National Institute for Health and Care Excellence

Draft for consultation

Atrial fibrillation

Detection diagnostic accuracy

NICE guideline

Diagnostic evidence review

September 2020

Draft for consultation

Developed by the National Guideline Centre, hosted by the Royal College of Physicians

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and, where appropriate, their carer or guardian.

Local commissioners and providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

ISBN

Contents

1 Detection diagnostic accuracy

1.1 Review question: What are the most accurate methods for detecting atrial fibrillation in people with cardiovascular risk factors for AF and/or symptoms suggestive of AF?

1.2 Introduction

Please see Evidence review A.

1.3 PICO table

For full details see the review protocol in Appendix A:.

Table 1: PICO characteristics of review question

Population	People aged over 18 with symptoms suggestive of AF (including breathlessness, palpitations, syncope/dizziness, and chest discomfort) and/or with cardiovascular risk factors for AF (including TIA, stroke, Heart Failure, hypertension, valve disease). Departures from this population are allowed, but the evidence will be downgraded for indirectness.
Target condition	Atrial fibrillation

Index test(s)	Any point of care tests used to detect AF
	For example (non-exhaustive list):
	• Manual pulse checking
	• Pulse oximeters
	• US devices
	• Blood pressure monitors
	  o Microlife BPM
	  o Watch BP Home A
	• Non-portable (but non-12 lead) ECG devices
	• Portable ECG devices
	  o My Diagnostick
	  o AliveCor Kardia
	• Smart portable devices e.g., phones, watches
	• 12 lead ECG (when gold standard is long-term loop recording – see section below)
	
	Where the same test is used with a differing number of recordings across studies, these should be regarded as separate test strategies, and should thus be dealt with separately.
	
	Tests using differing periods of recording will also be dealt with separately. For example, pulse oximeters for 2 minutes will be in a separate category of index test to pulse oximeters used for 1 hour, and they could be compared to each other as separate index tests.

Reference standard(s)	The reference standard that is used will determine the type of AF that the measured accuracy relates to. The analyses will therefore be stratified by the reference standards used, as follows:

	1. 12 lead ECG, adjudicated by an expert clinician (usually cardiologist). This will theoretically pick up all constant AF but only a small proportion of intermittent AF cases. It is therefore really only useful for determining how well an index test can pick up constant AF. 2. Ambulatory monitoring for >24 hrs [any device that gives a long term recording]. These should pick up all forms of AF. It is therefore a useful way gold standard for determining how well a test can pick up any AF. The ability of the tests to pick up AF vs no AF is being evaluated in this review, not the ability to differentiate between persistent and paroxysmal.
Outcomes	Diagnostic accuracy – sensitivity and specificity
Study design	Cross-sectional observational

1.4 Methods and process

This evidence review was developed using the methods and process described in Developing NICE guidelines: the manual.[156] Methods specific to this review question are described in the review protocol in Appendix A:.

1.5 Clinical evidence

1.5.1 Included studies

Sixty five studies were included in this review.[5, 6, 21, 22, 24, 46, 53, 56, 68, 69, 71, 73, 78, 79, 84, 89, 92, 104, 110, 113, 116, 117, 122, 124, 127, 132, 135, 138, 142, 143, 145, 146, 152, 154, 159, 167, 176-178, 188-190, 194, 198-200, 202, 213, 217, 220, 223, 233, 238, 245, 248, 251, 255, 258-261, 263, 264, 266, 267]

The characteristics of these studies are summarised in Table 2, and evidence from these studies are summarised in the clinical evidence summary (Table 5). Further details are available in the study selection flow chart in Appendix C:, sensitivity and specificity forest plots and receiver operating characteristics (ROC) curves in Appendix E:, and study evidence tables in Appendix D:.

Analysis was stratified by the gold standard used in the studies: 1)12 lead ECG interpreted by an expert (such as a cardiologist or electrophysiologist) or 2) ambulatory monitoring for >24 hours (such as Holter). This stratification was based on the AF that would be detected. 12 lead ECG should detect persistent AF but will only pick up paroxysmal AF during specific intervals of time, and is therefore only a valid gold standard for persistent AF. Ambulatory monitoring for >24 hours may be more useful at picking up AF of both persistent and paroxysmal types and so can be used as a more valid gold standard for any type of AF. Table 2 provides details of the reference standards used.

For each of the above separate strata, pre-hoc sub-grouping strategies (conditional on observed heterogeneity) for any diagnostic test meta-analyses were by

1) expertise of index test interpreter (automated reading / expert reader [such as cardiologist or electrophysiologist] / clinician [clinician such as nurse or GP that was not deemed to be an expert in analysis of ECG traces] / patient).
2) simultaneity of index and reference tests (yes/no)

Sub-grouping was only carried out for the ‘Alive Cor’ test because this was the only analysis where heterogeneity was evident and where there would be at least 3 studies in a sub-group. For the ‘AliveCor’ test, sub-grouping was carried out using the ‘expertise’ strategy and not the ‘simultaneity’ variable because there was evidence from the data that only the former sub-grouping variable could explain the original heterogeneity.

Only 5 diagnostic meta-analyses were possible because at least 3 studies are required for a valid pooling of results, and for most index tests only one or two studies were available. Where diagnostic meta-analysis was possible for a particular test, data from the same study that involved different interpreters were considered as separate data points. Such data were therefore entered alongside each other in the meta-analysis. This was necessary because expertise of examiners had been classified as a 'sub-grouping' (conditional stratification) variable rather than a 'stratification' (unconditional stratification) variable in the protocol. This meant that we could only stratify the meta-analysis by the expertise of interpreters if there was observable heterogeneity in the initial non-stratified analysis. This inclusion of more than one data point from the same study in the meta-analysis was not deemed to be 'double-counting' for two reasons. Firstly, the use of interpreters of different expertise was felt to make data points from the same study sufficiently 'different' to each other to the extent that they could be regarded as being from 'different studies' for the purposes of meta-analysis. Secondly, in many cases the samples of patients used for different interpreters within the same study were different or only overlapped partially.

In the vast majority of studies the unit of analysis was the person being tested, and if AF was detected once in that person then this was counted as a positive test result (regardless of how many times AF was detected in that person using that test) in the 2x2 table. This reflects the purpose of the tests – to find out if a specific patient has AF or not, and as soon as AF has been detected a diagnosis may be made. However in 5 studies[135, 152, 198, 248, 255], the unit of analysis was each of many separate measures done on each person (person-measures). Thus, if AF was detected on several occasions on one person, each event was considered a separate positive test. Since this may influence the strength of overall results, care should be taken with interpretation of these results. Therefore, where such results occur this has been highlighted (sections 1.5.5 and 1.5.6).

Most studies did not include the exact protocol population. For example, some studies contained people without symptoms suggestive of AF. Such studies were included with a quality downgrade for 'indirectness', as stated in the protocol. This flexibility was useful because very few studies were available that exactly met the protocol's population requirements. Furthermore, it was felt that the sensitivity and specificity of the devices would not be greatly influenced by variations in population characteristics, as it was felt implausible that any of these varying characteristics could significantly affect how easy it is to detect AF. It was accepted that different populations would have different prevalence of AF, and that this would therefore affect positive and negative predictive values. However, rather than to directly evaluate predictive values, the clinical aim of this review was to assess the sensitivity and specificity of tests, which independently measure their clinically important ability to differentiate people who have and who don't have the condition. Nevertheless, it was recognised that positive and negative predictive values are of great importance to health economic analysis, and so these will be calculated from the sensitivity and specificity data from the studies in conjunction with established UK prevalence rates (rather than the prevalence rates in individual studies) if tools are found with strong evidence of adequate sensitivity and specificity. Similarly, although 'screening' is outside the remit of this review, diagnostic papers with a reference to screening were included if they contained useful data on the accuracy of tests. The rationale for this is that the determined accuracy of a single device would be similar, whether it is part of a screening strategy or not.

Finally, there were some features of some of the data that should be clarified.

a. Occasionally, papers reported some data from the index test as unclear, and varied in whether they designated this as 'AF' or 'non-AF'. For the purposes of this review, any such data were designated 'non AF', regardless of how the paper designated the data. This approach was taken because this review is about *detection of AF*. If a data point is unclear then AF cannot be detected, so in a binary classification system it can only be designated 'non-AF'. However, if unclear data in papers were only designated as AF, and there was insufficient information in the paper to allow re-calculation, those data were used.

b. Sometimes a paper might have several index test interpreters who were *at the same level of expertise* (for example cardiologist 1, cardiologist 2, etc.) but their data were considered separately. In such cases only the first reported observer was included in this review, to avoid ‘double counting’ of similar data.
c. Destegne, 2017[53] provided data for a sample including people with pacemakers or implanted cardiac monitors and data for a sample with such people excluded. The latter sample was used for this review as people with pacemakers or implanted cardiac monitors were not part of the population in other studies, and had a significant effect on results

1.5.2 Excluded studies

Please see the excluded studies list in Appendix H:.

1.5.3 Summary of clinical studies included in the evidence review (Gold standard = 12 lead ECG stratum)

Table 2: Summary of studies included in the evidence review for detection of atrial fibrillation

Study	Population	Index test(s)	Reference standard
Antonicelli, 2012[5]	107 patients from Italy. Age 66; 57 men/50 women; Inclusion: Patients enrolled from the pre-surgical evaluation unit in the outpatient day surgery service at the National Research centre in Ancona Exclusion: None reported	• 3-lead tele-ECG	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Brito, 2018[21]	127 patients from Switzerland. Age 62; males 64.6%; MI 22.8%; CABG 6.3%; CorAngio 33.9%; valvular Sx 7.9%; sinus at baseline 85% Inclusion: 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; Patients with pacemaker or cardioverter defibrillator	• Beurer ME90 device – a handheld ECG recorder	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Bumgarner, 2018[24]	100 patients from USA. Age 68.2; female 17%; warfarin 32%; DOACs 68%; CV performed 85% Inclusion: 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: Implanted pacemaker; defibrillator	• the Kardia Band (KB) (AliveCor)	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Caldwell, 2012[27]	157 patients from UK. Details not reported Inclusion: Consecutive patients with chronic AF attending the anticoagulation clinic, and consecutive patients with no prior diagnosis of AF attending for a routine ECG Exclusion: None reported	• 6 lead ECG from conventionally positioned limb electrodes (4 limb-leads) • Supine 4-electrode 6-lead frontal plane ECG recording in supine using the prototype recorder	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
		placed on the lower thorax/abdomen • Seated 4-electrode 6-lead frontal plane ECG prototype recording with loosened clothing only	
Cunha, 2019[46]	101 patients from Portugal attending an outpatient cardiology unit. Inclusion: Aged >40 Exclusion: Previous diagnosis of atrial fibrillation being medicated with OACs; inability to communicate with the researcher; pacemakers; recent bypass; Wolff-Parkinson-White syndrome	• AliveCor	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Desteghe, 2017[53]	344 patients from Belgium 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 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: Patients admitted to cardiac wards in a tertiary hospital in Belgium; able to give informed consent Exclusion: Age <18 years, patients in isolation, and those who were unable to hold both devices properly.	• My Diagnostik • AliveCor	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Doliwa, 2009[56]	100 patients from Sweden. Inclusion: Patients with atrial fibrillation, atrial flutter or sinus rhythm recruited from cardiology department. Exclusion: None reported	• Thumb ECG device - Zenecor ECG	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Fallet, 2019[68]	17 patients from Switzerland. Age 57 years; 12/17 mean; referred for catheter ablation of cardiac arrhythmia (not all with AF) Inclusion: Patients undergoing catheter ablation of various	• Wrist-type photoplethysmographic (PPG) device	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	arrhythmias Exclusion: Not reported		
Fan, 2019[69]	112 patients from China. 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: Aged 18 or over Exclusion: Patients unable to use mobile phones and smart bands, with mental or memory problems, or with a pacemaker or implantable cardioverter defibrillator.	• Huawei mate 9 mobile phone – PPG measurements • Huawei Honor 7x mobile phone – PPG measurements • Smart band – Huawei band 2	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Gandolfo, 2015[71]	207 stroke unit inpatients from Italy; 103 women; mean age 77.7 years; 86.5% recent ischaemic CVA/TIA; 13.5% haemorrhagic stroke; within 48 hour window post stroke Inclusion: Patients admitted to stroke unit because of recent (<48 hours) TIA/stroke Exclusion: Patients with rhythm controlled by pacemakers or defibrillators	• Microlife AFib BP device	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Greg, 2008[73]	1785 patients from USA (1 ECG per patient). 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: Not reported Exclusion: ECGs with extreme artefact and paced rhythm	Using the Philips resting 12-lead ECG algorithm, the index tests were 1. Computer interpretation of full 12 lead ECG V1-V6 2. Computer interpretation of V2, V5 leads information only 3. Computer interpretation of V1, V4 leads information only	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Haberman, 2015[78]	130 patients from USA (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)	• AliveCor device	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	Age 59; male 56%; mean HR 72 Inclusion: Ambulatory cardiology patients Exclusion: Not reported		
Hald, 2017[79]	87 from Denmark 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. 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: 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: Not reported	12 lead ECG carried out and interpreted by GP/nurse	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Haverkamp, 2019[84]	94 patients from Norway.37% female; mean age 58; Inclusion: People having ongoing scECG cardiac surveillance who were admitted to the cardiac ward at a university hospital. Exclusion: None reported	• ECG check	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Kaleschke, 2009[104]	508 patients from Germany. 66% male; mean age 61.4; mean BMI 26.6; Inclusion: Clinical indication for 12 lead surface ECG; No other details provided. Exclusion: <18 years; pacemaker or defibrillator	• Omron Heartscan	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Kao, 2018[110]	63 (1 excluded as not fulfilling inclusion criteria) people from Taiwan, recruited from emergency department; age 67; 56% male; AF 29/62 Inclusion: Aged >20 years; either with AF or no AF	• Heart Spectrum Blood Pressure Monitor	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	(diagnosed by 12 lead ECG). Exclusion: People exposed to high frequency surgical equipment during testing' people with cardiac pacemakers or implantable defibrillators; pregnant women		
Kearley, 2014[113]	1000 patients from UK. 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: Participants aged 75 or over, living at home from 6 General practices in the UK Exclusion: People with pacemakers and defibrillators; unable to give consent; terminal illness; other reasons why participation is inappropriate at discretion of GP;	• Watch BP • Omron HCG-801 • Merlin ECG event recorder	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Koltowski, 2019[117]	100 patients from Poland. 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 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: Undergoing regular 12-lead ECG due to standard diagnosis on admission in stable state Exclusion: Need for urgent medical care	• Kardia mobile ECG (Lead I)	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Kristensen, 2016[122]	93 patients from Denmark. 54% male; age 67; IHD 11%; hypertension 54%; DM 21%; known AF diagnosis 36%; Medication affecting heart rhythm 47% Inclusion: 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.	• Portable ECG monitor (3 lead)	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	Exclusion: Patients with severe dementia, mental illness or poor ECG readings		
Kvist, 2019[124]	1340 people from Denmark. 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: Men aged 65-74 in Denmark Exclusion; None applied	• 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	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Langley, 2012[127]	2124 patients on a validation database from Tanzania. 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. Inclusion; Aged >70; residing in Hai district of Northern Tanzania; Exclusion: None reported	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	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Lewis, 2011[132]	594 patients from UK and USA. Aged >60 years; not specifically patients with cardiac symptoms or diagnoses. Inclusion and exclusion criteria not reported.	• Finger-probe instrument (as used in pulse oximetry) that	12 lead ECG, with interpretation by cardiologist

