated settings, according to the following non-adjustable algorithm, according the R-R variability within past 14 complexes: AF if 6/14 R-R intervals matched RRx – RRy > RRx/8 AND RRx – RRy < 2*RRx
Gold standard	24 hour external loop recorder, interpreted by 2 blinded qualified analysts
Expertise of index test interpreter	Not applicable as automated
Simultaneous index/gold vs non simultaneous	Yes, same devices used and the gold standard was simply the use of physicians rather than automated readings.
Results	Sensitivity 94.9, specificity 50.6; TP 56, FN 3, FP 1134, TN 1162
Source of funding	No funding declared
Limitations	Risk of bias (QUADAS 2 – risk of bias): No serious risk Indirectness (QUADAS 2 - applicability): none

Table 60 Haberman, 2015[90]

Reference	Haberman, 2015[90]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology outpatients
Country	USA
Sample size	130 (there were 251 other participants form other populations also analysed, such as athletes and asymptomatic students, but the 130 are the cardiology clinic patients of relevance to this review)
Sample characteristics	Age 59; male 56%; mean HR 72

Reference	Haberman, 2015[90]
Inclusion criteria	Ambulatory cardiology patients
Exclusion criteria	Not reported
Index test(s), including number of repetitions and duration	AliveCor device, which allowed user to record a wireless 30 sec ECG. To record the ECG the user touched the device with a finger of both hands. 30 seconds of waveforms were obtained, with the help of an iphone or ipad. Study organisers assisted, and patients able to collect their own ECG easily with 1-2 mins of training. Interpreted by automated algorithm. No detail of the algorithm.
Gold standard	12 lead ECG, interpreted by 2 board certified electrophysiologists.
Expertise of index test interpreter	Physician interpreted
Simultaneous index/gold vs non simultaneous	No, 12 lead taken immediately after index.
Results	Sensitivity 94.4, specificity 99.1; TP 17, FN 1, FP 1, TN 111
Source of funding	No funding declared
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 61. Poulsen, 2017[196]

Reference	Poulsen, 2017[196]
Study type	observational
Recruitment	consecutive
Setting	Neurology inpatients
Country	Denmark
Sample size	100
Sample characteristics	age 78; male 43/95; TIA 18/95; median CHADSVASC 5; median NIHSS 1; median time from stroke 4 days; median number of thumb ECG recordings 59; median duration of Holter monitoring 4.8 days
Inclusion criteria	>65 years; no history of AF who suffered an acute stroke or TIA of unknown origin in past 3 months verified by CT or MRI or clinically diagnosed; ability to handle thumb ECG

Reference	Poulsen, 2017[196]
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	30s thumb ECG (Zenicor Medical Systems AB) twice daily for 30 days (so should be 60). To be used additional time if palpitations. No further details provided
Gold standard	5 days Holter (Lifecard CF device). AF defined as irregular R-R intervals, absence of p waves and irregular atrial activity lasting 30 s. Initiated immediately after admission. Interpreted by a cardiologist and documented on a report that was confirmed by the second cardiologist.
Expertise of index test interpreter	Interpreted by same cardiologist who analysed gold standard and additionally by another cardiologist blinded to other cardiologist result (unclear if blinded to gold standard result). Consensus used to decide on final adjudication.
Simultaneous index/gold vs non simultaneous	Concurrent, so all time that index was recording, the gold standard was recording.
Results	Sensitivity 58.8, specificity 87.2; TP 10, FN 7, FP 10, TN 68
Source of funding	Department of neurology, Herlev Hospital and Carl and Ellen Hertz' grant to Danish medical and natural science
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 62. Renier, 2012[208]

Reference	Renier, 2012[208]
Study type	Observational
Recruitment	consecutive
Setting	Emergency department
Country	Belgium
Sample size	177
Sample characteristics	55 years; 45% men
Inclusion criteria	All consecutive patients visiting ED of University hospital in Belgium; any patients hospitalised in one respiratory, one gynaecological and one orthopaedic hospital ward on one day.

Reference	Renier, 2012[208]
Exclusion criteria	<18 years; unable to use right hand for heartscan device; did not understand language used by HCPs; no consent
Index test(s), including number of repetitions and duration	Heartscan is a hand-held device (121x67x24mm) that can be placed on the bare chest without cables, patches, suction heads or clamps, and is kept in place by patients right index finger for 30 seconds. Corresponds to the V3-V4 leads of a standard ECG. Provides traces and an automatic reading. Blinded.
Gold standard	12 lead ECG, taken and read at the same time by experienced university-hospital based cardiologist. Blinded.
Expertise of index test interpreter	Automated or by 2 GPs (one young and one experienced)
Simultaneous index/gold vs non simultaneous	No – 'immediately after' the index reading
Results	**AF/flutter** **Clinician interpretation of Heartscan (unclear which of the GPs, or whether was a majority rule or consensus decision)** Sensitivity 69.2, specificity 94.5; TP 9, FN 4, FP 9, TN 155 **Automated Heartscan** Sensitivity 92.3, specificity 100; TP 12, FN 1, FP 0, TN 164
Source of funding	NIHR programme grant RP-PG-0407-10347; Omron provided 10 Heartscan devices for free
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 63. Rizos, 2010[214]

Reference	Rizos, 2010[214]
Study type	Observational
Recruitment	consecutive
Settings	Tertiary care stroke unit

Reference	Rizos, 2010[214]
Country	Germany
Sample size	136
Sample characteristics	Patients admitted to a tertiary care stroke unit; age 72; male 58.8%; manifest stroke 88.2%; TIA 11.8%; duration of bedside ECG monitoring 97hrs; CHF 36%; MI 22.8%; HT 79.4%; DM 30.1%
Inclusion criteria	Patients > 60 years presenting with an acute ischemic stroke or TIA in the ER and who were subsequently admitted to the stroke unit of our hospital and underwent continuous ECG monitoring for a minimum period of 48 h were enrolled
Exclusion criteria	Patients with AF on the initial 12-channel ECG (ELI 350; Mortara Instruments, Milwaukee, Wisc., USA) in the ER or a history of paroxysmal or persistent AF were excluded
Index test(s), including number of repetitions and duration	6 channel Holter (H12+, Mortara Instruments) performed for 24 hours. 12-bit resolution digital ECG recoding for 1-2 hours. These ECG data were sent via internet to a computer where an unsupervised ASA was applied using the stroke risk analysis software (SRA; apoplex medical technologies, Pirmasens, Germany). The software employs an algorithm which creates an RR list of the ECG data, detects QRS complexes and then classifies atrial and ventricular beats. It performs time series analysis which includes 6 mostly nonlinear mathematical parameters. These parameters are derived from principle component analysis, RR difference plots, the ratio between shortest and longest interval of maximum 6 consecutive RR intervals, the number of atrial premature complexes, complexes without sinus nodal reset and approximate entropy of RR interval data. Based on this ASA analysis, the risk of pAF was estimated by the software and each patient was assigned to 1 of 5 predefined categories: (1) continuous sinus rhythm; (2) ventricular rhythm disorders; (3) intermediate risk of pAF; (4) high risk of pAF; (5) manifest episodes of AF. Reports for each patient were created by the system and sent to the clinical investigators via e-mail
Gold standard	Continuous ECG bedside monitoring for duration of stay in stroke unit (IQR 82-144 hrs, none <48hrs). Used Infinity Delta monitoring system. When AF suspected from monitor trace then a 12 channel ECG used and interpreted by cardiologist. AF defined as AF episode lasting >30s.
Expertise of index test interpreter	Holter: Results analysed and interpreted by a cardiologist using the H-Scribe software. ASA: automated
Simultaneous index/gold vs non simultaneous	Concurrent
Results	**Holter** Sensitivity 0.23, specificity 1; TP 3, FN 10, FP 0, TN 107 **ASA (threshold categories 3-5)** Sensitivity 0.72, specificity 0.63; TP 21, FN 8, FP 40, TN 67
Source of funding	Funding from the University of Heidelberg. Holter ECG recorders were provided by Spacelabs Healthcare. R.V. is supported by an Else-Kröner Memorial Scholarship.

Reference	Rizos, 2010[214]
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 64. Vukajlovic, 2010[271]

Reference	Vukajlovic, 2010[271]
Study type	Observational
Recruitment	consecutive
Setting	Elective DC cardioversion
Country	Serbia
Sample size	18 (but measured pre and post CV so 36 data points)
Sample characteristics	Age 33-77; 12 male;
Inclusion criteria	People with AF undergoing electrical DC cardioversion
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	Cardiobip, a portable handheld system for remote monitoring of patients. It has a mobile ECG device that is pocket sized and has two electrodes on the top of the device to connect with the patients' fingers (lead 1), and 3 on the bottom to connect with the patients precordium. 1-3 Cardiobip transmissions were performed 3-7 days before and up to 2 weeks after CV
Gold standard	12 lead ECGs recorded before and after CV, read by 2 expert and blinded readers (adjudicated)
Expertise of index test interpreter	2 expert blinded readers
Simultaneous index/gold vs non simultaneous	Does not appear to be simultaneous; certainly no direct reference to this being the case.
Results	The results below are not based on the main analysis in the paper, which was about concurrence between Cardiobip's reconstructed 12 lead trace and the 12 lead ECG trace *lead by lead* (not relevant to the actual diagnosis, which is made from a general impression of all the 12 leads). However stated in text that of the 36 data points, 22 were in AF on 12 lead ECG and 14 were in SR on 12 lead ECG. Also stated that Cardiobip and 12 lead were in complete concordance for the 22 deemed in AF by 12 lead (sensitivity 1) and similarly both were in complete concordance for the 14 deemed in SR by 12 lead (specificity 1). Therefore: Sensitivity 1, specificity 1; TP 22, FN 0, FP 0, TN 14

Reference	Vukajlovic, 2010[271]
Source of funding	No reports of funding
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 65. Ross, 2018[218]

Reference	Ross, 2018[218]
Study type	Observational
Recruitment	consecutive
Setting	Stroke Unit
Country	Germany
Sample size	798 patients (409 with stroke known to be due to AF and 389 with cryptogenic stroke)
Sample characteristics	Patients with stroke due to AF: 59% female; 81 years; 5% TIA; 95% CVA; NIHSS on admission 7 Patients with cryptogenic stroke: 41% female; 68 years; 12% TIA; 88% CVA; NIHSS on admission 7
Inclusion criteria	All patients on stroke unit – those with stroke due to known or newly diagnosed AF and those with cryptogenic stroke
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	SRAclinic, Apoplex medical Technologies. Stroke Risk Analysis (SRA) – software analysis of every hourly ECG snippet of continuous (non 12 lead) ECG monitoring, and report sent daily to stroke unit. The SRA algorithm first detects the QRS complexes to assess if normal to create an RR interval list for further processing. Based on the R-R intervals and via the use of a Lorenzplot, the algorithm gives one of 5 scores based on risk of AF: 0=SR (very low risk) to 4=very high risk for AF. Two risk score thresholds were tested: 1) 0-1=SR and 2-4=AF, and 2) 0-2 = SR and 3-4=AF.
Gold standard	Patients with stroke due to AF: repetitive 12 lead ECG Cryptogenic stroke: 24 Hour Holter Both evaluated by experienced cardiologists. Blinding not reported.
Expertise of index test interpreter	NA - automated

Reference	Ross, 2018[218]
Simultaneous index/gold vs non simultaneous	Concurrent
Results	**First threshold (0-1=SR and 2-4=AF)** Sensitivity 98 (95.19-99.04), specificity 27(22-32.17) **Second threshold (0-2=SR and 3-4=AF)** Sensitivity 84 (79.08-87.79), specificity 70(64.45-74.97) Raw data (TP, FN, FP, TN) not possible to calculate due to insufficient information provided by the paper
Source of funding	European Union (005-GW02-021A)
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 66. Lin, 2010[153]

Reference	Lin, 2010[153]
Study type	Observational
Recruitment	Consecutive, but paired analysis in that each patient was medically CV or not
Setting	Cardiology outpatients
Country	Taiwan
Sample size	20 people with AF (each with 60 x 6 second tests, each counting as a single test). Therefore 1200 data points (person-tests). Also 10 people with no AF (each with 20 x 15 sec tests, each counting as a single test). Therefore 200 data points (person-tests)
Sample characteristics	AF patients: Age 71.4 (range 50-89 years); AF based on 12 lead ECG Non-AF: Age 71.6 years (range 57-88 years); No AF based on 12 lead ECG
Inclusion criteria	Not reported
Exclusion criteria	Not reported
Index test(s), including number	Wearable and wireless 3-lead ECG device (Medi-Trace 200, Kendall) which can be connected to the user via disposable button electrodes. This can be connected to devices such as a notebook or mobile phone via Bluetooth. Signals then analysed according to 2 algorithms.

Reference	Lin, 2010[153]
of repetitions and duration	Algorithm 1: if the variation of consecutive R-R intervals is >150ms within 6 secs of computation Algorithm 2: if the variation of consecutive R-R intervals is >150ms AND SD of R-R intervals in each 6 second recording is >60 ms within 6 seconds of computation
Gold standard	12 lead ECG interpreted by cardiologists
Expertise of index test interpreter	Not reported
Simultaneous index/gold vs non simultaneous	Does not appear to be simultaneous; no direct reporting of this.
Results	The normal and AF data has not been superimposed as 1) the algorithm used for 'normals' is not reported and 2) the length of tests is different **Algorithm 1 in AF patients (n=1200 person-tests)** Sensitivity 92.83, specificity 0 (TP 1114, FN 78, FP 8, TN 0) **Algorithm 2 in AF patients (n=1200 person-tests)** Sensitivity 93.45, specificity 0 (TP 1135, FN 58, FP 7, TN 0) **Unknown algorithm in people with no AF (n=200 person-tests)** Sensitivity NA; specificity 1 (TP 0, FN 0, FP 0, TN 200)
Source of funding	Aiming For The Top University plan of National Chiao-Tung University
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 67. Fallet, 2019[76]

Reference	Fallet, 2019[76]
Study type	Observational
Recruitment	Consecutive
Setting	Patients referred for catheter ablation

Reference	Fallet, 2019[76]
Country	Switzerland
Sample size	17
Sample characteristics	Age 57 years; 12/17 mean; referred for catheter ablation of cardiac arrhythmia (not all with AF)
Inclusion criteria	Patients undergoing catheter ablation of various arrhythmias
Exclusion criteria	Not reported
Index test(s), including number of repetitions and duration	Wrist-type photoplethysmographic (PPG) device. Composed of 3 LEDs in reflection mode and an embedded 3-axis accelerometer. The PPG collects information on 'wave' and 'inter-beat interval (IBI)' features. Wave features: Adaptive organisation Index, variance of the slope of the phase difference, permutation entropy, fractional spectral radius and spectral purity index. IBI features: mean, SD, median, IQR, min, max and RMSSD. The actual thresholds used for each are not directly given.
Gold standard	12 lead ECG, interpreted by a team of 'local experts'.
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	Yes – ECG and PPG waveforms were 'temporally aligned'.
Results	**Using 'wave' features of PPG** Sensitivity 99.2, specificity 90.6 **Using 'IBI' features of PPG** Sensitivity 99.5, specificity 89.5 **Using all 'wave' and 'IBI' features of PPG** Sensitivity 99.7, specificity 92.4 Raw data not provided
Source of funding	Swiss NanoTera Initiative, NTF project MiniHolter
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 68. Kvist, 2019[140]