Study	Population	Index test(s)	Reference standard
		utilises the principle of photoplethysmography	/electrophysiologist
Lin, 2010[135]	20 people from Taiwan 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) 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 and exclusion criteria not reported.	• Wearable and wireless 3-lead ECG device (Medi-Trace 200, Kendall)	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Lown, 2018[138]	418 individuals from 3 general practices in UK 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: Aged>=65; from the 3 designated general practices Exclusion: 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.	• Watch BP • Alive Cor • PH7 • FirstBeat Bodyguard	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Mant, 2007[142] and Hobbs, 2005[92]	A random sample of 9866 people from UK 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. Inclusion: 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. Exclusion: None reported	• 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	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
		• 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)	
Marazzi, 2012[143]	550 patients from Italy. Mean age 67 years; 54.3% male; bp 139.8/86.9 Inclusion: Patients referred to hypertension clinic Exclusion: <18 years; pacemaker; implanted defibrillator	• Microlife BP A200 Plus • Omron M6 oscillometric device	12 lead ECG, with interpretation by cardiologist /electrophysiologist
McManus, 2013[146]	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) patients from USA. 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: Patients with persistent AF on a roster of patients scheduled to have elective cardioversion for AF Exclusion: Not reported	• iPhone 4S camera PPG measures on fingertip	12 lead ECG, with interpretation by cardiologist /electrophysiologist
McManus, 2016[145]	128 people from USA. Age 66.2yrs ; non-white 7%; 18% women; hypertension 75.7%; DM 28.2%; CAD 25%; CHF 32.8%; stroke 13.3% Inclusion: 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.	• PPG measures on an iphone 4S	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	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; Not reported		
Nigolian, 2018[159]	52 people from Switzerland. Age 69; male 58%; pacemaker 10%; hypertension 60%; DM 21%; COPD 8%; AF on 12 lead ECG 31%; OACs 40% Inclusion: Consecutive patients admitted to the cardiology department at a University Hospital Exclusion; <18 years; inability or unwilling to consent	• Beurer ME 80 device – a pocket sized (reconstructing 9 lead) ECG device	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Park, 2015[167]	17 patients from South Korea with palpitations. No other details given. Inclusion: Patients with palpitations Exclusion: None reported	• mobile ECG device ER-2000s	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Poon, 2005[176]	4297 ECGs had been taken from inpatients and outpatients in UK over a 3 week period	• 12 lead ECG interpreted by computer-based rhythm diagnosis (GE Healthcare Technologies MUSE software 005C, version 19)	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Proesmans, 2019[178]	223 patients from Belgium. 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: Known paroxysmal or persistent AF; aged >=65; other subjects without a history of AF. Exclusion: Active pacemakers	• Fibricheck app (PPG) • Single lead ECG using ECG-Bone	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Renier, 2012[188]	177 patients from Belgium aged 55 years; 45% men Inclusion: 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.	• Heartscan hand-held device	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	Exclusion: <18 years; unable to use right hand for heartscan device; did not understand language used by HCPs; no consent		
Rhys, 2013[190]	68 patients from UK 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. No inclusion or exclusion criteria reported.	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)	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Rozen, 2018[200]	99 patients from USA (but each patient contributed two sets of data – pre-cardioversion and post-cardioversion). Patients with paroxysmal AF were 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: Consecutive patients with a diagnosis of AF who were scheduled for elective direct current cardioversion. Exclusion: <18 years	• Cardio Rhythm Mobile Application (CRMA) (PPG)	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Slocum, 1992[217]	82 patients from USA (for validation study, which is the relevant part for this review; the developmental study to develop the algorithm involved 73 different rhythm traces).	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 the percent power in each lead II or V1 was >=32% the rhythm was considered	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
		AF. This algorithm was derived from the 'training set' of 72 rhythms in the developmental analysis.	
Somerville, 2000[220]	86 patients from UK. 30% with AF; no other details provided Inclusion: 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: None reported	• Pulse palpation • Bipolar ECGs • 12 lead ECG by non-expert interpreters	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Stergiou, 2009[223]	73 patients from Greece. 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: 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: age <35 years, presence of a pacemaker, and/or an implanted defibrillator and refusal to participate.	• Microlife BPA100 Plus	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Tieleman, 2014[238]	868 people from Holland. Inclusion: Patients visiting the outpatient cardiology clinic, or patients attending 2 GP clinics for influenza vaccination Exclusion: None reported	• My Diagnostik	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Vaes, 2014[245]	191 patients from Belgium. 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: 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	• My Diagnostik	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	history of atrial fibrillation. Exclusion: Pacemaker in active mode		
Vukajlovic, 2010[251]	18 (but measured pre and post CV so 36 data points) people from Serbia. Age 33-77; 12 male; Inclusion: People with AF undergoing electrical DC cardioversion Exclusion: None reported	• Cardiobip, a portable handheld system for remote monitoring of patients	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Wiesel, 2004[261]	450 people from USA contributing to 464 office visits (14 attended twice) 59% men; mean age 69; most common associated medical conditions were hypertension, CAD and DM Inclusion: Unselected outpatients followed by an urban cardiology practice who had an ECG performed during scheduled office visits. Exclusion: None reported	• Omron 712C sphygmomanometer	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Wiesel, 2009[260]	405 patients from USA. Mean age 73; male 51%; white 82%; black 8%; other 10%; CHF 6.7%; Hypertension 51.6%; DM 14.8%; CAD 37.3% Inclusion: Unselected general cardiology outpatients seen for scheduled visits in 2 cardiology centres in NY Exclusion: Pacemakers; defibrillators	• Microlife BP3MQ1-2D BP monitor	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Wiesel, 2014[259]	183 patients from USA. 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: All patients aged >50 attending 2 outpatient cardiology clinics Exclusion: Patients with pacemakers or defibrillators	• Microlife BP A 200 • Omron M6 comfort	12 lead ECG, with interpretation by cardiologist /electrophysiologist
William, 2018[263]	52 participants from USA with 225 sets of measurements 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: Patients with a diagnosis of AF admitted for AAD	• Kardia Mobile Cardiac Monitor	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	therapy; aged 35-85; history of PAF or persistent AF; baseline corrected QT interval <470 or 500 if QRS duration >120ms Exclusion: Patients with pacemakers; patients with defibrillators		
Williams, 2015[264]	99 patients from UK.29 with AF on ECG; other details not reported Inclusion: 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: None reported	• Alive Cor	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Winkler, 2011[266]	60 patients from Germany (details not provided) Inclusion: patients admitted to the cardiology department Exclusion: Not reported	• Handheld ECG device with dry electrodes that records 3 lead ECG (Einthiven I, II and III leads).	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Yan, 2018[267]	233 people from Hong Kong. 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: Patients admitted to the cardiology ward of the hospital for clinical reasons Exclusion: None reported	• Iphone units installed with Cardio Rhythm application for facial and fingertip photoplethysmographic (PPG) measures	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Himmelreich, 2019[89]	219 people from Holland. 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% 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	• KardiaMobile (AliveCor, Inc) - 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.	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
	participating general practices across the Netherlands. 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.		
Reverberi, 2019[189]	100 unselected ambulatory patients from Italy diagnosed with AF undergoing DC cardioversion; mean age 66.2; 21% female; CHADSVASC 2.3; successful CV 87.4% Inclusion: Age >18; AF undergoing CV; CHADSVASC >=2; Exclusion: Pacemaker/automatic internal cardioverter defibrillator	• RITMIA HR monitor using Bluetooth to communicate with iphone app. 10 minutes.	12 lead ECG, with interpretation by cardiologist /electrophysiologist
Sabar, 2019[202]	752 patients from UK attending a cardiology outpatient department for a routine 12 lead ECG or outpatient appointment. Age range 18-97; 51% female; no other information provided Inclusion: Age >=18; any patient attending the cardiology department for a routine 12 lead ECG or for an outpatient department Exclusion: 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	• 6 lead ECG using Rhythm Pad device (1 x 10s). The Rhythm Pad device (Cardiocity, Lancaster, UK) 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.	12 lead ECG, with interpretation by cardiologist /electrophysiologist

Study	Population	Index test(s)	Reference standard
		•	

1.5.4 Summary of clinical studies included in the evidence review (Gold standard = >24 hours ambulatory monitoring stratum)

Table 3: Summary of studies included in the evidence review for detection of atrial fibrillation

Study	Population	Index test(s)	Reference standard
Arevalo-Manso, 2016[6]	76 patients from Spain 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. 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: Age>18 years and having been admitted to the SU for an acute TIA or ischaemic stroke. Exclusion: History of AF	1. a monitor equipped with AF-RS (DASH 5000, General Electric Healthcare, Milwaukee, Wisconsin, USA) 2. Standard ECG monitoring devices without AF-RS.	12-lead ECG and 24 hr Holter ECG
Brown, 2019[22]	265 patients on a stroke unit in USA. Inclusion: Ischaemic stroke or TIA in 6 bed stroke unit; 18 or over; discharged with diagnosis of acute ischaemic stroke or TIA Exclusion: Pacemaker	>24 hrs telemetry with ‘electrocardiomatrix’	>24 hours telemetry
Kollias, 2018[116]	100 patients attending a hypertension clinic in Greece. Age	• Microlife WatchBP O3	24 hour Holter recording

Study	Population	Index test(s)	Reference standard
	70.6; BMI 29.1; 52.9% male; 11% stroke; 85% hypertension; 20% DM; 7% CAD; 82% antihypertensive treatment; CHADSVASC score 3.06 Inclusion: 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: Pacemaker implantation	Afib oscillometric device with measurements programmed at 20-minute intervals for 24 hours.	using the SpiderView (ELA Medical, Sorin Group) multichannel system recorder which was performed simultaneously with 24-hour ABPM.
Mulder, 2012[153]	96 patients from Holland 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: Patients who had undergone PVI 12 months previously for paroxysmal AF Exclusion: None reported	• Holter for 1,2,3,4,5,6 days	7 day Holter
Muller, 2009[154]	48 people from Germany. Mean age 62; 29/48 male; 24 with AF; consecutive patients at an internal medicine department. Inclusion: Presence of an indication for 24 hr Holter ECG Exclusion; Antibradycardic pacemakers; implantable cardioverters and defibrillators	• Vitaphone 3100 BT external loop recorder	24 hours 3 channel ECG (Holter).
Poulsen, 2017[177]	100 patients from Denmark. 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 >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 None reported	• 30s thumb ECG (Zenicor Medical Systems AB) twice daily for 30 days	5 days Holter (Lifecard CF device).
Rizos, 2010[194]	136 patients from Germany 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: Patients > 60 years presenting with an acute	• 6 channel Holter (H12+, Mortara Instruments) performed for 24 hours. • 12-bit resolution digital	Continuous ECG bedside monitoring for duration of stay in stroke unit

Study	Population	Index test(s)	Reference standard
	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: 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	ECG recoding for 1-2 hours.	
Ross, 2018[198]	798 patients (409 with stroke known to be due to AF and 389 with cryptogenic stroke) from Germany. 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: All patients on stroke unit – those with stroke due to known or newly diagnosed AF and those with cryptogenic stroke Exclusion: None reported	• SRAclinic, Apoplex medical Technologies. Stroke Risk Analysis (SRA) – software analysis of every hourly ECG snippet of continuous (non 12 lead) ECG monitoring	Patients with stroke due to AF: repetitive 12 lead ECG Cryptogenic stroke: 24 Hour Holter
Roten, 2012[199]	88 patients from Switzerland (12 patients undergoing ablation included twice, before and after ablation) – therefore 100 datasets 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: 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: Patients with persistent AF; patients unable to handle the devices independently.	• 7 day triggered ECG (R.Test Evolution 3).	7 day continuous Holter (Lifecard CF).
Sejr, 2019[213]	1412 patients from Denmark. 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: 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	• R.Test Evolution 4 (NorDiaTech, Paris, France) was device used as External loop recording (ELR).	Continuous ECG monitoring for 48 hours

Study	Population	Index test(s)	Reference standard
	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: See above		
Velthuis, 2013[248]	153 patients from Holland. 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: Consecutive patients aged >18 years admitted with a provisional diagnosis of acute ischaemic stroke Exclusion: Patients with known history of AF	• 24 hour external loop recorder (single channel device 3100 BT, Vitaphone, Mannheim) using automated settings	24 hour external loop recorder, interpreted by 2 blinded qualified analysts
Wiesel, 2013[258]	160 patients from USA. 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: 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: Pacemakers; defibrillators	• AF-BP monitor for 30 days	Electrocardiographic event monitor (Hearttrack 2) [regarded as a Holter equivalent] for 30 days
Wasserlauf, 2019[255]	26 patients from USA with an implanted cardiac monitor. Mean 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: Patients with previously implanted ICMs and a history of paroxysmal AF were eligible for enrolment. Exclusion: not reported	• Kardia-Band (KB; AliveCor, Mountain View, CA) - smartwatch accessory that allows a patient to record a 30-second lead I rhythm strip. Watch worn during waking hours (mean 11.3 hrs/day, over a mean of 110 days)	Insertable Cardiac Monitor

1 See Appendix D: for full evidence tables.

1.5.5 Quality assessment of clinical studies included in the evidence review

For measurement of imprecision, clinical decision thresholds for sensitivity and specificity were set at 0.90 and 0.60.