Reference	Kvist, 2019[140]
Study type	Observational
Recruitment	Consecutive
Setting	Entire subset of population of men aged 65-74
Country	Denmark
Sample size	1340
Sample characteristics	Age 69; 100% male; BMI 27.3; self-reported AF 7.9%; DM 10.9%; Hypertension 42.4%; Ischaemic stroke 6.1%; acute MI 6.2%; PAD 2.2%; CABG or PCI 8.3%; COPD 6.8%; never smoked 33.9%; OACs 8.5%; AADs 1.1%; statins 35.6%
Inclusion criteria	Men aged 65-74 in Denmark
Exclusion criteria	None applied
Index test(s), including number of repetitions and duration	1. CT-related single-lead ECG assessed by radiographer (radiograph-CT-ECG). The CT scans were performed with a 320-slice volume CT scanner (Aquilion One, Toshiba Medical Systems, Japan). As the CT scan uses single-lead ECG (extremity lead I) to trigger the processing of the images during diastole, the radiographers were allowed to screen for AF. The average duration of a single-lead ECG recording was 5–10 min. The single-lead ECG recordings could not be stored for later re-evaluation. During the study period, one of eight alternating radiographers examined each single-lead ECG for AF. The radiographers had oral and written training in ECG assessment with a focus on the ECG characteristics of AF. A research nurse trained in cardiology was responsible for the training. The training session consisted of a thorough introduction to the normal ECG, and subsequently an electrocardiographic description of cardiac arrhythmias with emphasis on AF, in particular the identification of no distinct P waves and irregular RR intervals. Furthermore, the training included case-based exercises. During the first 2 weeks of the study, the radiographers had access to supervision by cardiac nurses. The written training material was available for the radiographers throughout the entire screening period. 2. Within a maximum of 1 hour after the CT scan, the participants had a 12-lead ECG recorded (Schiller Cardiovit AT-102, Schiller Cardiovit AT-102 Plus or Philips PageWriter Trim II). The 12-lead ECGs were examined for AF by one of four study nurses. All of the four nurses had training in ECG and experience with patients with AF from working at a cardiology ward for 4–20 years. The nurses had no access to the radiographer's interpretations of the single-lead ECGs, but they did have knowledge about the participant-reported medical history and medication.
Gold standard	Same 12 lead ECG interpreted by 2 independent cardiologists, who examined all of the 12-lead ECG recordings, which were used as the reference standard for the verification of AF. In the case of any disagreements, a consensus was made between the two cardiologists. The cardiologists had no knowledge of the related medical history and the use of medications, and the cardiologists were blinded to the reports from both the radiographers and the nurses.

Reference	Kvist, 2019[140]
Expertise of index test interpreter	Radiographer/nurse
Simultaneous index/gold vs non simultaneous	Not simultaneous – within 1 hour
Results	Radiograph-CT-ECG Sensitivity 60.3(47.7-72), specificity 97.2(96.2-98.1); TP 41 , FN 27 , FP 35 , TN 1235 Nurse 12 lead ECG Sensitivity 97.1(89.8-99.6), specificity 100(99.7-100); TP 66 , FN 2 , FP 0 , TN 1270
Source of funding	This work was supported by the Region of Southern Denmark, the Danish Heart Foundation, the Elitary Research Centre of Individualized Medicine in Arterial Disease (CIMA), the Odense University Hospital, and the Free National Research Councils and Helsefonden. The CT scan and room facilities were provided by the Silkeborg Regional Hospital.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 69. Antonicelli, 2012[6]

Reference	Antonicelli, 2012[6]
Study type	Observational
Recruitment	Consecutive
Setting	Pre-surgical evaluation unit
Country	Italy
Sample size	107
Sample characteristics	Age 66; 57 men/50 women;
Inclusion criteria	Patients enrolled from the pre-surgical evaluation unit in the outpatient day surgery service at the National Research centre in Ancona
Exclusion criteria	None reported

Reference	Antonicelli, 2012[6]
Index test(s), including number of repetitions and duration	3-lead tele-ECG; This was performed using a personal ECG recorder with three wires (CG-7100, Card Guard Scientific Survival Ltd, Rehovot, Israel). Twelve phases were calculated as follows: rhythm leads and leads I, II, III, aVR, aVL, aVF, V1, V2 in the first phase, leads V3 and V4 in the second phase, and leads V5 and V6 in the third phase; 12-lead tele-ECG; This was performed using a portable 12-lead ECG recorder (CG-7000DX-BT, Card Guard Scientific Survival Ltd, Rehovot, Israel). All recordings were performed in the hospital on the same day. The tele-ECG recordings were transmitted from outpatient examination rooms (Day Surgery Service) to the Telemedicine Call Centre of the Division of Cardiology in the same hospital using telephone transmission with specific call centre software (Heartline version 6.5.0.15, Aerotel Medical Systems, Israel). Interpreted of these in blinded manner by 2 cardiologists unaware of study protocol.
Gold standard	Conventional 12 lead ECG interpreted by the same 2 blinded cardiologists. This was performed using a standard ECG recorder (Archimed 42–20, Esaote Biomedical, Florence Italy);
Expertise of index test interpreter	Cardiologist
Simultaneous index/gold vs non simultaneous	No – same day
Results	This study was not designed to assess diagnostic accuracy of detection of AF, and more to evaluate inter-rater agreement between assessors. Nevertheless contains enough data to allow diagnostic accuracy to be assessed. Results difficult to interpret because several rhythm abnormalities were evaluated but appears that for AF there was only 1 case that was picked up by both index tests. It also appears that there were no false positives, giving a sensitivity of 100% and specificity of 100% for both tele-tests. The paper states: “Both tele-ECG recordings correctly diagnosed sinus rhythm in 106 patients and one atrial fibrillation. Thus, rhythm analysis was 100% correct.” **3-lead tele-ECG** Sensitivity 100, specificity 100; TP 1, FN 0, FP 0, TN 106 **12-lead tele-ECG** Sensitivity 100, specificity 100; TP 1, FN 0, FP 0, TN 106
Source of funding	Not reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 70. Lewis, 2011[150]

Reference	Lewis, 2011[150]
Study type	Observational
Recruitment	Consecutive
Setting	Hospital outpatients or inpatients at 2 hospitals in South wales and New York.
Country	UK and USA
Sample size	594
Sample characteristics	Aged >60 years; not specifically patients with cardiac symptoms or diagnoses
Inclusion criteria	Not reported
Exclusion criteria	Not reported
Index test(s), including number of repetitions and duration	The screening technique involves a finger-probe instrument (as used in pulse oximetry) that utilises the principle of photoplethysmography. In the study, each patient's pulse rhythm was assessed by fitting the probe around the tip of their index finger and recording, and storing on a laptop computer, the pulse waveform pattern for 30 seconds. This pattern was then analysed by the specifically developed software, Fast Fourier Transform Analysis, to determine pulse rate variability, and expressed as an index of deviation from normal sinus wave form. As the pulse in AF is classically 'irregularly irregular', this formed the basis for detecting AF. During the study, the interpretation of records was undertaken later, although 'blinded' to the results of pulse palpation and electrocardiography. Single reading performed.
Gold standard	A 12-lead ECG was recorded immediately after the finger probe had been disconnected. Later, the ECG was interpreted by a consultant cardiologist who reported on the presence or absence of AF without knowledge of the patients' histories, their pulse rates or rhythms, or the findings of the finger probe device.
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	No – immediately afterwards
Results	Modifying the threshold to Index=0.2, led to a sensitivity of 100%. There were zero false negatives and 53 (reported as 8.9%) false positives. Sensitivity is definitely 100% at this threshold (must be correct as 0 false negatives always implies a sensitivity of 100%), but specificity incorrectly stated to be 91.1%. This was based on 53 false positive events which were stated to be 8.9%. But 8.9% of

Reference	Lewis, 2011[150]
	what? Had this false positive figure been 8.9% of those WITHOUT AF then this would have implied that 91.1% were true negatives, and so, by definition, a specificity of 91.1 would have been correct. However 53 is actually 8.9% of *594*, which is the entire cohort (both WITH and WITHOUT AF). Thus the specificity is likely to be far lower than 91.1%, as 53 out of a lower denominator than 594 must be more than 8.9%, and so the specificity would be less than 91.1%. However the actual value cannot be known. There is insufficient raw data provided to allow calculation of TP, etc. (e.g. no numbers with AF). False positives and false negatives given at other indices (0.25 and 0.30) but again the figures prevent us knowing the true sensitivity and specificity.
Source of funding	The study was funded by Melys AFS Ltd and by the authors
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 71. Poon, 2005[195]

Reference	Poon, 2005[195]
Study type	Observational
Recruitment	Consecutive
Setting	University teaching hospital (inpatients and outpatients)
Country	UK
Sample size	4297
Sample characteristics	No information given, apart from the fact that the 4297 ECGs had been taken from inpatients and outpatients over a 3 week period
Inclusion criteria	Not reported
Exclusion criteria	Not reported
Index test(s), including number of repetitions and duration	12 lead ECG interpreted by computer-based rhythm diagnosis (GE Healthcare Technologies MUSE software 005C, version 19)
Gold standard	12 lead ECG, over-read by an experienced electrocardiographer. If there was a discrepancy between the algorithm interpretation and the electrocardiographer interpretation then a second electrocardiographer also looked at the recording and consensus was reached. Clearly not blinded.

Reference	Poon, 2005[195]
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	Simultaneous
Results	Sensitivity 90.8%, specificity 98.9%; TP 227, FN 23, FP 41, TN 3663
Source of funding	None reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 72. Greg, 2008[82]

Reference	Greg, 2008[82]
Study type	Observational
Recruitment	Consecutive
Setting	2 teaching hospitals
Country	USA
Sample size	1785 (1 ECG per patient)
Sample characteristics	Male 1090/1785; age 62 (male) and 63 (female); 109/1785 with AF on gold standard 12 lead testing; no other information given, apart from the fact that the 1785 ECGs had been taken from a random selection of 50000 ECGs collected from 2 teaching hospitals
Inclusion criteria	Not reported
Exclusion criteria	ECGs with extreme artefact and paced rhythm
Index test(s), including number of repetitions and duration	Using the Philips resting 12-lead ECG algorithm, the index tests were 1. Computer interpretation of full 12 lead ECG V_1-V_6 2. Computer interpretation of V_2, V_5 leads information only 3. Computer interpretation of V_1, V_4 leads information only
Gold standard	Full 10 second 12 lead ECG (sampled at 500 samples/sec), over-read by an 2 cardiologists
Expertise of index test interpreter	Automated

Reference	Greg, 2008[82]
Simultaneous index/gold vs non simultaneous	Simultaneous
Results	**Computer interpretation of full 12 lead ECG V1-V6** Sensitivity 89 (82-94), specificity 99 (99-99); TP 97, FN 12, FP 17, TN 1659 **Computer interpretation of V_2, V_5 leads information only** Sensitivity 84 (76-90), specificity 99 (98-99); TP 92, FN 17, FP 17, TN 1659 **Computer interpretation of V_1, V_4 leads information only** Sensitivity 88 (81-93), specificity 99 (98-99); TP 96, FN 13, FP 17, TN 1659
Source of funding	None reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF]

Table 73. Hobbs, 2005[104]

Reference	Hobbs, 2005[104]
Study type	Cross-sectional analysis of diagnostic accuracy data within a large scale RCT of 15,000 people
Recruitment	Consecutive
Setting	50 GP practices in UK
Country	UK
Sample size	2595 ECGs done on 2595 patients
Sample characteristics	Mean age 73.5, 46.9% male; white British 93.2%, white other 2.3%, black African 0.0004%, black Caribbean 3.2%, Chinese 0.1%, Indian 0.9%, Pakistani 0.2%, Asian other 0.1%; AF 8.4%
Inclusion criteria	Random sample of patients from 50 GP practices aged >65
Exclusion criteria	None

Reference	Hobbs, 2005[104]
Index test(s), including number of repetitions and duration	GPs and practice nurses from both intervention practices (who had received education on ECG interpretation) and control practices (who had received no education) were sent ECGs to interpret for the presence or absence of AF. All ECGs recorded within the study were printed off as 12-lead, single-lead thoracic placement or limb-lead recordings. Allocation to ECG type was random and resulted in three equal ECG groups. In order for each interpreter to read all three types of ECG, batches of 100 ECGs were collated with the same numbers of each type of ECG. Allocation to a batch was also random. In total, there were 25 batches of ECGs to match the number of practices in each arm. The GP and practice nurse from the same practice read the same batch of ECGs and each batch was read by one control practice and one intervention practice. Therefore, each ECG was read by two GPs and two practice nurses. All ECGs were anonymised, and practices did not receive any ECGs from their own practice. The interpreters were given a sheet to fill in to indicate for each ECG the presence or absence of AF. A smaller scale process was undertaken with the study cardiologists. They were given a small sample of limb-lead and single-lead ECGs (50 of each) to diagnose in order to calculate diagnostic statistics. All ECGs (as 12- lead) were also analysed by the specific software package accompanying the electronic ECG and results recorded. Pulse palpation was also evaluated, carried out by GPs and nurses. Therefore the index tests were: 1. GP 12 lead 2. GP single thoracic lead 3. GP limb lead 4. Nurse 12 lead 5. Nurse single thoracic lead 6. Nurse limb lead 7. Cardiologists single limb lead 8. Cardiologist limb lead 9. Automated 12 lead 10. Pulse palpation
Gold standard	12 lead ECG interpreted by 2 cardiologists. Where disagreement a third cardiologist made the decision.
Expertise of index test interpreter	GP and nurse
Simultaneous index/gold vs non simultaneous	Yes – all based on the same 12 lead measurements – just portions were used for index tests
Results	Where index test interpreter could not decide on a diagnosis this was given a rating of –ve (=sinus rhythm) **GP 12 lead; sens 79.8(70.9-86.5), spec 91.6(90-92.9)** TP 79, FN 20, FP 114, TN 1241 (n=1454) **GP single thoracic lead; sens 85.4(78.5-90.5), spec 86.4(84.4-88.1)**

Reference	Hobbs, 2005[104]
	TP 112, FN 20, FP 180, TN 1145 (n= 1457) **GP limb lead; sens 82.5(75-88.2), spec88.4(86.6-90)** TP 104, FN 22, FP 156, TN 1202 (n=1484) **Nurse 12 lead; sens 77.1(67.7-84.4), spec 85.1(83-86.9)** TP 74, FN 22, FP 198, TN 1132 (n=1426) **Nurse single thoracic lead; sens 68.7(60.4-75.9), spec 82.7(80.5-84.7)** TP 92, FN 42, FP 222, TN 1066 (n=1422) **Nurse limb lead; sens 73.3(64.6-80.5), spec 83.3(81.2-85.2)** TP 85, FN 33, FP 220, TN 1107 (n=1445) **Cardiologists single limb lead; sens 92.9, spec 98.8** No raw data **Cardiologist limb lead; sens 100, spec 100** No raw data **Automated 12 lead; sens 87.3(82.1-91.2), spec 99.1(98.6-99.4)** TP 179, FN 40, FP 21, TP 2352 **Pulse (by GP or nurse); sens 87.2(82.1-91.1); spec 81.3(79.7-82.8)** TP 190, FN 28, FP 441, TP 1919
Source of funding	HTA funding source
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): Serious [population not only that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF – also contains other people]

Table 74. Langley, 2012[145]

Reference	Langley, 2012[145]
Study type	Derivation and external validation study of algorithms for 12 lead ECG
Recruitment	Consecutive
Setting	Community based cohort from Tanzania
Country	Tanzania
Sample size	The validation database comprised 2124 patients. There was also a derivation database comprising 167 patients from UK, but these were used to derive the thresholds of algorithms and not pertinent to this review.
Sample characteristics	Aged >70; residing in Hai district of Northern Tanzania;
Inclusion criteria	See above
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	12 lead ECG, using the following automated detection algorithms, each based on a short 10s recording, were tested: 1. Based on a co-efficient of variation of the beat intervals (CV). Threshold set at 0.12 2. Based on the mean successive beat interval difference (defined as the mean absolute successive beat interval difference divided by the mean beat interval (Delta). Threshold set at 0.11 3. Based on the co-efficient of sample entropy (COSEn). Threshold set at -1.19
Gold standard	12 lead ECG interpreted by 'expert' and also validated by researcher. Not stated that the ECG was 12 lead, but the machine [GE MAC 1200] is a 12 lead machine, so the assumption has been made that the recordings were 12 lead.
Expertise of index test interpreter	Algorithm
Simultaneous index/gold vs non simultaneous	Yes – all based on the same 12 lead measurements.
Results	**CV algorithm** Sensitivity 90.5%, specificity 89.6% **Delta algorithm** Sensitivity 90.5%, specificity 89.3% **COSEn algorithm**

Reference	Langley, 2012[145]
	Sensitivity 95.2%, specificity 93.4%
Source of funding	Peel Travelling fellowship; no reported conflicts of interest
Limitations	Risk of bias (QUADAS 2 – risk of bias): No serious risk of bias Indirectness (QUADAS 2 - applicability): Serious [population not only that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF – also contains other people]

Table 75. Rhys, 2013[210]