<u>STRATUM 1: 12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard</u>

Table 4: Clinical evidence summary: diagnostic test accuracy for mobile ECG devices (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
Alive Cor handheld lead I ECG *1 measure of 30s (for Lown,2018 it lasted a single bp cycle which is assumed to be similar)*	8 Cunha, 2019[46] Desteghe, 2017[53] Desteghe, 2017[53] Haberman, 2015[78] Himmelreich, 2019[89] Himmelreich, 2019[89] Koltowski, 2019[117] Lown, 2018[138] William, 2018[263] William, 2018[263] Williams, 2015[264] Williams, 2015[264]	1544	 auto auto expert expert auto expert expert auto auto expert clinician expert	 no no no no yes yes no no no no yes yes	 0.91[0.71-0.99] 0.55 [0.32, 0.76] 1.00 [0.83, 1.00] 0.94 [0.73, 1.00] 0.87 [0.66-0.97] 1.00 [0.85, 1.00] 0.928[f] 0.88 [0.79, 0.94] 0.71 [0.60, 0.81] 0.94 [0.86, 0.98] 0.90 [0.73, 0.98] 0.93 [0.77, 0.99] **Pooled: 0.91(0.82-0.96)[e]**	 0.97[0.91-1.00] 0.98 [0.95, 0.99] 0.98 [0.95, 0.99] 0.99 [0.95, 1.00] 0.98 [0.95,1.00] 1.00 [0.98, 1.00] 1.000 [f] 0.99 [0.97, 1.00] 0.67 [0.59, 0.75] 0.87 [0.80, 0.92] 0.86 [0.76, 0.94] 0.76 [0.64, 0.85] **Pooled: 0.96(0.90-0.99)[e]**	**Sensitivity**				
							Very serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	No serious inconsistency	no serious imprecision	VERY LOW
	5 Desteghe, 2017[53] Lown, 2018[138] William, 2018[263] Himmelreich, 2019[89] Cunha, 2019[46]		AUTO SUBGROUP		 0.55 [0.32, 0.76] 0.88 [0.79, 0.94] 0.71 [0.60, 0.81] 0.87 [0.66-0.97] 0.91[0.71-0.99] **Pooled: 0.81(0.61-0.92)[e]**	 0.98 [0.95, 0.99] 0.99 [0.97, 1.00] 0.67 [0.59, 0.75] 0.98 [0.95,1.00] 0.97 [0.91-1.00] **Pooled: 0.96(0.83-0.99) [e]**	**Sensitivity**				
							Very serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
							Specificity				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
							Very serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
	6 Desteghe, 2017[53] Haberman, 2015[78] Williams, 2015[264] Koltowski, 2019 [117] William, 2018[263] Himmelreich, 2019[89]		EXPERT SUBGROUP		1.00 [0.83, 1.00] 0.94 [0.73, 1.00] 0.93 [0.77, 0.99] 0.928[f] 0.94 [0.86, 0.98] 1.00 [0.85, 1.00] **Pooled: 0.95(0.88-0.99)** [e]	0.98 [0.95, 0.99] 0.99 [0.95, 1.00] 0.76 [0.64, 0.85] 1.000[f] 0.87 [0.80, 0.92] 1.00 [0.98, 1.00] **Pooled: 0.96(0.81-0.99)** [e]	**Sensitivity**				
							Very serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
	1 Williams, 2015[264]		CLINICIAN SUBGROUP		0.90 [0.73, 0.98]	0.86 [0.76, 0.94]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Kardia band (Alive Cor) watch device (equivalent to Lead I) *1 measure for 30s*	1 Bumgarner, 2018[24] Bumgarner, 2018[24]	169	auto expert	no no	0.69 [0.59, 0.78] 0.88 [0.79, 0.94] **Median[g]: 0.69[0.59, 0.78]**	0.91 [0.82, 0.96] 0.86 [0.76, 0.93] **0.91 [0.82, 0.96]**	**Sensitivity**				
							serious[a]	serious[b]	Serious[d]	Very serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
Beurer ME90 device – lead I ECG	1 Brito, 2018[21]	126	auto	no	0.89 [0.65, 0.99]	0.62 [0.52, 0.71]	**Sensitivity**				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
1 measure of 30s	Brito, 2018[21]		expert	no	0.84 [0.60, 0.97] **Median[g]: 0.84[0.60, 0.97]**	1.00 [0.97, 1.00] **1.00[0.97, 1.00]**	Very serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Serious[d]	No serious imprecision	VERY LOW
Beurer ME90 device – lead I and mv4 leads *1 measure of 30s*	1 Brito, 2018[21]	126	auto	no	0.88 [0.64, 0.99]	0.84 [0.76, 0.91]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Beurer ME90 device – mv4 lead *1 measure of 30s*	1 Brito, 2018[21] Brito, 2018[21]	126	auto expert	no no	0.94 [0.71, 1.00] 0.84 [0.60, 0.97] **Median[g]: 0.84[0.60, 0.97]**	0.76 [0.67, 0.84] 1.00 [0.97, 1.00] **1.00 [0.97, 1.00]**	**Sensitivity**				
							Very serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Serious[d]	no serious imprecision	VERY LOW
Beurer ME 80 device – a pocket sized (reconstructing 9 lead) ECG device. *1 measure of unknown*	1 Nigolian, 2018[159] Nigolian, 2018[159]	52	clinician expert	no no	0.75 [0.48, 0.93] 1.00 [0.79, 1.00] **Median[g]: 0.75 [0.48, 0.93]**	0.89 [0.74, 0.97] 0.94 [0.81, 0.99] **0.89 [0.74, 0.97]**	**Sensitivity**				
							serious[a]	serious[b]	No serious inconsistency	Very serious[c]	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
duration							**Specificity**				
							serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
ECG Check, an FDA-approved mobile heart monitor manufactured by Cardiac Designs. By putting two fingers on the ECG Check, it registers a lead I ECG *1 measure of 30s*	1 Haverkamp, 2019[84]	94	auto	no	1.00 [0.72, 1.00]	0.94 [0.86, 0.98]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Merlin ECG event recorder (single lead) *1 measure of 30s*	1 Kearley, 2014[113]	1000	expert	no	0.939 [f]	0.901[f]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
My Diagnostik handheld lead I ECG *1 measure of 60s*	2 Desteghe, 2017[53] Desteghe, 2017[53] Tieleman, 2014[238]	1125	auto expert auto	no no no	0.82 [0.60, 0.95] 0.85 [0.62, 0.97] 1.00 [0.97, 1.00] **Pooled:0.94(0.52-0.99)** [e]	0.94 [0.91, 0.97] 0.95 [0.92, 0.98] 0.98 [0.97, 0.99] **Pooled:0.97(0.85-0.99)** [e]	**Sensitivity**				
							serious[a]	serious[b]	No serious inconsistency	Very serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	No serious inconsistent	Serious[c]	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
									ncy		
My Diagnostik handheld lead I ECG *3 measures of 60s (majority rule)*	1 Vaes, 2014[245]	181	auto	no	0.94 [0.87, 0.98]	0.93 [0.85, 0.97]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Omron Heartscan HCG 801 E single lead device *1 measure of 30s (though duration not stated in Kearley, 2014[113])*	3 Kearley, 2014[113] Kearley, 2014[113] Renier, 2012[188] Renier, 2012[188] Kaleschke, 2009[104]	1684	auto expert auto clinician expert	no no no no no	0.99 [0.93, 1.00] 0.944 [f] 0.92 [0.64, 1.00] 0.69 [0.39, 0.91] 0.99(0.96-1.00)[f] **Pooled:0.93(0.50-0.99)** [e]	0.76 [0.73, 0.79] 0.946 [f] 1.00 [0.98, 1.00] 0.95 [0.90, 0.97] 0.96(0.94-0.98)[f] **Pooled:0.95(0.52-0.99)** [e]	**Sensitivity**				
							serious[a]	serious[b]	No serious inconsistency	Very serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Serious[d]	Very serious[c]	VERY LOW
ECG Bone – single lead *Unclear reps and duration*	1 Proesmans, 2019[178]	223	expert	no	0.90(0.824-0.951)	0.968(0.919-0.991)	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	no serious imprecision	VERY LOW
Zenicor ECG thumb device (bipolar lead I) *1 measure of 10s*	1 Doliwa, 2009[56]	100	expert	no	0.96 [0.86, 1.00]	0.92 [0.81, 0.98]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not	Serious[c]	VERY

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
									applicable		LOW
Polar H7 heart rate monitor – heart rate sensor derives ECG data via chest electrodes. Can detect RR intervals accurately *1 measure over duration of a bp measurement cycle*	1 Lown, 2018[138]	418	auto	no	0.96 [0.90, 0.99]	0.98 [0.96, 0.99]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	no serious imprecision	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	no serious imprecision	VERY LOW
RITMIA HR monitor *1 measure over 10 minutes*	1 Reverberi, 2019[189]	100	auto	no	0.97 [0.91, 0.99]	0.96 [0.89, 0.99]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Firstbeat Bodyguard 2 – delivers single lead data *1 measure over duration of a bp measurement cycle*	1 Lown, 2018[138]	418	auto	no	0.96 [0.90, 0.99]	0.99 [0.97, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	no serious imprecision	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
							Specificity				
							Very serious[a]	serious[b]	Not applicable	no serious imprecision	VERY LOW
Cardiobip, a portable 5 lead handheld ECG system for remote monitoring of patients. *1-3 measures of unknown duration*	1 Vukajlovic, 2010[251]	36	expert	no	1.00 [0.85, 1.00]	1.00 [0.77, 1.00]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Mobile ECG device ER-2000s. Mode 1 uses three ECG electrodes that are attached to the anterior chest wall. patients were instructed to push the record button when they believed they were experiencing a cardiac symptom. *1 measure of unknown duration*	1 Park, 2015[167]	17	expert	yes	1.000[f]	1.000[f]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	Not assessed	LOW
							Specificity				
							Very serious[a]	No serious indirectness	Not applicable	Not assessed	LOW
Mobile ECG device ER-2000s. Mode 2 uses the side chest channel and finger channel. Patients were instructed to push the record button when they believed they were experiencing a cardiac	1 Park, 2015[167]	17	expert	yes	1.000[f]	1.000[f]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	Not assessed	LOW
							Specificity				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
symptom. *1 measure of unknown duration*							Very serious[a]	No serious indirectness	Not applicable	Not assessed	LOW
Huawei band 2 smartband *1 measure of 3mins*	1 Fan, 2019[69]	624	auto	yes	0.92 [0.88, 0.95]	1.00 [0.98, 1.00]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.

(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect

(c) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.

(d) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.

(e) Pooled sensitivity/specificity from diagnostic meta-analysis

(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses

(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given

Table 5: Clinical evidence summary: diagnostic test accuracy for blood pressure monitors (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
Microlife BP3MQ1-2D oscillometric device *3 readings, with 'majority rule' of 10 beat intervals*	2 Gandolfo, 2015[71] Wiesel, 2009[260]	612	auto auto	no no	0.89 [0.75, 0.97] 0.97 [0.91, 0.99] **Median[g]: 0.89 [0.75, 0.97]**	0.99 [0.96, 1.00] 0.89 [0.85, 0.92] **0.99 [0.96, 1.00]**	**Sensitivity**				
							serious[a]	serious[b]	No serious inconsistency	Serious[d]	VERY LOW
							Specificity				
							serious[a]	serious[b]	serious[c]	No serious imprecision	VERY LOW
Microlife BP3MQ1-2D oscillometric device *Single readings of 10 beat intervals*	1 Wiesel, 2009[260]	1215	auto	no	0.95 [0.92, 0.97]	0.86 [0.84, 0.89]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
Microlife BPA 200 oscillometric device *3 readings, with majority rule of 10 beat intervals*	2 Marazzi, 2012[143] Wiesel, 2014[259]	686	auto auto	yes no	0.92 [0.85, 0.97] 1.00 [0.88, 1.00] **Median[g]: 0.92 [0.85, 0.97]**	0.97 [0.95, 0.98] 0.92 [0.87, 0.96] **0.97 [0.95, 0.98]**	**Sensitivity**				
							Serious[a]	No serious indirectness	No serious inconsistency	serious[d]	LOW
							Specificity				
							Serious[a]	No serious	No serious	No serious imprecisio	MOD

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
								indirectness	inconsistency	n	
Microlife BPA100 Plus BP oscillometric device *3 readings (majority rule) of 10 beat intervals*	1 Stergio, 2009[223]	72	auto	yes	1.00 [0.87, 1.00]	0.89 [0.76, 0.96]	**Sensitivity**				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
							Specificity				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
Microlife BPA100 Plus BP oscillometric device *3 readings (minority rule) of 10 beat intervals*	1 Stergio, 2009[223]	72	auto	yes	1.00 [0.87, 1.00]	0.69 [0.53, 0.82]	**Sensitivity**				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
							Specificity				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
Microlife BPA100 Plus BP oscillometric device *1st reading only of 10 beat intervals*	1 Stergio, 2009[223]	72	auto	yes	0.93 [0.76, 0.99]	0.89 [0.76, 0.96]	**Sensitivity**				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
							Specificity				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
Microlife BPA100 Plus BP oscillometric device *1st 2 readings of 10 beat intervals*	1 Stergio, 2009[223]	72	auto	yes	1.00 [0.87, 1.00]	0.76 [0.60, 0.87]	**Sensitivity**				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
							Specificity				
							Serious[a]	No serious indirectness	Not applicable	No serious imprecision	MOD
Microlife Watch BP oscillometric device *Readings over up to 3 bp measurement cycles*	2 Kearley, 2014[113] Lown, 2018[138]	1417	auto auto	no no	0.95 [0.88, 0.99] 0.96 [0.90, 0.99] **Median[g]: 0.95 [0.88, 0.99]**	0.90 [0.88, 0.92] 0.93 [0.90, 0.96] **0.90 [0.88, 0.92]**	**Sensitivity**				
							Very serious[a]	serious[b]	No serious inconsistency	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	No serious inconsistency	serious[d]	VERY LOW
Heart Spectrum Blood Pressure Monitor algorithm 1 (see evidence tables) *1 reading(unknown*	1 Kao, 2018[110]	62	expert	yes	0.97 [0.82, 1.00]	0.97 [0.84, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
duration)							**Specificity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
Heart Spectrum Blood Pressure Monitor algorithm 2 (see evidence tables) *1 reading(unknown duration)*	1 Kao, 2018[110]	62	expert	yes	0.90 [0.73, 0.98]	1.00 [0.89, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
Heart Spectrum Blood Pressure Monitor algorithm 3 (see evidence tables) *1 reading(unknown duration)*	1 Kao, 2018[110]	62	expert	yes	1.00 [0.88, 1.00]	0.94 [0.80, 0.99]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
Omron 712C automatic sphygmomanometer *2 readings of 10-40 secs each*	1 Wiesel, 2004[261]	450	auto	no	1.00 [0.93, 1.00]	0.91 [0.88, 0.94]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	No serious imprecision	LOW
							Specificity				
							Very	No	Not	serious[d]	VERY

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
							serious[a]	serious indirectness	applicable		LOW
Omron M6 Comfort *1 reading(unknown duration)*	2 Marazzi, 2012[143] Wiesel, 2014[259]	686	yes no	auto auto	1.00 [0.96, 1.00] 0.33 [0.17, 0.53] **Median[g]: 0.33 [0.17, 0.53]**	0.94 [0.92, 0.96] 0.97 [0.93, 0.99] **0.97 [0.93, 0.99]**	**Sensitivity**				
							Serious[a]	No serious indirectness	serious[c]	no serious imprecision	LOW
							Specificity				
							Serious[a]	No serious indirectness	No serious imprecision	no serious imprecision	MOD

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.
(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.
(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(e) Pooled sensitivity/specificity from diagnostic meta-analysis
(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses
(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.

Table 6: Clinical evidence summary: diagnostic test accuracy for pulse palpation (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
Pulse palpation *1 reading(unknown duration)*	2 Hobbs, 2005[92] Somerville, 2000[220] Somerville, 2000[220]	2616	Clinician Clinician expert	yes no no	0.87 [0.82, 0.91] 0.92 [0.64, 1.00] 1.00 [0.87, 1.00] **Pooled:0.92(0.71-0.99)** [d]	0.81 [0.80, 0.83] 0.84 [0.64, 0.95] 0.77 [0.64, 0.87] **Pooled:0.81(0.56-0.94)** [d]	**Sensitivity**				
							serious[a]	serious[b]	No serious inconsistency	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	No serious inconsistency	very serious[d]	VERY LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.
(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(d) Pooled sensitivity/specificity from diagnostic meta-analysis

Table 7: Clinical evidence summary: diagnostic test accuracy for photoplethysmography (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard)

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
iPhone 4s app - 2 minute pulse waveforms with PULSESMART app (using RMSSD, ShE and Poincare thresholds) from fingertip pulse recordings *1 reading (2 mins)*	1 McManus, 2016[145]	189	auto	yes	0.97 [0.91, 0.99]	0.93 [0.86, 0.98]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
iPhone 4S camera using both RMSSD and Shannon entropy thresholds from fingertip pulse recordings *1 reading (2 mins)*	1 McManus, 2013[146]	76	Expert, calculated algorithm	no	0.962[d]	0.975[d]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
iPhone 4S camera using just RMSSD threshold from fingertip pulse recordings *1 reading (2 mins)*	1 McManus, 2013[146]	76	Expert, calculated algorithm	no	0.982[d]	0.915[d]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
iPhone 4S camera using just Shannon entropy threshold from fingertip pulse recordings *1 reading (2 mins)*	1 McManus, 2013[146]	76	Expert, calculated algorithm	no	0.975[d]	0.822[d]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
Fingertip Cardio Rhythm Mobile Application (CRMA) iphone *3 x 20s readings (majority rule)*	1 Rozen, 2018[200]	189	auto	no	0.93[0.86, 0.97]	0.91[0.83, 0.96]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Fingertip Cardio Rhythm Mobile Application (CRMA) iphone *3 x 20s readings (minority rule)*	1 Yan, 2018[267]	217	auto	no	0.95 [0.87, 0.99]	0.93 [0.87, 0.97]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
Facial Cardio Rhythm Mobile Application (CRMA) iphone *3 x 20s readings (minority rule)*	1 Yan, 2018[267]	217	auto	no	0.95 [0.87, 0.99]	0.96 [0.91, 0.98]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
FibriCheck app. A PPG signal (fingertip) was acquired with the rear camera of an iPhone 5S. *3 x 1min readings*	1 Proesmans, 2019[178]	214	auto	no	0.87 [0.79, 0.93]	0.97 [0.92, 0.99]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
Huawei Honor 7A fingertip/LED device (auto) *1 x 3mins readings*	1 Fan, 2019[69]	108	auto	yes	0.956(0.902-0.982) No raw data in paper – CIs provided by paper)	0.99.4(0.962-1.00) No raw data in paper – CIs provided by paper)	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
Huawei Mate 9 fingertip/LED device (auto) 3mins *1 x 3mins readings*	1 Fan, 2019[69]	108	auto	yes	0.944(0.889-0.974) No raw data in paper – CIs provided by paper)	1.00(0.972-1.00) No raw data in paper – CIs provided by paper)	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Serious[c]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not	No serious	LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
									applicable	imprecision	
The screening technique involves a finger-probe instrument (as used in pulse oximetry) that utilises the principle of photoplethysmography. *1 x 30s reading*	1 Lewis, 2011[132]	594	auto	no	1.00[d]	Incorrect calculation in paper but insufficient data to correct it	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Not assessed	LOW
							Specificity				
							Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
Wrist-type photoplethysmographic (PPG) device. Using inter-beat interval (IBI) features (mean, SD, median, IQR, min, max and RMSSD. *1 reading (unknown duration)*	1 Fallet, 2019[68]	17	auto	yes	0.995[d]	0.895[d]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
Wrist-type photoplethysmographic (PPG) device. Using 'wave' features (Adaptive organisation	1 Fallet, 2019[68]	17	auto	yes	0.992[d]	0.906[d]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
Index, variance of the slope of the phase difference, permutation entropy, fractional spectral radius and spectral purity index) *1 reading (unknown duration)*											
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
Wrist-type photoplethysmographic (PPG) device. Using BOTH IBI and wave features *1 reading (unknown duration)*	1 Fallet, 2019[68]	17	auto	yes	0.997[d]	0.924[d]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.
(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(d) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses

Table 8: Clinical evidence summary: diagnostic test accuracy for 3-lead tele ECG (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
CG-7100, Card Guard Scientific Survival Ltd 3-lead tele-ECG. *2 readings (unknown duration)*	1 Antonicelli, 2012[5]	107	expert	no	1.00 [0.03, 1.00]	1.00 [0.97, 1.00]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Very serious[d]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
Handheld tele ECG device with dry electrodes that records 3 lead ECG. *1 reading of 120s*	1 Winkler, 2011[266]	60	automated	no	0.929[f]	0.909[f]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	Not assessed	VERY LOW
Portable 3-lead ECG monitor (PEM) *1 reading of 30s*	1 Kristensen, 2016[122]	89	clinician	yes	0.87 [0.60, 0.98]	0.99 [0.93, 1.00]	**Sensitivity**				
							No serious risk of bias	serious[b]	Not applicable	serious[d]	LOW
							Specificity				
							No serious risk of bias	serious[b]	Not applicable	No serious imprecision	MOD

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
Wearable and wireless 3-lead ECG device (Medi-Trace 200, Kendall) connected to the user via disposable button electrodes. Signals then analysed according to algorithm 1: if the variation of consecutive R-R intervals is >150ms within 6 secs of computation *1 reading of 6 min*	1 Lin, 2010[135]	1200	automated	no	0.93 [0.92, 0.95]**	0.00 [0.00, 0.37]**	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
Wearable and wireless 3-lead ECG device (Medi-Trace 200, Kendall) connected to the user via disposable button electrodes. Signals then analysed according to 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 *1 reading of 6min*	1 Lin, 2010[135]	1200	automated	no	0.95 [0.94, 0.96]**	0.00 [0.00, 0.41]**	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.

(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect

(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.

(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.

(e) Pooled sensitivity/specificity from diagnostic meta-analysis

(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses

(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.

***unit of analysis was each of the separate measures done on each person*

Table 9: Clinical evidence summary: diagnostic test accuracy for 6 lead ECG (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
6 lead ECG with prototype recorder placed on thorax/abdomen in sitting with loosed clothing only *1 measure of 5 seconds*	1 Caldwell, 2012[27]	157	expert	no	0.97 [0.91, 1.00]	1.00 [0.95, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
6 lead ECG with prototype recorder placed on thorax/abdomen in	1 Caldwell, 2012[27]	157	expert	no	0.97 [0.91, 1.00]	1.00 [0.95, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
supine *1 measure of 5 seconds*											
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
6 lead ECG with standard electrode positions *1 measure of 5 seconds*	1 Caldwell, 2012[27]	157	expert	no	0.97 [0.91, 1.00]	1.00 [0.95, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
6 lead ECG using Rhythm Pad device *1 measure of 10s*	1 Sabar, 2019[202]	752	Auto Expert	No no	0.95 [0.87, 0.99] 0.94 [0.85, 0.98] **Median[g]: 0.94 [0.85, 0.98]**	0.99 [0.97, 1.00] 0.97 [0.95, 0.98] **0.97 [0.95, 0.98]**	**Sensitivity**				
							Very serious[a]	No serious indirectness	No serious inconsistency	serious [d]	VERY LOW
							Specificity				
							Very serious[a]	No serious indirectness	No serious inconsistency	No serious imprecision	LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.

(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect

(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.

(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.

(e) Pooled sensitivity/specificity from diagnostic meta-analysis
(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses
(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.

Table 10: Clinical evidence summary: diagnostic test accuracy for other non-12 lead ECG (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
Limb lead ECG *1 measure of unknown duration*	1 Mant, 2007[142] Mant, 2007[142]	1484	Clinician (GP) Clinician (nurse)	Yes yes	0.83 [0.75, 0.89] 0.72 [0.63, 0.80] **Median[g]: 0.72 [0.63, 0.80]**	0.89 [0.87, 0.90] 0.83 [0.81, 0.85] **0.83 [0.81, 0.85]**	**Sensitivity**				
							No serious risk of bias	No serious indirectness	No serious inconsistency	No serious imprecision	HIGH
							Specificity				
							No serious risk of bias	No serious indirectness	serious[c]	No serious imprecision	MOD
Chest lead ECG *1 measure of unknown duration*	1 Mant, 2007[142] Mant, 2007[142]	1484	Clinician (GP) Clinician (nurse)	Yes yes	0.85 [0.78, 0.90] 0.69 [0.60, 0.76] **Median[g]: 0.69 [0.60, 0.76]**	0.86 [0.84, 0.88] 0.98 [0.97, 0.99] **0.98 [0.97, 0.99]**	**Sensitivity**				
							No serious risk of bias	No serious indirectness	serious[c]	No serious imprecision	MOD
							Specificity				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
							No serious risk of bias	No serious indirectness	serious[c]	No serious imprecision	MOD
Computer interpretation of V1, V4 leads information only using Philips resting 12 lead algorithm *1 measure of 10s*	1 Greg, 2008[73]	1785	automated	yes	0.88 [0.80, 0.93]	0.99 [0.98, 0.99]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
Computer interpretation of V2, V5 leads information only using Philips resting 12 lead algorithm *1 measure of 10s*	1 Greg, 2008[73]	1785	automated	yes	0.84 [0.76, 0.91]	0.99 [0.98, 0.99]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
bipolar lead 1 ECG (no other details) *1 measure of unknown duration*	1 Somerville, 2000[220]	86	Clinician (GP) Expert (nurse)	No no	0.96 [0.80, 1.00] 0.92 [0.75, 0.99] **Median[g]: 0.92 [0.75, 0.99]**	0.98 [0.91, 1.00] 0.88 [0.77, 0.95] **0.88 [0.77, 0.95]**	**Sensitivity**				
							Very serious[a]	serious[b]	serious[c]	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	serious[c]	serious[d]	VERY LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.
(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.
(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(e) Pooled sensitivity/specificity from diagnostic meta-analysis
(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses
(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.