Reference	Rhys, 2013[210]
Study type	Observational
Recruitment	Consecutive
Setting	Single general practice in UK (screening for AF at flu clinic)
Country	UK
Sample size	68 patients with abnormal pulses, from a screening study of 573 people, who were not already diagnosed with AF. The 68 patients with abnormal pulses were all invited to ECG but only 39 attended.
Sample characteristics	Patients
Inclusion criteria	See above
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	1. 12 lead ECG interpreted by algorithm in Cardioview interpretive software (not described) 2. 12 lead ECG interpreted by GP specialty trainee (interpretation done before sent to gold standard interpretation, so effectively blinded to gold standard) The study also looked at pulse measurement but these were not evaluated for diagnostic accuracy because those with normal pulses were not given ECG
Gold standard	12 lead ECG interpreted by 'cardiac physiologist or nurse specialist' with peer review by a cardiologist. Not stated that the ECG was 12 lead, but the machine [Biolog 3000] is a 12 lead machine, so the assumption has been made that the recordings were 12 lead.
Expertise of index test interpreter	Algorithm / SP specialty trainee

Reference	Rhys, 2013[210]
Simultaneous index/gold vs non simultaneous	Yes – all based on the same 12 lead measurements.
Results	**12 lead ECG interpreted by Cardioview algorithm** Sensitivity 100%, specificity 100% TP 2, FN 0, FP 0, TN 30 **12 lead ECG interpreted by GP specialty trainee** Sensitivity 100%, specificity 100% TP 2, FN 0, FP 0, TN 30
Source of funding	Report of no funding
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias Indirectness (QUADAS 2 - applicability): Serious [population not only that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF – also contains other people]

Table 76. Slocum, 1992[237]

Reference	Slocum, 1992[237]
Study type	Observational
Recruitment	Database of rhythms taken from people in AF, in sinus rhythm and people in what was deemed to be an ambiguous rhythm
Setting	Unclear, as based on database of rhythms
Country	USA
Sample size	82 (for validation study, which is the relevant part for this review; the developmental study to develop the algorithm involved 73 different rhythm traces).
Sample characteristics	Not provided
Inclusion criteria	Not reported
Exclusion criteria	Not reported
Index test(s), including number	Algorithm for reading 12 lead ECGs. This first tested for the presence of noncoupled P waves. If noncoupled P waves were detected the rhythm was considered nonatrial fibrillation and no further testing was done. If the rhythm did not have noncoupled P waves, and

Reference	Slocum, 1992[237]
of repetitions and duration	the percent power in each lead II or V1 was >=32% the rhythm was considered AF. This algorithm was derived from the 'training set' of 72 rhythms in the developmental analysis.
Gold standard	12 lead ECG interpreted by a cardiologist.
Expertise of index test interpreter	Automated algorithm
Simultaneous index/gold vs non simultaneous	Yes, same traces used
Results	Algorithm sensitivity 68.3%, specificity 87.8%; TP 28, FN 13, FP 5, TN 36
Source of funding	Not stated
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): Serious [population not only that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF – also contains other people]

Table 77. Hald, 2017[91]

Reference	Hald, 2017[91]
Study type	Observational
Recruitment	Consecutive
Setting	General practices
Country	Denmark
Sample size	87 patients who had irregular pulse on palpation, who were also given ECG by GP/nurse (index test) and ECG by cardiologist (gold standard). The entire study looked at 970 people who were all given pulse palpation. However the larger group of 970 are not considered here because the only people given the gold standard (ECG interpreted by AF specialist) were the 87 with the irregular pulse. Hence the accuracy of pulse palpation is not determinable as we have no gold standard data on those who were negative on pulse palpation.
Sample characteristics	Data not available for subset who had irregular pulse; however for our subset all had irregular pulse on palpation which makes them have a high prevalence of AF (11%)
Inclusion criteria	Any person aged >=65 from the GP practices; no previous AF; presentation was for a genuine medical reason and not for the screening itself; also positive palpation findings, but that is only for the diagnostic accuracy analysis pertinent to this review.
Exclusion criteria	Not reported

Reference	Hald, 2017[91]
Index test(s), including number of repetitions and duration	12 lead ECG carried out and interpreted by GP/nurse
Gold standard	12 lead ECG interpreted by 2 AF specialists
Expertise of index test interpreter	GP/nurse
Simultaneous index/gold vs non simultaneous	Yes, same traces used
Results	GP/nurse 12 lead Sensitivity 100%, specificity 96.1% TP 10, FN 0, FP 3, TN 74 The above had to be derived from the paper as not described directly. Reported that the gold standard result was 10 AF, 77 non AF and that index tests demonstrated 13 AF and 74 non AF. The paper also states that '3 GP suspicions and interpretations of the ECG results… were disapproved by the specialists in representing AF'. This means that there must have been 3 false positives, leaving 10 true positives. Since there were only 10 gold standard positives this implies that there were no false negatives. The rest (n=74) must therefore have been true negatives.
Source of funding	Pfizer Denmark (industry) paid the investigators
Limitations	Risk of bias (QUADAS 2 – risk of bias): No serious risk of bias Indirectness (QUADAS 2 - applicability): Serious [population not only that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF – also contains other people]

Table 78. Himmelreich, 2019[101]

Reference	Himmelreich, 2019[101]
Study type	Observational
Recruitment	Consecutive
Setting	Primary care
Country	Holland
Sample size	219

Reference	Himmelreich, 2019[101]
Sample characteristics	Mean age 64.1; 53.7% male; hypertension 40.7%; DM 30.8%; hypercholesterolaemia 25.2%; known AF or AFL 10.7%; CHD 9.8%; TIA/stroke 6.1%; HF 3.7%; PVD 8.9%; CRF 12.1%; indication for inclusion: 44.4% palpitations, 43.5% other chest symptoms, 21.3% dyspnea, 14.8% lightheadedness 14.8%; fatigue 13%, collapse 2.8%, other 15.7%
Inclusion criteria	Eligible patients were aged 18 years or older who were assigned to 12L-ECG for any non-acute indication as ordered by the local primary care physician in 1 of 10 participating general practices across the Netherlands.
Exclusion criteria	Exclusion criteria were a clinically acute indication for ECG as defined by the local primary care physician (eg, suspicion of acute coronary syndrome) and presence of a pacemaker rhythm on 12L-ECG. We categorized patients according to indication for 12L-ECG either because of presentation with new symptoms (symptom-driven ECG) or as an integral part of protocolized care for primary or secondary prevention of cardiovascular disease (protocol-driven ECG).
Index test(s), including number of repetitions and duration	The KardiaMobile (AliveCor, Inc) is a smartphone-connected, 1L-ECG device that displays ECG recordings in real time (30 seconds) via a smartphone application with a built-in AF detection algorithm. The 1L-ECG recordings were assessed in 2 ways as follows: 1. The AF detection algorithm assessed all 1L-ECG recordings. It classified recordings as either possible AF, normal, or unreadable, or provided no classification. We marked all recordings classified as possible AF as positive for AF. We marked all other algorithm classifications, or when no classification was provided, as negative for AF. The algorithm did not provide a classification for when a 1L-ECG recording was truncated (<30 seconds) 2. 2. Cardiologists (M.L.H., R.N., J.R.dG.) assessed all 1L-ECG recordings in randomized order. The evaluation consisted of scoring each recording for the presence of arrhythmias, ectopic beats, and conduction abnormalities according to a scoring template designed for this study
Gold standard	12 lead ECG interpreted by 2 study cardiologists
Expertise of index test interpreter	Study cardiologists
Simultaneous index/gold vs non simultaneous	Yes
Results	**Automated** Sensitivity 87%, specificity 97.9% TP 20, FN 3, FP 4, TN 187 **Expert** Sensitivity 100%, specificity 100% TP 23, FN 0, FP 0, TN 191

Reference	Himmelreich, 2019[101]
Source of funding	This work was supported by the Netherlands Organisation for Health Research and Development (ZonMw) (80-83910-98-13046). Salary support for Dr Harskamp was provided by a Rubicon fellowship of the Netherlands Organization for Scientific Research (NWO). Dr de Groot is supported by a personal VIDI grant from NWO/ZonMW (016.146.310), reports research grants through his institution from Abbott, Atricure, Boston Scientific, and Medtronic, and received consultancy/speakers fees from Atricure, Bayer, Daiichi Sankyo, Johnson & Johnson, Medtronic, Novartis, and Servier; all outside the scope of this study. All devices and research efforts were paid from university funds. The authors received no funding from the device's producer or local distributor. The authors report no ties to the manufacturer of the investigated device and had full autonomy in the design, conduct, and reporting of this manuscript.
Limitations	Risk of bias (QUADAS 2 – risk of bias): No serious risk of bias Indirectness (QUADAS 2 - applicability): Serious [population not only that defined in protocol – people with cardiovascular risk factors for AF (other than just age) and/or symptoms suggestive of AF – also contains other people]