Table 11: Clinical evidence summary: diagnostic test accuracy for 12 lead ECG interpreted by automated algorithm or non-expert interpreters (12 lead ECG interpreted by expert cardiologist/electrophysiologist as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
12 lead ECG interpreted by non-	5 Hald, 2017[79]	2999	Clinician (GP/nurse)	yes	1.00 [0.69, 1.00]	0.96 [0.89, 0.99]	**Sensitivity**				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
expert interpreter (GP or nurse *1 measure of unknown duration*	Kvist, 2019[124] Mant, 2007[142] Mant, 2007[142] Rhys, 2013[190] Somerville, 2000[220] Somerville, 2000[220]		Clinician (nurse) Clinician (GP) Clinician (nurse Clinician (GP) Clinician (GP) Clinician (nurse)	no Yes Yes yes No No	0.97 [0.90, 1.00] 0.78 [0.69, 0.86] 0.77 [0.67, 0.85] 1.00 [0.16, 1.00] 1.00 [0.87, 1.00] 1.00 [0.75, 1.00] **Pooled:0.89(0.77-0.96)** [e]	1.00 [1.00, 1.00] 0.92 [0.90, 0.93] 0.85 [0.83, 0.87] 1.00 [0.88, 1.00] 0.98 [0.91, 1.00] 0.76 [0.55, 0.91] **Pooled:0.97(0.85-0.99)** [e]	serious[a]	serious[b]	serious[c]	serious[d]	VERY LOW
							Specificity				
							serious[a]	serious[b]	serious[c]	serious[d]	VERY LOW
12 lead ECG interpreted by non-expert interpreter (GP or nurse) combined with interpretive algorithm *1 measure of unknown duration*	1 Mant, 2007[142]	1454	Clinician + automated		0.92 [0.85, 0.96]	0.91 [0.89, 0.93]	**Sensitivity**				
							No serious risk of bias	No serious indirectness	Not applicable	serious[d]	MOD
							Specificity				
							No serious risk of bias	No serious indirectness	Not applicable	serious[d]	MOD
12 lead ECG detection algorithm based on a co-efficient of variation of the beat intervals (CV). Threshold set at 0.12 *1 measure of 10s*	1 Langley, 2012[127]	2124	automated	yes	0.905[f]	0.896[f]	**Sensitivity**				
							No serious risk of bias	serious[b]	Not applicable	Not assessed	MOD
							Specificity				
							No serious risk of bias	serious[b]	Not applicable	Not assessed	MOD
12 lead ECG	1	2124					**Sensitivity**				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
detection algorithm based on the co-efficient of sample entropy (COSEn). Threshold set at -1.19 *1 measure of 10s*	Langley, 2012[127]		automated	yes	0.952[f]	0.934[f]	No serious risk of bias	serious[b]	Not applicable	Not asses sed	MOD
							Specificity				
							No serious risk of bias	serious[b]	Not applicable	Not asses sed	MOD
12 lead ECG detection algorithm 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 *1 measure of 10s*	1 Langley, 2012[127]	2124	automated	yes	0.905[f]	0.893[f]	**Sensitivity**				
							No serious risk of bias	serious[b]	Not applicable	Not asses sed	MOD
							Specificity				
							No serious risk of bias	serious[b]	Not applicable	Not asses sed	MOD
12 lead ECG interpreted by algorithm in Cardioview interpretive software (not 1described) *measure of unknown duration*	1 Rhys, 2013[190]	32	automated	yes	1.00 [0.16, 1.00]	1.00 [0.88, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	Very seriou s[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	seriou s[d]	VERY LOW
12 lead ECG interpreted by	1 Poon, 2005[176]	3954	automated	yes	0.91 [0.87, 0.94]	0.99 [0.99, 0.99]	**Sensitivity**				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
computer-based rhythm diagnosis (GE Healthcare Technologies MUSE software 005C, version 19) (automated) *1 measure of unknown duration*							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
12 lead interpretive software (no details given) *1 measure of unknown duration*	1 Mant, 2007[142]	2556	automated	yes	0.83 [0.78, 0.88]	0.99 [0.99, 0.99]	**Sensitivity**				
							No serious risk of bias	No serious indirectness	Not applicable	No serious imprecision	HIGH
							Specificity				
							No serious risk of bias	No serious indirectness	Not applicable	No serious imprecision	HIGH
Algorithm for reading 12 lead ECGs. This first tested for the presence of non-coupled P waves. If non-coupled P	1 Slocum, 1992[217]	82	automated	yes	0.68 [0.52, 0.82]	0.88 [0.74, 0.96]	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
waves were detected the rhythm was considered non-atrial fibrillation and no further testing was done. If the rhythm did not have non-coupled P waves, and 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. *1 measure of unknown duration*							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
Computer interpretation of full 12 lead ECG V1-V6 using Philips resting 12 lead algorithm *1 measure of 10s*	1 Greg, 2008[73]	1785	automated	yes	0.89 [0.82, 0.94]	0.99 [0.98, 0.99]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.

(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.
(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(e) Pooled sensitivity/specificity from diagnostic meta-analysis
(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses
(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.

1.5.6 Quality assessment of clinical studies included in the evidence review

STRATUM 2: >24 hour ambulatory monitoring [such as Holter] as gold standard

Table 12: Clinical evidence summary: diagnostic test accuracy for blood pressure monitors (>24 hour ambulatory monitoring as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
24 hr ambulatory Microlife Afib device Watch BP *1 reading every 20 mins over 24 hrs*	1 Kollias, 2018[116]	5778	automated	yes	0.93 [0.91, 0.94]	0.98 [0.98, 0.99]	**Sensitivity**				
							Serious[a]	No serious indirectness	Not applicable	no serious imprecision	MOD
							Specificity				
							Serious[a]	No serious indirectness	Not applicable	no serious imprecision	MOD
AF-BP monitor device, to take home and use daily for 30 days *If event detected had to take 2 additional readings. If 2 or3 indicated AF took a final reading 1 hr later*	1 Wiesel, 2013[258]	139	automated	yes	1.00 [0.77, 1.00]	0.90 [0.83, 0.94]	**Sensitivity**				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
							Specificity				
							Serious[a]	No serious indirectness	Not applicable	serious[d]	LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.

(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.
(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(e) Pooled sensitivity/specificity from diagnostic meta-analysis
(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses
(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.

Table 13: Clinical evidence summary: diagnostic test accuracy for <7 day Holter devices (7 day Holter as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
1 day Holter	1 Mulder, 2012[153]	96	automated	yes	0.52 [0.30, 0.74]**	1.00 [0.95, 1.00]**	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Sensitivity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
2 day Holter	1 Mulder, 2012[153]	96	automated	yes	0.67 [0.43, 0.85]**	1.00 [0.95, 1.00]**	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Sensitivity				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
3 day Holter	1 Mulder, 2012[153]	96	automated	yes	0.81 [0.58, 0.95]**	1.00 [0.95, 1.00]**	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	Very serious[d]	VERY LOW
							Sensitivity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
4 day Holter	1 Mulder, 2012[153]	96	automated	yes	0.86 [0.64, 0.97]**	1.00 [0.95, 1.00]**	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Sensitivity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW
5 day Holter	1 Mulder, 2012[153]	96	automated	yes	0.90 [0.70, 0.99]**	1.00 [0.95, 1.00]**	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Sensitivity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
6 day Holter	1 Mulder, 2012[153]	96	automated	yes	0.95 [0.76, 1.00]**	1.00 [0.95, 1.00]**	**Sensitivity**				
							serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							serious[a]	serious[b]	Not applicable	No serious imprecision	LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.
(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and 'not applicable' was recorded.
(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(e) Pooled sensitivity/specificity from diagnostic meta-analysis
(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses
(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.
***unit of analysis was each of the separate measures done on each person*

Table 14: Clinical evidence summary: other longer term devices (>24 hour ambulatory monitoring as gold standard). Where 95% CIs are provided in round brackets (or no 95% CIs are given), raw data were not available and Forest Plots or pooled analyses were not possible.

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecision	GRADE
R.Test Evolution 3 – triggered ECG	1 Roten, 2012[199]	100	expert	yes	0.88 [0.74, 0.96]	1.00 [0.94, 1.00]	**Sensitivity**				
							Very serious[a]	serious[b]	Not applicable	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	serious[b]	Not applicable	No serious imprecision	VERY LOW
R.Test Evolution 4 (NorDiaTech, Paris, France) - External loop recorder (ELR) *1 measure x 48 hours*	1 Sejr, 2019[213] Sejr, 2019[213]	1412	Automated expert	Yes yes	0.92 [0.79, 0.98] 0.84 [0.69, 0.94] **Median[g]: 0.84 [0.69, 0.94]**	0.87 [0.85, 0.89] 0.98 [0.97, 0.99] **0.98 [0.97, 0.99]**	**Sensitivity**				
							No serious risk of bias	No serious indirectness	No serious inconsistency	serious[d]	MOD
							Specificity				
							No serious risk of bias	No serious indirectness	serious[c]	No serious imprecision	
Vitaphone 3100 BT external loop recorder. Recorded event ECGs manually when triggered by the patient or automatically	2 Muller, 2009[154] Velthuis, 2013[248]	104	Automated automated	Yes yes	1.00 [0.86, 1.00] 0.95 [0.86, 0.99]** **Median[g]: 0.95 [0.86, 0.99]**	0.50 [0.29, 0.71] 0.51 [0.49, 0.53]** **0.51 [0.49, 0.53]**	**Sensitivity**				
							serious[a]	serious[b]	No serious inconsistency	serious[d]	VERY LOW
							Specificity				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
1 measure x 24 hours							serious[a]	serious[b]	No serious inconsistency	No serious imprecision	LOW
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.(automated) threshold of 0-1=SR and 2 or more =AF *1 measure of >24 hours*	1 Ross, 2018[198]	798	automated	yes	0.98(0.952-0.990)** No raw data in paper – CIs provided by paper)	0.27(0.223-0.322)** No raw data in paper – CIs provided by paper)	**Sensitivity**				
							serious[a]	No serious indirectness	Not applicable	No serious imprecision	MOD
							Specificity				
							serious[a]	No serious indirectness	Not applicable	No serious imprecision	MOD
SRAclinic, Apoplex medical Technologies. Stroke Risk Analysis (SRA) – software analysis of every hourly ECG snippet	1 Ross, 2018[198]	798	automated	yes	0.840(0.790-0.878)** No raw data in paper – CIs provided by paper)	0.700(0.645-0.749)** No raw data in paper – CIs provided by paper)	**Sensitivity**				
							serious[a]	No serious indirectness	Not applicable	No serious imprecision	MOD
							Specificity				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
of continuous (non 12 lead) ECG monitoring, and report sent daily to stroke unit.(automated) threshold of 0-2=SR and 3 or more =AF *1 measure of >24 hours*							serious[a]	No serious indirectness	Not applicable	No serious imprecision	MOD
Standard ECG monitoring devices without AF-RS - with ECG confirmation if positive (thus specificity was 100%) *1 measure of 48 hours*	1 Arevalo-Manso, 2016[6]	59	Automated/clinician	No	0.08 [0.00, 0.36]	1.00 [0.92, 1.00]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	No serious imprecision	LOW
							Specificity				
							Very serious[a]	No serious indirectness	Not applicable	No serious imprecision	LOW
4 lead AF-RS monitor - with ECG confirmation if positive (thus specificity was 100%) *1 measure of 48 hours*	1 Arevalo-Manso, 2016[6]	17	Automated/clinician	No	0.57 [0.18, 0.90]	1.00 [0.69, 1.00]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	Very serious[d]	LOW
							Specificity				
							Very serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
12-bit resolution	1	136					**Sensitivity**				

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
digital ECG recoding for 1-2 hours *1 measure of 1-2 hours*	Rizos, 2010[194]		automated	yes	0.72 [0.53, 0.87]	0.63 [0.53, 0.72]	serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
							Specificity				
							serious[a]	No serious indirectness	Not applicable	serious[d]	LOW
6 channel Holter (H12+, Mortara Instruments) *1 measure of 24 hours*	1 Rizos, 2010[194]	136	expert	yes	0.23 [0.05, 0.54]	1.00 [0.97, 1.00]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	No serious imprecision	LOW
							Specificity				
							Very serious[a]	No serious indirectness	Not applicable	No serious imprecision	LOW
Thumb ECG (Zenicor Medical Systems AB) *30s measure twice daily for 30 days*	1 Poulsen, 2017[177]	95	expert	yes	0.59 [0.33, 0.82]	0.87 [0.78, 0.94]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	serious[d]	VERY LOW
							Specificity				
							Very serious[a]	No serious indirectness	Not applicable	serious[d]	VERY LOW
Kardia-Band *Mean 11.3 hrs per day for 110 days*	1 Wasserlauf, 2019[255]	26	automated	yes	0.977 [f]	0.989 [f]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	Not assessed	LOW
							Specificity				
							Very serious[a]	No serious indirectness	Not applicable	Not assessed	LOW

Index Test	Number of studies	n	Interpreter of index test	Gold standard simultaneous with index test?	Sensitivity (95% CI)	Specificity (95% CI)	Risk of bias	Indirectness	Inconsistency	Imprecsion	GRADE
>24 hour telemetry with ‘electrocardiomatrix’ *Median 46 hours*	1 Brown, 2019[22]	260	automated	yes	0.98 [0.95, 0.99]	0.86 [0.71, 0.95]	**Sensitivity**				
							Very serious[a]	No serious indirectness	Not applicable	No serious imprecision	LOW
							Specificity				
							Very serious[a]	No serious indirectness	Not applicable	serious[d]	VERY LOW

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.
(b) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect
(c) Inconsistency was assessed by visual inspection of the sensitivity/specificity plots, or data (if 2 studies). The evidence was downgraded by 1 increment if there was no overlap of 95% confidence intervals. For single studies no evaluation was made and ‘not applicable’ was recorded.
(d) Imprecision was assessed based on inspection of the confidence region in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the confidence interval around the point estimate crossed one of the clinical thresholds (0.90 or 0.60), and downgraded by 2 increments when the confidence interval around the point estimate crossed both of the clinical thresholds. The upper clinical threshold marked the point above which recommendations would be possible, and the lower clinical threshold marked the point below which the tool would be regarded as of little clinical use.
(e) Pooled sensitivity/specificity from diagnostic meta-analysis
(f) indicates that because the raw data were not provided by the paper, and were also not able to be calculated by the reviewer, these were not able to be included in any diagnostic meta-analyses
(g) For non-pooled analyses, the median sensitivity and the paired specificity were reported. If there were an even number of studies the lower middle value was given.
***unit of analysis was each of the separate measures done on each person*

1.6 Economic evidence

Please see evidence review A.

1.7 The committee's discussion of the evidence

1.7.1 Interpreting the evidence

1.7.1.1 The outcomes that matter most

For the diagnostic accuracy review, the outcomes were sensitivity and specificity. For a test that is suitable to be used alone as a definitive diagnostic test (in place of 12 lead ECG), both sensitivity and specificity are of equal value, as a definitive test needs to have almost perfect sensitivity and specificity. High sensitivity is essential to avoid people with true AF being missed and therefore untreated, as this can lead to serious sequelae such as stroke. High specificity is equally important to prevent people without AF being misdiagnosed as having it, which may lead to unnecessary prescription of anticoagulants, antiarrhythmic drugs, or invasive procedures, all of which carry a burden of serious adverse effects.

In contrast, for tests that might be used as the first part of a two stage testing process (an example of such a two stage process is pulse palpation followed by 12 lead ECG in people who test positive) then sensitivity may be more important than specificity. Reasons for this are as follows. In a two-test scenario, the initial test is used as a filter to decide who goes on to the resource-intensive 12 lead ECG, and this could be achieved by *either* an extremely sensitive initial test or an extremely *specific* initial test. With a highly sensitive initial test, only initial positives go on to the next stage of testing (where the false positives resulting from the sub-optimal specificity of the initial test can be 'weeded out' by 12 lead ECG). Initial negatives can be safely discarded from the diagnostic process when the initial test has high sensitivity, because very high sensitivity means that the initial negatives should contain hardly any people with true AF. In contrast, with an extremely *specific* initial test, only initial negatives go on to further testing (where the false negatives resulting from the sub-optimal sensitivity of the initial test can be 'weeded out' by 12 lead ECG). The initial positives can be regarded as diagnostic in the presence of high specificity because high specificity means that almost all initial positives will have true AF. Because there are likely to be fewer initial positives than initial negatives, using a highly sensitive test is likely to lead to fewer people going on to the 12-lead test than use of a highly specific test. A highly sensitive initial test is therefore the preferable option for a two-stage process because the purpose of a two-stage process is to limit use of the resource-intensive 12 lead ECG.