Table 79. Reverberi, 2019[209]

Reference	Reverberi, 2019[209]
Study type	Observational
Recruitment	Consecutive
Setting	Elective CV clinic
Country	Italy
Sample size	100 – each provided a reading before and after cardioversion. 95 analysed, thus 190 data points derived.
Sample characteristics	Unselected ambulatory patients diagnosed with AF undergoing DC cardioversion; mean age 66.2; 21% female; CHADSVASC 2.3; successful CV 87.4%
Inclusion criteria	Age >18; AF undergoing CV; CHADSVASC >=2;
Exclusion criteria	Pacemaker/automatic internal cardioverter defibrillator
Index test(s), including number of repetitions and duration	RITMIA HR monitor using Bluetooth to communicate with iphone app. 10 minutes. Every patient was monitored with a personal chest belt HR sensor, connected via bluetooth to a dedicated smartphone running the RITMIA app. The data collected by the chest belt HR sensor were analysed in real-time by the algorithm of the RITMIA app (using beat to bear R-R interval data) and directly uploaded and collected for review in the cloud-based server. The automated algorithm classifies each acquired beat as "probable

Reference	Reverberi, 2019[209]
	AF," "unclassified non-AF arrhythmia," or "normal rhythm" and updates the diagnosis second by second. The result is a map of coloured dots plotted on a graph that display time on the x-axis and RR interval (HR) on the y-axis.
Gold standard	12 lead ECG interpreted by 2 blinded cardiologists
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	No – 12L ECG preceded the pre-CV index measure and followed the post-CV index measure
Results	**Automated** Sensitivity 97%, specificity 95.6% TP 96, FN 3, FP 4, TN 87
Source of funding	No funding information. Dr Reverberi is one of the cofounders of theHeartsentinel srl which conceived the RITMIA patent-pending algorithm. All the other authors have no conflicts of interest to declare.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): serious

Table 80. Sabar, 2019[222]

Reference	Sabar, 2019[222]
Study type	Observational
Recruitment	Consecutive
Setting	Cardiology department in secondary care
Country	UK
Sample size	752 (only latter 648 cases used for validation as initial 103 used for refining of the algorithm).
Sample characteristics	Age range 18-97; 51% female; no other information provided
Inclusion criteria	Age >=18; any patient attending the cardiology department for a routine 12 lead ECG or for an outpatient department
Exclusion criteria	Allergies to Velcro or metal used in device; medical condition affecting the wrists that may be interfered with by the attachment of the RhythmPad, such as a fracture necessitating a cast; pacemakers or implantable cardiac devices

Reference	Sabar, 2019[222]
Index test(s), including number of repetitions and duration	6 lead ECG using Rhythm Pad device (1 x 10s). The Rhythm Pad device (Cardiocity, Lancaster, UK) (Figure 1) is a CE-marked medical device that consists of electric potential titanium-based sensors which are placed around both arms of the patient and the right leg, using Velcro straps. The system is attached via leads to a hardware device consisting of a tablet computer that displays and stores the six-lead ECG data. An automated diagnostic report is generated at the same time, using a bespoke algorithm to determine heart rhythm and rate. The Rhythm Pad device does not require the patient to undress or lie flat. The ECG waveform definition is based upon a modified list of 34 data statements that were derived from a list generated by the bespoke analysis algorithm. Data were stored on the Rhythm Pad's hard drive. The Rhythm Pad offers six-lead ECGs from the limb and augmented leads to overcome the low QRS displayed in a single-lead ECG when acquired from the hands. This also overcomes some of the limitations of single-lead ECG systems which can be hampered by poor conductivity attributed to skin condition and a vertical heart alignment. Training for ECG acquisition with the Rhythm Pad is simpler than for a standard 12-lead ECG. As for the ECG interpretation skills, the Rhythm Pad software focuses on rhythm disturbances for which the algorithms are highly accurate when producing the automated diagnoses.
Gold standard	10s 12 lead ECG interpreted by 2 blinded cardiologists
Expertise of index test interpreter	Cardiologists (blinded)
Simultaneous index/gold vs non simultaneous	No – 12L ECG done prior to the index measure
Results	**Expert** Sensitivity 93.85%, specificity 96.84% TP 62, FN 4, FP 18, TN 555 **Automated** Sensitivity 95.38%, specificity 98.77% TP 63, FN 3, FP 7, TN 566
Source of funding	No funding information. The RhythmPad device was provided by the UK-based company CardiocityUKLtd, togetherwith technical support.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias Indirectness (QUADAS 2 - applicability): No serious indirectness

Table 81. Wasserlauf, 2019[275]

Reference	Wasserlauf, 2019[275]
Study type	Observational
Recruitment	Consecutive
Setting	Secondary care
Country	USA
Sample size	Validation cohort of 26 (7500 used as a derivation cohort)
Sample characteristics	All had ICMs previously implanted; age 72.1; female 34.6%; stroke 15.4%; TIA 7.7%; CHF 0%; DM 7.7%; Hypertension 69.2%; CAD 15.4%; prior MI 7.7%; CHADSVASC 2 or more 92.2%; AADs 34.6%; OACs 84.6%
Inclusion criteria	Patients with previously implanted ICMs (Reveal LINQ; Medtronic Inc, Minneapolis, MN) and a history of paroxysmal AF were eligible for enrolment.
Exclusion criteria	None reported
Index test(s), including number of repetitions and duration	Kardia-Band (KB; AliveCor, Mountain View, CA) is a Food and Drug Administration–cleared smartwatch accessory that allows a patient to record a 30-second lead I rhythm strip. Coupled with an investigational application that provides continuous assessment of heart rate, heart rate variability, and activity along with automatic rhythm adjudication, the device has the capability of functioning as a continuous, wearable AF monitor with real-time patient notification that also provides data on AF duration. Watch worn during waking hours (mean 11.3 hrs/day, over a mean of 110 days)
Gold standard	Insertable Cardiac Monitor
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	Yes
Results	**Automated** Duration sensitivity 97.7%, Duration specificity 98.9% The primary outcome was accuracy in detection of AF>1 hr, which is outside the protocol for this review. Moreover the analysis of detection of AF>1 hr did not yield specificity. The results described here were for 'duration accuracy' merely describing the degree of temporal overlap between AF traces on the index and gold standards. For example there were 1101.1 hrs of AF picked up by the index test, out of 1127.1 hours detected by the gold standard, which yielded the value of 97.7%.
Source of funding	No funding information. The RhythmPad device was provided by the UK-based company CardiocityUKLtd, togetherwith technical support.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias Indirectness (QUADAS 2 - applicability): No serious indirectness

Table 82. Cunha, 2019[49]

Reference	Cunha, 2019[49]
Study type	Cross-sectional
Recruitment	consecutive
Setting	Outpatient unit of cardiology unit
Country	Portugal
Sample size	101 undertook accuracy testing (subset of 205 who were part of a larger study)
Sample characteristics	Unclear, as the data provided do not concern the 101 in the diagnostic accuracy study.
Inclusion criteria	Aged >40
Exclusion criteria	Previous diagnosis of atrial fibrillation being medicated with OACs; inability to communicate with the researcher; pacemakers; recent bypass; Wolff-Parkinson-White syndrome
Index test(s), including number of repetitions and duration	Alive-Cor Cardia mobile device.
Gold standard	12 lead ECG, interpreted by a cardiologist
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	No. Timing unclear
Results	**Automated** Sensitivity 90.91, specificity 97.44 TP 20, FN 2, FP 2, TN 76 There were also 29 index traces that were unclassified or unreadable but it was not specified what the corresponding gold standard designation was for these. Thus it was not possible to usefully assign unclassified or unreadable traces to the lower left and lower right cells in the 2x2 table (based on unclassified or unreadable = 'negative index test')
Source of funding	FCT-Foundation for Science and Technology (non-commercial)
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias

Reference	Cunha, 2019[49]
	Indirectness (QUADAS 2 - applicability): Serious indirectness

Table 83. Brown, 2019[24]

Reference	Brown, 2019[24]
Study type	Cross-sectional
Recruitment	consecutive
Setting	Stroke unit
Country	USA
Sample size	265
Sample characteristics	Age 68.4; 57% male; 82% white, 3% Asian, 10% Black, 6% unknown, 4% Hispanic; AF 15%; hypertension 72%; DM 28%; hyperlipidaemia 44%; CAD 16%; CHF 8%; previous stroke 29%
Inclusion criteria	Ischaemic stroke or TIA in 6 bed stroke unit; 18 or over; discharged with diagnosis of acute ischaemic stroke or TIA
Exclusion criteria	Pacemaker
Index test(s), including number of repetitions and duration	Telemetry data from the cardiac monitor (unspecified) of all stroke unit beds that were continually exported to hard drives and then converted to electrocardiomatrix data that were analysed remotely. The electrocardiomatrix used filters and algorithms to produce a colour coded display of the telemetry data that was supposed to be easier to interpret. This visual display was interpreted by study staff for evidence of AF. Median of 46 hours.
Gold standard	Standard telemetry (median 46 hours) analysed by unblended cardiologist
Expertise of index test interpreter	Unclear
Simultaneous index/gold vs non simultaneous	Yes
Results	**Automated** Sensitivity 0.978, specificity 0.864 TP 218, FN 5, FP 5, TN 32
Source of funding	Michigan Translational Research and Commercialisation Grant and T3N grant (both non-commercial)
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias

Reference	Brown, 2019[24]
	Indirectness (QUADAS 2 - applicability): no indirectness

Table 84. Chen, 2020[36]

Reference	Chen, 2020[36]
Study type	Cross-sectional
Recruitment	Unclear
Setting	Inpatients and outpatients in cardiovascular department
Country	China
Sample size	401
Sample characteristics	197 female, 204 male AF/no AF: age 70.4/59.3; hypertension 47.3%/42.2%; CHD 17.3%/26.3%
Inclusion criteria	>18 years; stable heart rhythm at time of study
Exclusion criteria	Situations where wristband could not be used such as bilateral UL disabilities, wrist colour 'abnormalities', severe occlusive disease, or significant UL oedema; implanted pulse generator
Index test(s), including number of repetitions and duration	Amazfit Health band – a wearable wristband device that combines a single channel ECG recorder with a high precision PPG optical sensor. Works with a smartphone application via Bluetooth. A single lead ECG is recorded for 60seconds when initiated by the wearer, and transmitted to a smartphone and then to an AI algorithm on an internet server. PPG signal is then acquired for 71 seconds and evaluated using an AI algorithm on the writsband; if suspected AF is detected there is a repeated PPG test and two tests 3 minutes apart as deemed to be AF. If the second is negative 'no AF' is designated.
Gold standard	12 lead ECG read by an experienced senior ECG physician
Expertise of index test interpreter	NA - automated
Simultaneous index/gold vs non simultaneous	No
Results	**Wristband PPG** **Automated** Sensitivity 0.88, specificity 0.992 TP 132, FN 18, FP 2, TN 249

Reference	Chen, 2020[36]
	Note that unclear readings taken as no AF in our analysis **Wristband ECG** **Automated** Sensitivity 0.873, specificity 1.00 TP 131, FN 19, FP 0, TN 251 Note that unclear readings taken as no AF in our analysis **Combined PPG and ECG (unclear how these were combined)** **Automated** Sensitivity 0.80, specificity 0.968 TP 120, FN 30, FP 8, TN 243 The ECG records were also evaluated by ECG physicians and an electrophysiologist: the sensitivity was 0.9667 and specificity was 0.981. However this does not reflect how the device would be used in practice.
Source of funding	Anhui Huami Information Technology Co Ltd.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): no indirectness (cardiovascular population)

Table 85. Diamantino, 2020[59]

Reference	Diamantino, 2020[59]
Study type	Cross-sectional
Recruitment	Appears to be consecutive
Setting	Primary care cardiovascular screening clinic
Country	Brazil

Reference	Diamantino, 2020[59]
Sample size	334
Sample characteristics	Data only available for sample that were +ve on AFSD. Age 61.8, female 50.5%, hypertension 72.2%, HF 40.2%, CAD 22.7%, major HD on standard echo 76.1%
Inclusion criteria	Unclear, but would need to come from an area of Brazil conducting cardiovascular screening; awaiting echo screening
Exclusion criteria	Unclear
Index test(s), including number of repetitions and duration	Atrial Fibrillation Screening Device, incorporating a 1 lead ECG recording. Patients held the AFSD in a steady seating position for 1 minute with both hands.
Gold standard	12 lead ECG read by experienced cardiologists blinded to AFSD results. This was done after AFSD screening.
Expertise of index test interpreter	NA - automated
Simultaneous index/gold vs non simultaneous	No
Results	**TP 37, FN 4, FP 47, TN 246** **Sensitivity: 0.902(0.77 – 0.973)** **Specificity: 0.84(0.793-0.88)**
Source of funding	This study was funded by Edwards Lifesciences Foundation, USA. The AFSD devices were purchased by the project, and the manufacturer did not have any relationship with the conduct of the study, the collection, analysis, and interpretation of the data.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): no indirectness (waiting list for echo)

Table 86. Karunadas, 2020[126]

Reference	Karunadas, 2020[126]
Study type	Cross-sectional

Reference	Karunadas, 2020[126]
Recruitment	Unclear
Setting	Department of cardiology
Country	India
Sample size	141
Sample characteristics	Age of the patents ranged from 9 years to 77 years with maximum number of patients in the age group of 40-60 (mean age 44.41 years, SD 19.409). Majority were females (n ¼ 74, 52.5%).
Inclusion criteria	Patients who needed AECG monitoring as part of their clinical workup and who consented for simultaneous evaluation with the two AECG systems were included.
Exclusion criteria	Critically ill patients, those with implanted devices like permanent pacemaker or implantable cardioverter defibrillator were excluded.
Index test(s), including number of repetitions and duration	Android App basedWebCardio using WiPatch is an ambulatory ECG system which records ECG in two leads for 72 h. Patients had WiPatch applied in the left upper part of chest after skin preparation. The patch was applied immediately after the connection of conventional Holter. Only data from the 24hours simultaneous with Holter were used for analysis. Analysis performed by proprietary software by a qualified technician.
Gold standard	Conventional 24 hour Holter using either a 3 channel or a 12 channel recorder (Hanix- DL- 820/Hanix 820-DL pro). Soft gel adhesive electrodes were placed on the chest using the Mason Likar system after preparation of chest and the leads of the Holter recorder were connected. The recorder was secured to the body using a belt and the wires fixed using adhesive tape to minimise the movement and artefacts. Note that this was a comparison study where Holter was not designated as the gold standard. However this study is still eligible as 2x2 tables of data were provided, which allowed calculation of accuracy data when making the assumption that Holter is the gold standard.
Expertise of index test interpreter	NA - automated
Simultaneous index/gold vs non simultaneous	Yes
Results	**TP 3, FN 0, FP 0, TN 138** **Sensitivity: 1.0** **Specificity: 1.0**
Source of funding	Not reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): no indirectness

Table 87. Lyckhage, 2020[160]

Reference	Lyckhage, 2020[160]
Study type	Prospective cohort
Recruitment	Consecutive
Setting	Primary care
Country	Denmark
Sample size	366
Sample characteristics	Age 70; 34.4% female; 3.9 years since stroke; >1 clinical stroke or TIA 28.4%; CHADSVASC 4; IHD 7.4%; HF 0.3%; hypertension 69.7%; DM 13.7%; KD 3.6%
Inclusion criteria	AF-naive, had ischaemic stroke over 1 year before enrolment and were older than 49 at stroke onset. Participants with an acute infection or surgery were included at least 1 month after remission. Participants taking OAC for other indications than AF were included.
Exclusion criteria	Participants with a systemic infection or taking antiarrhythmic drugs (class I and III, digoxin, flecainide, and non-dihydropyridine calcium-channel blockers), who had cECG within 1 year before inclusion, and who had an implanted loop recorder, cardioverter defibrillator or pacemaker were not eligible.
Index test(s), including number of repetitions and duration	12 lead ECG, performed minutes before or after application of Holter equipment (number and duration not given but appears to be once) Pulse palpation – radial pulse for at least 20s. Unclear when this was done
Gold standard	7 day Holter (mean use 6.9 days), using Pathfinder SL software for automatic AF detection, and adjudicated by primary investigator and 2 specialist raters
Expertise of index test interpreter	Unclear
Simultaneous index/gold vs non simultaneous	Unclear – appears to have occurred for some 12 lead ECG.
Results	**12 lead ECG** **TP 3, FN 14, FP 0, TN 349; sen 0.176, spec 1.0** **Pulse palpation** **TP 8, FN 9, FP 115, TN 234; sen 0.47, spec 0.67**