Positive predictive value (PPV) and negative predictive value (NPV) are important for health economic considerations but are less important for evaluating clinical utility, and are often unreliable when calculated from study data as they are dependent on the prevalence which may not always be representative in studies. The aim had been to calculate PPV and NPV for any tools that had good evidence of adequate sensitivity and specificity in relation to an agreed prevalence rate of AF. However, this was not carried out because no tools were identified.

For the RCT review, outcomes were quality of life, mortality, stroke and thromboembolism, Major bleeding, all cause hospitalisation, confirmed diagnosis of AF and initiated anticoagulants for AF. All were regarded as critical by the committee, but quality of life, stroke/systemic embolism, mortality, and confirmed diagnosis of AF were deemed the most relevant for decision-making. These were prioritised over other critical outcomes because 'quality of life' was felt to provide the most comprehensive measure of benefit to the patient, 'stroke and systemic thromboembolism' was regarded as the major serious complication of AF, 'mortality' was felt to best characterise the harms of treatment, and 'confirmed diagnosis of AF' was thought to best characterise the benefits of treatment.

1.7.1.2 The quality of the evidence

For the diagnostic accuracy evidence, most data were rated as at serious or very serious risk of bias, because of a lack of simultaneity between index and reference tests, and because of a lack of blinding in some studies. Indirectness was also often rated as serious because the populations in studies differed from the protocol definition. Overall, most data were rated as low or very low. For the RCT evidence, a similar picture existed. Serious or very serious risk of bias was largely due to issues around selection and attrition bias, and again indirectness of populations was a major issue. Outcomes were therefore mostly rated as low or very low.

1.7.1.3 Benefits and harms

The diagnostic accuracy data for the different index test devices in relation to the gold standard of 12 lead ECG interpreted by a cardiologist/electrophysiologist were initially discussed. These devices included mobile ECG devices, HR monitors, blood pressure measurements, photoplethysmographic technique, pulse palpation, other ECG measures and 12 lead ECG not interpreted by an expert. The sensitivity and specificity of the majority of these devices were regarded by the committee as insufficiently high to permit their use as a single diagnostic test. Some devices, such as HR monitors, BP devices or plethysmographic devices, did approach 100% sensitivity and specificity, but these had often been tested in small samples leading to imprecise estimates. Alternatively, such estimates were from large but solitary studies. The committee noted that accuracy differed quite widely between different studies looking at the same test and they were therefore unable to make recommendations based on results from single studies.

Having decided that none of the tests could be used as an individual (definitive) diagnostic test, the committee discussed whether any of the tests could be used as a first-line test, prior to 12 lead ECG (please see 'outcomes that matter the most' above for an explanation of this process). The committee realised that such tests would need perfect or almost perfect sensitivity to avoid losing some people with AF from the diagnostic process (with enough specificity to allow a worthwhile reduction in the number going on to 12 lead testing compared to 12 lead testing used alone). The current recommendation is to use pulse palpation as the initial test, and thus an alternative test would need to have clear superiority in sensitivity over pulse palpation (with similar specificity) to justify replacement of pulse palpation. Some of the devices had sensitivity point estimates that exceeded those of pulse palpation, with upper 95% confidence intervals that extended closer to maximal sensitivity than those for pulse palpation. This provided weak evidence that some of the devices might be of greater use as a first line test than pulse palpation. However, the confidence intervals of the devices overlapped with those of pulse palpation, demonstrating a level of uncertainty about such superiority in the population. The committee were of the opinion that this level of uncertainty was insufficient to change the established practice of pulse palpation, which is a core clinical skill in widespread use, and which is extremely quick and low-cost to carry out. However, they felt that new devices had promise, which might be manifested in further high-quality research, and so a research recommendation was proposed, alongside a continuation of the current recommendation.

It is important to note a subtle change to the recommendations regarding the definitive test to be used if pulse irregularities are observed. In the previous guideline the recommendation had been to use 'ECG' as the definitive test, whereas in the present guideline we are specifying ' 12-lead ECG' as the definitive test. This change was noted by the committee to be very important to prevent non-12 lead ECG such as lead I devices (which this review has shown to be lacking in adequate accuracy compared to 12 lead ECG) being used as the definitive test.

The diagnostic accuracy for the devices tested in relation to a longer-term gold standard (>24-hour ambulatory monitoring) were also considered by the committee. This evidence was regarded as particularly important as it was the only evidence able to inform the accuracy of detection of paroxysmal AF (12 lead ECG usually lasts only 10 seconds and so whilst it is perfectly good as a gold standard for detecting persistent AF it is often inadequate for detecting paroxysmal AF). The committee again noted that the evidence did not suggest that any specific test or device should be recommended but did note that the evidence clearly demonstrated that the accuracy of detection increased with the duration of testing. Therefore, the committee recommended that testing for suspected paroxysmal AF should be continued for as long as possible by any form of continuous or loop monitoring.

The committee agreed that the RCT review did not offer particularly useful evidence to inform recommendations, over and above the data provided by the diagnostic accuracy review. In particular, the committee highlighted that the follow up periods of the included studies were too short to allow a meaningful picture of downstream clinical outcomes. The RCT review was also noted to have serious gaps in terms of many of the available tests not having been studied.

1.7.2 Cost effectiveness and resource use

One cost-utility analysis was identified comparing single time point lead-I ECG devices with manual pulse palpation (MPP) followed by a 12-lead ECG in primary or secondary care for the detection of AF in people presenting to primary care with signs or symptoms of AF and who have an irregular pulse. This cost utility analysis was conducted as part of the NICE Diagnostic Guidance DG35 published in 2019 for lead-I devices. The study found that in all base case scenarios (these varied the time to and location of confirmatory 12 lead ECG) Kardia mobile, where treatment for AF is initiated following a positive result, ahead of confirmatory 12-lead ECG test, was the more cost-effective than the standard diagnostic pathway where no treatment is initiated until 12 lead ECG testing is complete. Furthermore, Kardia Mobile dominated (less costly and more effective) all other lead-I devices included in the analysis. This study was partially applicable as it did not include all comparators in the protocol for this question. There were potential serious limitations, primarily due to the fact the sensitivity and specificity data used in this analysis was from studies conducted in asymptomatic patients, and so this was indirect evidence. Furthermore, the economic evaluation is only relevant to primary care practices where patients have to wait at least 48 hours between an initial consultation with the GP and a 12-lead ECG.

In addition to this study, unit costs for different methods of detecting AF were presented, including current practice that is manual pulse palpitation followed by 12-lead ECG in those with an irregular pulse. The committee noted that although the lead-I devices do not appear particularly costly per use; they may add a significant resource burden in terms of the need for expert interpretation. This would either require training of GPs or would necessitate sending lead-I results to cardiologists for guidance and advice.

The committee considered the published health economic analysis alongside the clinical evidence and concluded that there was insufficient direct evidence to support replacing the current methods of detecting AF. In particular, the health economic evidence is based on indirect clinical evidence and there is uncertainty as to whether the sensitivity and specificity can be translated from an asymptomatic to a symptomatic AF population. This is in line with the guidance from DG35.

Overall, therefore the committee have kept the previous recommendations, only adjusting the wording to make these clearer. As they represent current practice, no resource impact is anticipated.

1.7.3 Other factors the committee took into account

The committee noted that the use of hand-held devices could improve diagnosis in people who find it impossible or difficult to access EEG services, for example people in care homes.

References

1. Adami A, Gentile C, Hepp T, Molon G, Gigli GL, Valente M et al. Electrocardiographic RR interval dynamic analysis to identify acute stroke patients at high risk for atrial fibrillation episodes during stroke unit admission. Translational Stroke Research. 2019; 10(3):273-278

2. Afzal MR, Gunda S, Waheed S, Sehar N, Maybrook RJ, Dawn B et al. Role of outpatient cardiac rhythm monitoring in cryptogenic stroke: A systematic review and meta-analysis. Pacing and Clinical Electrophysiology. 2015; 38(10):1236-1245

3. Alshraideh H, Otoom M, Al-Araida A, Bawaneh H, Bravo J. A web based cardiovascular disease detection system. Journal of Medical Systems. 2015; 39(122)

4. Alves M, Narciso MR, Cruz J, Rocha M, Fonseca T. Paroxysmal atrial fibrillation detection in patients with acute ischemic stroke through prolonged Holter: prospective study. Aging-Clinical and Experimental Research. 2019; 31(4):469-474

5. Antonicelli R, Ripa C, Abbatecola AM, Capparuccia CA, Ferrara L, Spazzafumo L. Validation of the 3-lead tele-ECG versus the 12-lead tele-ECG and the conventional 12-lead ECG method in older people. Journal of Telemedicine and Telecare. 2012; 18(2):104-8

6. Arevalo-Manso JJ, Martinez-Sanchez P, Fuentes B, Ruiz-Ares G, Sanz-Cuesta BE, Prefasi D et al. Can we improve the early detection of atrial fibrillation in a stroke unit? Detection rate of a monitor with integrated detection software. European Journal of Cardiovascular Nursing. 2016; 15(1):64-71

7. Athif M, Yasawardene PC, Daluwatte C. Detecting atrial fibrillation from short single lead ECGs using statistical and morphological features. Physiological Measurement. 2018; 39(6):064002

8. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet. 2019; 394(10201):861-867

9. Barrett PM, Komatireddy R, Haaser S, Topol S, Sheard J, Encinas J et al. Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring. American Journal of Medicine. 2014; 127(1):95.e11-95.e17

10. Barthelemy JC, Feasson-Gerard S, Garnier P, Gaspoz JM, Da Costa A, Michel D et al. Automatic cardiac event recorders reveal paroxysmal atrial fibrillation after unexplained strokes or transient ischemic attacks. Annals of Noninvasive Electrocardiology. 2003; 8(3):194-9

11. Bell C, Kapral M. Use of ambulatory electrocardiography for the detection of paroxysmal atrial fibrillation in patients with stroke. Canadian Journal of Neurological Sciences. 2000; 27(1):25-31

12. Berge T, Brynildsen J, Larssen HKN, Onarheim S, Jenssen GR, Ihle-Hansen H et al. Systematic screening for atrial fibrillation in a 65-year-old population with risk factors for stroke: data from the Akershus Cardiac Examination 1950 study. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2018; 20(FI3):f299-f305

13. Berge T, Bryildsen J, Larssen HKN, Onarheim S, Jenssen GR, Ihle-Hansen H et al. Systematic screening for atrial fibrillation in 65-year-olds with risk factors for stroke. Data from the ACE 1950 Study. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2017; 19 (Supplement 3):iii355

14. Bettin M, Dechering D, Kochhauser S, Bode N, Eckardt L, Frommeyer G et al. Extended ECG monitoring with an implantable loop recorder in patients with cryptogenic stroke: time schedule, reasons for explantation and incidental findings (results from the TRACK-AF trial). Clinical Research in Cardiology. 2019; 108(3):309-314

15. Beukema R, Beukema WP, Sie HT, Misier AR, Delnoy PP, Elvan A. Monitoring of atrial fibrillation burden after surgical ablation: relevancy of end-point criteria after radiofrequency ablation treatment of patients with lone atrial fibrillation. Interactive Cardiovascular and Thoracic Surgery. 2009; 9(6):956-9

16. Bonomi AG, Schipper F, Eerikainen LM, Margarito J, Van Dinther R, Muesch G et al. Atrial fibrillation detection using a novel cardiac ambulatory monitor based on photo-plethysmography at the wrist. Journal of the American Heart Association. 2018; 7(15):e009351

17. Botto GL, Padeletti L, Santini M, Capucci A, Gulizia M, Zolezzi F et al. Presence and duration of atrial fibrillation detected by continuous monitoring: crucial implications for the risk of thromboembolic events. Journal of Cardiovascular Electrophysiology. 2009; 20(3):241-8

18. Bourdillon PJ, Kilpatrick D. Clinicians, the Mount Sinai program and the Veterans' Administration program evaluated against clinico-pathological data derived independently of the electrocardiogram. European Journal of Cardiology. 1978; 8(4-5):395-412

19. Brasier N, Raichle CJ, Dorr M, Becke A, Nohturfft V, Weber S et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2019; 21(1):41-47

20. Brembilla-Perrot B, Luporsi JD, Louis S, Kaminsky P. Long-term follow-up of patients with myotonic dystrophy: An electrocardiogram every year is not necessary. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2011; 13(2):251-257

21. Brito R, Mondouagne LP, Stettler C, Combescure C, Burri H. Automatic atrial fibrillation and flutter detection by a handheld ECG recorder, and utility of sequential finger and precordial recordings. Journal of Electrocardiology. 2018; 51(6):1135-1140

22. Brown DL, Xu G, Belinky Krzyske AM, Buhay NC, Blaha M, Wang MM et al. Electrocardiomatrix Facilitates Accurate Detection of Atrial Fibrillation in Stroke Patients. Stroke. 2019; 50(7):1676-1681

23. Buechi R, Faes L, Bachmann LM, Thiel MA, Bodmer NS, Schmid MK et al. Evidence assessing the diagnostic performance of medical smartphone apps: A systematic review and exploratory meta-analysis. BMJ Open. 2017; 7(12):e018280

24. Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K et al. Smartwatch algorithm for automated detection of atrial fibrillation. Journal of the American College of Cardiology. 2018; 71(21):2381-2388

25. Burkowitz J, Merzenich C, Grassme K, Bruggenjurgen B. Insertable cardiac monitors in the diagnosis of syncope and the detection of atrial fibrillation: A systematic review and meta-analysis. European Journal of Preventive Cardiology. 2016; 23(12):1261-72

26. Busch MC, Gross S, Alte D, Kors JA, Volzke H, Ittermann T et al. Impact of atrial fibrillation detected by extended monitoring-a population-based cohort study. Annals of Noninvasive Electrocardiology. 2017; 22(6)

27. Caldwell JC, Borbas Z, Donald A, Clifford A, Bolger L, Black A et al. Simplified electrocardiogram sampling maintains high diagnostic capability for atrial fibrillation: Implications for opportunistic atrial fibrillation screening in primary care. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2012; 14(2):191-196

28. Callizo J, Feltgen N, Ammermann A, Ganser J, Bemme S, Bertelmann T et al. Atrial fibrillation in retinal vascular occlusion disease and non-arteritic anterior ischemic optic neuropathy. PloS One. 2017; 12(8):e0181766

29. Censi F, Calcagnini G, Mattei E, Gargaro A, Biancalana G, Capucci A. Simulation of monitoring strategies for atrial arrhythmia detection. Annali dell'Istituto Superiore di Sanità. 2013; 49(2):176-82

30. Chan PH, Wong CK, Poh YC, Pun L, Leung WW, Wong YF et al. Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting. Journal of the American Heart Association. 2016; 5(7):21

31. Chan PH, Wong CK, Pun L, Wong YF, Wong MM, Chu DW et al. Diagnostic performance of an automatic blood pressure measurement device, Microlife WatchBP Home A, for atrial fibrillation screening in a real-world primary care setting. BMJ Open. 2017; 7(6):e013685

32. Chan PH, Wong CK, Pun L, Wong YF, Wong MMY, Chu DWS et al. Head-to-head comparison of the AliveCor heart monitor and Microlife WatchBP Office AFIB for atrial fibrillation screening in a primary care setting. Circulation. 2017; 135(1):110-112

33. Charitos EI, Stierle U, Ziegler PD, Baldewig M, Robinson DR, Sievers HH et al. A comprehensive evaluation of rhythm monitoring strategies for the detection of atrial fibrillation recurrence: insights from 647 continuously monitored patients and implications for monitoring after therapeutic interventions. Circulation. 2012; 126(7):806-814

34. Chen YH, Hung CS, Huang CC, Hung YC, Hwang JJ, Ho YL. Atrial Fibrillation Screening in Nonmetropolitan Areas Using a Telehealth Surveillance System With an Embedded Cloud-Computing Algorithm: Prospective Pilot Study. JMIR MHealth and UHealth. 2017; 5(9):e135

35. Choe WC, Passman RS, Brachmann J, Morillo CA, Sanna T, Bernstein RA et al. A Comparison of Atrial Fibrillation Monitoring Strategies After Cryptogenic Stroke (from the Cryptogenic Stroke and Underlying AF Trial). American Journal of Cardiology. 2015; 116(6):890-893

36. Chong JW, Cho CH, Tabei F, Le-Anh D, Esa N, McManus DD et al. Motion and noise artifact-resilient atrial fibrillation detection using a smartphone. IEEE Journal on Emerging and Selected Topics in Circuits & Systems. 2018; 8(2):230-239

37. Chong JW, Esa N, McManus DD, Chon KH. Arrhythmia discrimination using a smart phone. IEEE journal of biomedical and health informatics. 2015; 19(3):815-824

38. Chovancik J, Bulkova V, Wichterle D, Toman O, Rybka L, Januska J et al. Comparison of two modes of long-term ECG monitoring to assess the efficacy of

catheter ablation for paroxysmal atrial fibrillation. Biomedical Papers of the Medical Faculty of Palacky University in Olomouc, Czech Republic. 2019; 163(1):54-60

39. Christensen LM, Krieger DW, Hojberg S, Pedersen OD, Karlsen FM, Jacobsen MD et al. Paroxysmal atrial fibrillation occurs often in cryptogenic ischaemic stroke. Final results from the SURPRISE study. European Journal of Neurology. 2014; 21(6):884-889

40. Ciconte G, Saviano M, Giannelli L, Calovic Z, Baldi M, Ciaccio C et al. Atrial fibrillation detection using a novel three-vector cardiac implantable monitor: the atrial fibrillation detect study. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2017; 19(7):1101-1108

41. Conroy T, Guzman JH, Hall B, Tsouri G, Couderc JP. Detection of atrial fibrillation using an earlobe photoplethysmographic sensor. Physiological Measurement. 2017; 38(10):1906-1918

42. Cooke G, Doust J, Sanders S. Is pulse palpation helpful in detecting atrial fibrillation? A systematic review. Journal of Family Practice. 2006; 55(2):130-4

43. Couderc JP, Kyal S, Mestha LK, Xu B, Peterson DR, Xia X et al. Detection of atrial fibrillation using contactless facial video monitoring. Heart Rhythm. 2015; 12(1):195-201

44. Coutts SB, Choi PMC. Seven days of non-invasive cardiac monitoring early postischaemic stroke or TIA increases atrial fibrillation detection rate compared with current guideline-based practice. Evidence-Based Medicine. 2014; 19(4):152

45. Cuadrado-Godia E, Benito B, Ois A, Valles E, Rodriguez-Campello A, Giralt-Steinhauer E et al. Ultra-early continuous cardiac monitoring improves atrial fibrillation detection and prognosis of patients with cryptogenic stroke. European Journal of Neurology. 2019; 27(2):244-250

46. Cunha S, Antunes E, Antoniou S, Tiago S, Relvas R, Fernandez-Llimos F et al. Raising awareness and early detection of atrial fibrillation, an experience resorting to mobile technology centred on informed individuals. Research In Social and Administrative Pharmacy. 2019; doi: 10.1016/j.sapharm.2019.08.036

47. Dagres N, Kottkamp H, Piorkowski C, Weis S, Arya A, Sommer P et al. Influence of the duration of Holter monitoring on the detection of arrhythmia recurrences after catheter ablation of atrial fibrillation implications for patient follow-up. International Journal of Cardiology. 2010; 139(3):305-306

48. Damiano RJ, Jr., Lawrance CP, Saint LL, Henn MC, Sinn LA, Kruse J et al. Detection of atrial fibrillation after surgical ablation: Conventional versus continuous monitoring. Annals of Thoracic Surgery. 2016; 101(1):42-47

49. de Voogt WG, van Hemel NM, van de Bos AA, Koistinen J, Fast JH. Verification of pacemaker automatic mode switching for the detection of atrial fibrillation and atrial tachycardia with Holter recording. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2006; 8(11):950-61

50. DeBoard Z, Doty JR. Evaluation of new generation loop recorders placed during surgical ablation for atrial fibrillation. Journal of Cardiac Surgery. 2018; 33(7):416-419

51. Defaye P, Dournaux F, Mouton E. Prevalence of supraventricular arrhythmias from the automated analysis of data stored in the DDD pacemakers of 617 patients: the AIDA study. The AIDA Multicenter Study Group. Automatic Interpretation for Diagnosis Assistance. Pacing and Clinical Electrophysiology. 1998; 21(1 Pt 2):250-5

52. Derkac WM, Finkelmeier JR, Horgan DJ, Hutchinson MD. Diagnostic yield of asymptomatic arrhythmias detected by mobile cardiac outpatient telemetry and autotrigger looping event cardiac monitors. Journal of Cardiovascular Electrophysiology. 2017; 28(12):1475-1478

53. Desteghe L, Raymaekers Z, Lutin M, Vijgen J, Dilling-Boer D, Koopman P et al. Performance of handheld electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2017; 19(1):29-39

54. Diamantopoulos A, Sawyer LM, Lip GY, Witte KK, Reynolds MR, Fauchier L et al. Cost-effectiveness of an insertable cardiac monitor to detect atrial fibrillation in patients with cryptogenic stroke. International Journal of Stroke. 2016; 11(3):302-12

55. Dimarco AD, Onwordi EN, Murphy CF, Walters EJ, Willis L, Mullan NJ et al. Diagnostic utility of real-time smartphone ECG in the initial investigation of palpitations. British Journal of Cardiology. 2018; doi:10.5837/bjc.2018.006

56. Doliwa PS, Frykman V, Rosenqvist M. Short-term ECG for out of hospital detection of silent atrial fibrillation episodes. Scandinavian Cardiovascular Journal. 2009; 43(3):163-8

57. Dorr M, Nohturfft V, Brasier N, Bosshard E, Djurdjevic A, Gross S et al. The WATCH AF Trial: SmartWATCHes for Detection of Atrial Fibrillation. JACC: Clinical Electrophysiology. 2019; 5(2):199-208

58. Dussault C, Toeg H, Nathan M, Wang ZJ, Roux JF, Secemsky E. Electrocardiographic monitoring for detecting atrial fibrillation after ischemic stroke or transient ischemic attack. Circulation: Arrhythmia and Electrophysiology. 2015; 8(2):263-269

59. Edgerton JR, Mahoney C, Mack MJ, Roper K, Herbert MA. Long-term monitoring after surgical ablation for atrial fibrillation: How much is enough? Journal of Thoracic and Cardiovascular Surgery. 2011; 142(1):162-165

60. Eitel C, Husser D, Hindricks G, Fruhauf M, Hilbert S, Arya A et al. Performance of an implantable automatic atrial fibrillation detection device: Impact of software adjustments and relevance of manual episode analysis. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2011; 13(4):480-485

61. Elijovich L, Josephson SA, Fung GL, Smith WS. Intermittent atrial fibrillation may account for a large proportion of otherwise cryptogenic stroke: a study of 30-day cardiac event monitors. Journal of Stroke and Cerebrovascular Diseases. 2009; 18(3):185-9

62. Engdahl J, Andersson L, Mirskaya M, Rosenqvist M. Stepwise screening of atrial fibrillation in a 75-year-old population: implications for stroke prevention. Circulation. 2013; 127(8):930-937

63. Engdahl J, Holmen A, Rosenqvist M, Stromberg U. A prospective 5-year follow-up after population-based systematic screening for atrial fibrillation. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2018; 20(FI_3):f306-f311

64. Ermini G, Filippi A, Salera M. Switching from traditional to automatic sphygmomanometer increases opportunistic detection of atrial fibrillation in hypertensive patients. British Journal of Medical Practitioners. 2013; 6(1):A616

65. Etiwy M, Akhrass Z, Gillinov L, Alashi A, Wang R, Blackburn G et al. Accuracy of wearable heart rate monitors in cardiac rehabilitation. Cardiovascular Diagnosis and Therapy. 2019; 9(3):262-271

66. Evans GF, Shirk A, Muturi P, Soliman EZ. Feasibility of using mobile ECG recording technology to detect atrial fibrillation in low-resource settings. Global heart. 2017; 12(4):285-289

67. Eysenck W, Ammar A, Kanthasamy V, Freemantle N, Veasey RA, Patel NR et al. A trial of three non-invasive blood pressure monitors compared with invasive blood pressure assessment in atrial fibrillation and sinus rhythm. International Journal of Clinical Practice. 2019:e13410

68. Fallet S, Lemay M, Renevey P, Leupi C, Pruvot E, Vesin JM. Can one detect atrial fibrillation using a wrist-type photoplethysmographic device? Medical and Biological Engineering and Computing. 2019; 57(2):477-487

69. Fan YY, Li YG, Li J, Cheng WK, Shan ZL, Wang YT et al. Diagnostic performance of a smart device with photoplethysmography technology for atrial fibrillation detection: pilot study (Pre-mAFA II Registry). JMIR MHealth and UHealth. 2019; 7(3):e11437

70. Gaillard N, Deltour S, Vilotijevic B, Hornych A, Crozier S, Leger A et al. Detection of paroxysmal atrial fibrillation with transtelephonic EKG in TIA or stroke patients. Neurology. 2010; 74(21):1666-70

71. Gandolfo C, Balestrino M, Bruno C, Finocchi C, Reale N. Validation of a simple method for atrial fibrillation screening in patients with stroke. Neurological Sciences. 2015; 36(9):1675-1678

72. Godin R, Yeung C, Baranchuk A, Guerra P, Healey JS. Screening for atrial fibrillation using a mobile, single-lead electrocardiogram in canadian primary care clinics. Canadian Journal of Cardiology. 2019; 35(7):840-845

73. Gregg RE, Zhou SH, Lindauer JM, Feild DQ, Helfenbein ED. Where do derived precordial leads fail? Journal of Electrocardiology. 2008; 41(6):546-52

74. Grond M, Jauss M, Hamann G, Stark E, Veltkamp R, Nabavi D et al. Improved detection of silent atrial fibrillation using 72-hour holter ecg in patients with ischemic stroke: A prospective multicenter cohort study. Stroke. 2013; 44(12):3357-3364

75. Gunalp M, Atalar E, Coskun F, Yilmaz A, Aksoyek S, Aksu NM et al. Holter monitoring for 24 hours in patients with thromboembolic stroke and sinus rhythm diagnosed in the emergency department. Advances in Therapy. 2006; 23(6):854-860

76. Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y et al. Mobile health technology for atrial fibrillation screening using photoplethysmography-based smart devices: the HUAWEI heart study. Journal of the American College of Cardiology. 2019; 22:22

77. Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y et al. Mobile photoplethysmographic technology to detect atrial fibrillation. Journal of the American College of Cardiology. 2019; 74(19):2365-2375

78. Haberman ZC, Jahn RT, Bose R, Tun H, Shinbane JS, Doshi RN et al. Wireless smartphone ECG enables large-scale screening in diverse populations. Journal of Cardiovascular Electrophysiology. 2015; 26(5):520-526

79. Hald J, Poulsen PB, Qvist I, Holm L, Wedell-Wedellsborg D, Dybro L et al. Opportunistic screening for atrial fibrillation in a real-life setting in general practice in Denmark-The Atrial Fibrillation Found On Routine Detection (AFFORD) non-interventional study. PloS One. 2017; 12(11):e0188086

80. Hanke T, Charitos EI, Stierle U, Karluss A, Kraatz E, Graf B et al. Twenty-four-hour holter monitor follow-up does not provide accurate heart rhythm status after surgical atrial fibrillation ablation therapy: up to 12 months experience with a novel

permanently implantable heart rhythm monitor device. Circulation. 2009; 120(11 Suppl):S177-84

81. Harju J, Tarniceriu A, Parak J, Vehkaoja A, Yli-Hankala A, Korhonen I. Monitoring of heart rate and inter-beat intervals with wrist plethysmography in patients with atrial fibrillation. Physiological Measurement. 2018; 39(6):065007

82. Harris K, Edwards D, Mant J. How can we best detect atrial fibrillation? The journal of the Royal College of Physicians of Edinburgh. 2012; 42(Suppl 18):5-22

83. Hartikainen S, Lipponen JA, Hiltunen P, Rissanen TT, Kolk I, Tarvainen MP et al. Effectiveness of the chest strap electrocardiogram to detect atrial fibrillation. American Journal of Cardiology. 2019; 123(10):1643-1648

84. Haverkamp HT, Fosse SO, Schuster P. Accuracy and usability of single-lead ECG from smartphones - A clinical study. Indian Pacing and Electrophysiology Journal. 2019; 20:20