Reference	Lyckhage, 2020[160]
Source of funding	The study was supported by Bayer, Boehringer Ingelheim, the Department of Neurology, Zealand University Hospital, 'Region Sjællands Ordinære pulje', 'Grosserer L.F. Foghts Fond', 'A.P. Møller og Hustru Chastine Mc-Kinney Møllers Fond til almene Formaal-Fonden til Lægevidenskabens Fremme' and 'Hans og Nora Buchards Fond'.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias Indirectness (QUADAS 2 - applicability): no indirectness

Table 88. Rajakariar, 2020[201]

Reference	Rajakariar, 2020[201]
Study type	Prospective multicentre validation study
Recruitment	Consecutive
Setting	Tertiary University hospitals
Country	Australia
Sample size	200
Sample characteristics	No AF/AF: age 64/76; male 64%/52%; IHD 32%/50%; hypertension 51%/50%; HF 13%/44%; DM 20%/25%; stroke/TIA 16%/17%; known AF 9%/95%
Inclusion criteria	Patients≥18 years of age admitted to the medical, cardiac or intensive care ward
Exclusion criteria	Patients with cardiac implantable electronic devices, those unable to independently use the device, or in contact isolation were excluded from the study
Index test(s), including number of repetitions and duration	Alive-Cor KardiaBand. The KB strap was attached to an Apple Watch and paired with an iPhone 6 smartphone (Apple, Cupertino, California, USA) using the AliveCor Kardia application V.5.0.2 (AliveCor, Mountain View, California, USA). The device obtains a 30-second continuous lead-I recording that can be viewed in real-time on the iPhone and is remotely transmitted to a secure server for storage and subsequent clinician analysis.
Gold standard	A 12-lead ECG was performed immediately following the KB trace. Interpreted by a cardiologist
Expertise of index test interpreter	automated

Reference	Rajakariar, 2020[201]
Simultaneous index/gold vs non simultaneous	Not simultaneous
Results	**TP:36 FN:2 FP:29 TN:133** **Sen:0.944, spec 0.819**
Source of funding	This work was supported by the Eastern Health Foundation Research Grant [EHFRG2017_029]. The sponsor had no role in study design, collection, analysis, interpretation of data and in the decision to submit the article for publication.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): serious indirectness (included all people admitted to medical and ICU)

Table 89. Zwart, 2020[295]

Reference	Zwart, 2020[295]
Study type	Pragmatic prospective cohort study
Recruitment	Consecutive
Setting	Outpatient Geriatric Clinic
Country	Netherlands
Sample size	439
Sample characteristics	Age 78.4, female 54.4%, hypertension 63.3%, DM 22.3%, CHADSVASC 3.8, HASBLED 1.5, any stroke 15.5%, HF 11.2%, IHD 21.6%
Inclusion criteria	All consecutive patients aged ≥ 65 years at the outpatient geriatric clinic, memory clinic, or Fall and Syncope day clinic (FSC)
Exclusion criteria	Patients with pacemakers or implantable cardioverter defibrillators (ICD) or patients unable or unwilling to provide informed consent were excluded

Reference	Zwart, 2020[295]
Index test(s), including number of repetitions and duration	MyDiagnostik, a single lead ECG device. The measurement was repeated on an average of 3.5 occasions (coinciding with repeated visits to the department
Gold standard	12 lead ECG on study entry. However also stated that a 'confirmatory ECG' was done after each positive index test result
Expertise of index test interpreter	Expert cardiologists
Simultaneous index/gold vs non simultaneous	Not simultaneous
Results	Not possible to calculate raw data because insufficient and unclear information given, and what information was provided did not tally with the accuracy results. Sen: 0.90, spec 0.99
Source of funding	Not reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias Indirectness (QUADAS 2 - applicability): serious indirectness (included all geriatric patients)

Table 90. Lai, 2020[144]

Reference	Lai, 2020[144]
Study type	Prospective cohort study
Recruitment	Consecutive
Setting	Department of Cardiovascular Ultrasound and Cardiology
Country	China
Sample size	40 for data relevant to review
Sample characteristics	Age 68, female 5%, AF patients (35 persistent, 2 paroxysmal and 18 SR) having prior ablation
Inclusion criteria	All consecutive patients with a history of paroxysmal or persistent AF
Exclusion criteria	Patients with pacemakers or defibrillators

Reference	Lai, 2020[144]
Index test(s), including number of repetitions and duration	Single lead (MP1*) patch-based ambulatory ECG monitor worn for 24 hours; This used an automated AF detection algorithm on the basis of a convoluted neural network. *MP1 = single lead patch placed at the 3-4th intercostal space on the midline of the clavicle, 45 degrees from the upper right to the lower left
Gold standard	12 lead Holter ECG for 24 hours; blinded annotations of two clinicians
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	Simultaneous
Results	MP1 electrode position: sensitivity: 0.931, specificity: 0.934
Source of funding	National Natural Science Foundation of China
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): serious indirectness (included patients post ablation)

Table 91. Osca Asensi, 2020[184]

Reference	Osca Asensi, 2020[184]
Study type	Prospective cohort study
Recruitment	Not reported
Setting	Cardiology outpatient clinic
Country	Spain
Sample size	167
Sample characteristics	SR/SF: age 54/67; hypertension 39%/54%; DM 10%/19%; OACs 54%/85%

Reference	Osca Asensi, 2020[184]
Inclusion criteria	Patients aged >18 referred to a cardiology department for cardioversion for AF or for a general consultation (SR or AF)
Exclusion criteria	Atrial flutter or implanted pacemaker
Index test(s), including number of repetitions and duration	Rithmi heart rhythm monitor: wrist monitor using PPG and ECG lead. One repetition for 3 minutes in seating
Gold standard	12 lead ECG read by 2 expert cardiologists
Expertise of index test interpreter	Automated
Simultaneous index/gold vs non simultaneous	Not simultaneous; ECG done first
Results	PPG algorithm: sensitivity: 0.91, specificity: 0.96 ECG algorithm: sensitivity: 0.94, specificity: 0.96 Raw data not available, nor possible to calculate
Source of funding	Not reported
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious risk of bias Indirectness (QUADAS 2 - applicability): serious indirectness (included patients for cardioversion and therefore known to be in AF)

Table 92. Guan, 2020[86]

Reference	Guan, 2020[86]
Study type	Cross-sectional study
Recruitment	Random
Setting	Community

Reference	Guan, 2020[86]
Country	China
Sample size	1479
Sample characteristics	Male: 51.7%; hypertension: 86.9%; DM: 27.4%; history of stroke: 19.3%
Inclusion criteria	Aged >50 years
Exclusion criteria	Tremors; unable to use index device properly
Index test(s), including number of repetitions and duration	Snap ECG – portable single lead (blinded). This is an intelligent palmar portable ECG home monitor device. It can provide a single timepoint (1 minute) single-lead ECG tracing when participants thumbs are on the electrodes in sitting. Readings sent to cardiologist via Bluetooth.
Gold standard	12 lead supine ECG (10s) read by 1 cardiologist (blinded)
Expertise of index test interpreter	Cardiologist
Simultaneous index/gold vs non simultaneous	Not simultaneous; <2 hour delay
Results	Sensitivity: 0.65(0.41-0.85); specificity: 0.99(0.99-1.00); no raw data available
Source of funding	National Natural Science Foundation of China, Natural Science Foundation of Jiangsu Province and Jiangsu Commission of Health
Limitations	Risk of bias (QUADAS 2 – risk of bias): Serious risk of bias Indirectness (QUADAS 2 - applicability): serious indirectness (included patients with no symptoms)