85. Hendrikx T, Rosenqvist M, Wester P, Sandstrom H, Hornsten R. Intermittent short ECG recording is more effective than 24-hour Holter ECG in detection of arrhythmias. BMC Cardiovascular Disorders. 2014; https://doi.org/10.1186/1471-2261-14-41

86. Hickey KT, Biviano AB, Garan H, Sciacca RR, Riga T, Warren K et al. Evaluating the utility of mHealth ECG heart monitoring for the detection and management of atrial fibrillation in clinical practice. Journal of Atrial Fibrillation. 2017; 9(5):1546

87. Higgins P, Dawson J, Macfarlane PW, McArthur K, Langhorne P, Lees KR. Predictive value of newly detected atrial fibrillation paroxysms in patients with acute ischemic stroke, for atrial fibrillation after 90 days. Stroke. 2014; 45(7):2134-2136

88. Higgins P, MacFarlane PW, Dawson J, McInnes GT, Langhorne P, Lees K. ECG monitoring strategy to identify AF after stroke. ISRCTN Registry. Springer Nature, 2010. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-00983010/full 21/01/2020

89. Himmelreich JCL, Karregat EPM, Lucassen WAM, van Weert HCPM, de Groot JR, Handoko ML et al. Diagnostic accuracy of a smartphone-operated, single-lead electrocardiography device for detection of rhythm and conduction abnormalities in primary care. Annals of Family Medicine. 2019; 17(5):403-411

90. Hindricks G, Pokushalov E, Urban L, Taborsky M, Kuck KH, Lebedev D et al. Performance of a new leadless implantable cardiac monitor in detecting and quantifying atrial fibrillation: Results of the XPECT trial. Circulation: Arrhythmia and Electrophysiology. 2010; 3(2):141-7

91. Hisazaki K, Miyazaki S, Hasegawa K, Kaseno K, Amaya N, Shiomi Y et al. The P wave morphology in lead V7 on the synthesized 18-lead ECG is a useful parameter for identifying arrhythmias originating from the right inferior pulmonary vein. Heart and Vessels. 2019; 35(2):246-251

92. Hobbs FD, Fitzmaurice DA, Mant J, Murray E, Jowett S, Bryan S et al. A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study. Health Technology Assessment. 2005; 9(40)

93. Hochstadt A, Chorin E, Viskin S, Schwartz AL, Lubman N, Rosso R. Continuous heart rate monitoring for automatic detection of atrial fibrillation with novel bio-sensing technology. Journal of Electrocardiology. 2019; 52:23-27

94. Ip JE. Wearable devices for cardiac rhythm diagnosis and management. JAMA. 2019; 321(4):337-338

95. Ip JH, Viqar-Syed M, Grimes D, Xie Y, Jager K, Boak J et al. Surveillance of AF recurrence post-surgical AF ablation using implantable cardiac monitor. Journal of Interventional Cardiac Electrophysiology. 2012; 33(1):77-83

96. Israel C, Kitsiou A, Kalyani M, Deelawar S, Ejangue LE, Rogalewski A et al. Detection of atrial fibrillation in patients with embolic stroke of undetermined source by prolonged monitoring with implantable loop recorders. Thrombosis and Haemostasis. 2017; 117(10):1962-1969

97. Israel CW, Hugl B, Unterberg C, Lawo T, Kennis I, Hettrick D et al. Pace-termination and pacing for prevention of atrial tachyarrhythmias: results from a multicenter study with an implantable device for atrial therapy. Journal of Cardiovascular Electrophysiology. 2001; 12(10):1121-8

98. Jabaudon D, Sztajzel J, Sievert K, Landis T, Sztajzel R. Usefulness of ambulatory 7-day ECG monitoring for the detection of atrial fibrillation and flutter after acute stroke and transient ischemic attack. Stroke. 2004; 35(7):1647-1651

99. Jacobs MS, Kaasenbrood F, Postma MJ, van Hulst M, Tieleman RG. Cost-effectiveness of screening for atrial fibrillation in primary care with a handheld, single-lead electrocardiogram device in the Netherlands. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2018; 20(1):12-18

100. Jiang K, Huang C, Ye SM, Chen H. High accuracy in automatic detection of atrial fibrillation for Holter monitoring. Journal of Zhejiang University: Science B. 2012; 13(9):751-756

101. Kaasenbrood F, Hollander M, Rutten FH, Gerhards LJ, Hoes AW, Tieleman RG. Yield of screening for atrial fibrillation in primary care with a hand-held, single-lead electrocardiogram device during influenza vaccination. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2016; 18(10):1514-1520

102. Kabutoya T, Imai Y, Hoshide S, Kario K. Diagnostic accuracy of a new algorithm to detect atrial fibrillation in a home blood pressure monitor. Journal of Clinical Hypertension. 2017; 19(11):1143-1147

103. Kabutoya T, Takahashi S, Watanabe T, Imai Y, Uemoto K, Yasui N et al. Diagnostic accuracy of an algorithm for detecting atrial fibrillation in a wrist-type pulse wave monitor. Journal of Clinical Hypertension. 2019; 21(9):1393-1398

104. Kaleschke G, Hoffmann B, Drewitz I, Steinbeck G, Naebauer M, Goette A et al. Prospective, multicentre validation of a simple, patient-operated electrocardiographic system for the detection of arrhythmias and electrocardiographic changes. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2009; 11(10):1362-8

105. Kalidas V, Tamil LS. Detection of atrial fibrillation using discrete-state Markov models and Random Forests. Computers in Biology and Medicine. 2019; 113:103386

106. Kallmunzer B, Bobinger T, Kahl N, Kopp M, Kurka N, Hilz MJ et al. Peripheral pulse measurement after ischemic stroke: A feasibility study. Neurology. 2014; 83(7):598-603

107. Kallmunzer B, Breuer L, Hering C, Raaz-Schrauder D, Kollmar R, Huttner HB et al. A structured reading algorithm improves telemetric detection of atrial fibrillation after acute ischemic stroke. Stroke. 2012; 43(4):994-999

108. Kane SA, Blake JR, McArdle FJ, Langley P, Sims AJ. Opportunistic detection of atrial fibrillation using blood pressure monitors: a systematic review. Open Heart. 2016; 3(1):e000362

109. Kang SH, Joe B, Yoon Y, Cho GY, Shin I, Suh JW. Cardiac auscultation using smartphones: pilot study. JMIR MHealth and UHealth. 2018; 6(2):e49

110. Kao WF, Hou SK, Huang CY, Chao CC, Cheng CC, Chen YJ. Assessment of the clinical efficacy of the heart spectrum blood pressure monitor for diagnosis of atrial fibrillation: An unblinded clinical trial. PloS One. 2018; 13(6):e0198852

111. Karaoguz MR, Yurtseven E, Aslan G, Deliormanli BG, Adiguzel O, Gonen M et al. The quality of ECG data acquisition, and diagnostic performance of a novel adhesive patch for ambulatory cardiac rhythm monitoring in arrhythmia detection. Journal of Electrocardiology. 2019; 54:28-35

112. Kashiwa A, Koyama F, Miyamoto K, Kamakura T, Wada M, Yamagata K et al. Performance of an atrial fibrillation detection algorithm using continuous pulse wave monitoring. Annals of Noninvasive Electrocardiology. 2019; 24(2):e12615

113. Kearley K, Selwood M, Van Den Bruel A, Thompson M, Mant D, Hobbs FDR et al. Triage tests for identifying atrial fibrillation in primary care: A diagnostic accuracy study comparing single-lead ECG and modified BP monitors. BMJ Open. 2014; 4(5):e004565

114. Kircher S, Hindricks G, Sommer P. Long-term success and follow-up after atrial fibrillation ablation. Current Cardiology Reviews. 2012; 8(4):354-361

115. Kishore A, Vail A, Majid A, Dawson J, Lees KR, Tyrrell PJ et al. Detection of atrial fibrillation after ischemic stroke or transient ischemic attack: a systematic review and meta-analysis. Stroke. 2014; 45(2):520-526

116. Kollias A, Destounis A, Kalogeropoulos P, Kyriakoulis KG, Ntineri A, Stergiou GS. Atrial fibrillation detection during 24-hour ambulatory blood pressure monitoring: comparison with 24-hour electrocardiography. Hypertension. 2018; 72(1):110-115

117. Koltowski L, Balsam P, Gllowczynska R, Rokicki JK, Peller M, Maksym J et al. Kardia Mobile applicability in clinical practice: A comparison of Kardia Mobile and standard 12-lead electrocardiogram records in 100 consecutive patients of a tertiary cardiovascular care center. Cardiology Journal. 2019; 15:15

118. Kong D, Zhu J, Wu S, Duan C, Lu L, Chen D. A novel IRBF-RVM model for diagnosis of atrial fibrillation. Computer Methods and Programs in Biomedicine. 2019; 177:183-192

119. Korompoki E, Del Giudice A, Hillmann S, Malzahn U, Gladstone DJ, Heuschmann P et al. Cardiac monitoring for detection of atrial fibrillation after TIA: A systematic review and meta-analysis. International Journal of Stroke. 2017; 12(1):33-45

120. Koshy AN, Sajeev JK, Nerlekar N, Brown AJ, Rajakariar K, Zureik M et al. Smart watches for heart rate assessment in atrial arrhythmias. International Journal of Cardiology. 2018; 266:124-127

121. Koshy AN, Sajeev JK, Nerlekar N, Brown AJ, Rajakariar K, Zureik M et al. Utility of photoplethysmography for heart rate estimation among inpatients. Internal Medicine Journal. 2018; 48(5):587-591

122. Kristensen AN, Jeyam B, Riahi S, Jensen MB. The use of a portable three-lead ECG monitor to detect atrial fibrillation in general practice. Scandinavian Journal of Primary Health Care. 2016:1-5

123. Krivoshei L, Weber S, Burkard T, Maseli A, Brasier N, Kuhne M et al. Smart detection of atrial fibrillation. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2017; 19(5):753-757

124. Kvist LM, Vinter N, Urbonaviciene G, Lindholt JS, Diederichsen ACP, Frost L. Diagnostic accuracies of screening for atrial fibrillation by cardiac nurses versus radiographers. Open Heart. 2019; 6(1):e000942

125. Kwon S, Hong J, Choi EK, Lee E, Hostallero DE, Kang WJ et al. Deep learning approaches to detect atrial fibrillation using photoplethysmographic signals: Algorithms development study. JMIR MHealth and UHealth. 2019; 7(6):e12770

126. Lahdenoja O, Hurnanen T, Iftikhar Z, Nieminen S, Knuutila T, Saraste A et al. Atrial fibrillation detection via accelerometer and gyroscope of a smartphone. IEEE journal of biomedical and health informatics. 2018; 22(1):108-118

127. Langley P, Dewhurst M, Di Marco LY, Adams P, Dewhurst F, Mwita JC et al. Accuracy of algorithms for detection of atrial fibrillation from short duration beat interval recordings. Medical Engineering and Physics. 2012; 34(10):1441-7

128. Lau J, Lowres N, Neubeck L, Brieger D, Sy R, Galloway C et al. Performance of an automated iPhone ECG algorithm to diagnose atrial fibrillation in a community AF screening program (search-AF). Heart Lung and Circulation. 2013; 22(Suppl 1):S205

129. Lauschke J, Busch M, Haverkamp W, Bulava A, Schneider R, Andresen D et al. New implantable cardiac monitor with three-lead ECG and active noise detection. Herz. 2017; 42(6):585-592

130. Lee J, Nam Y, McManus DD, Chon KH. Time-varying coherence function for atrial fibrillation detection. IEEE Transactions on Biomedical Engineering. 2013; 60(10):2783-2793

131. Levin LA, Husberg M, Sobocinski PD, Kull VF, Friberg L, Rosenqvist M et al. A cost-effectiveness analysis of screening for silent atrial fibrillation after ischaemic stroke. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2014; 17(2):207-214

132. Lewis M, Parker D, Weston C, Bowes M. Screening for atrial fibrillation: sensitivity and specificity of a new methodology. British Journal of General Practice. 2011; 61(582):38-9

133. Li KHC, White FA, Tipoe T, Liu T, Wong MC, Jesuthasan A et al. The current state of mobile phone apps for monitoring heart rate, heart rate variability, and atrial fibrillation: Narrative review. JMIR MHealth and UHealth. 2019; 7(2):e11606

134. Liao J, Khalid Z, Scallan C, Morillo C, O'Donnell M. Noninvasive cardiac monitoring for detecting paroxysmal atrial fibrillation or flutter after acute ischemic stroke: a systematic review. Stroke. 2007; 38(11):2935-40

135. Lin CT, Chang KC, Lin CL, Chiang CC, Lu SW, Chang SS et al. An intelligent telecardiology system using a wearable and wireless ECG to detect atrial fibrillation. IEEE Transactions on Information Technology in Biomedicine. 2010; 14(3):726-33

136. Liu J, Fang PH, Hou Y, Li XF, Liu Y, Wang YS et al. The value of transtelephonic electrocardiogram monitoring system during the "Blanking Period" after ablation of atrial fibrillation. Journal of Electrocardiology. 2010; 43(6):667-72

137. Lowe A, Oh TH, Stewart R. Screening for atrial fibrillation during automatic blood pressure measurements. IEEE Journal of Translational Engineering in Health and Medicine. 2018; 6:4400307

138. Lown M, Yue AM, Shah BN, Corbett SJ, Lewith G, Stuart B et al. Screening for atrial fibrillation using economical and accurate technology (from the SAFETY study). American Journal of Cardiology. 2018; 122(8):1339-1344

139. Lowres N, Mulcahy G, Gallagher R, Ben Freedman S, Marshman D, Kirkness A et al. Self-monitoring for atrial fibrillation recurrence in the discharge period post-cardiac surgery using an iPhone electrocardiogram. European Journal of Cardio-Thoracic Surgery. 2016; 50(1):44-51

140. Lowres N, Neubeck L, Salkeld G, Krass I, McLachlan AJ, Redfern J et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies: The SEARCH-AF study. Thrombosis and Haemostasis. 2014; 111(6):1167-1176

141. Lumikari TJ, Putaala J, Kerola A, Sibolt G, Pirinen J, Pakarinen S et al. Continuous 4-week ECG monitoring with adhesive electrodes reveals AF in patients with recent embolic stroke of undetermined source. Annals of Noninvasive Electrocardiology. 2019; 21(5):e12649

142. Mant J, Fitzmaurice DA, Hobbs FD, Jowett S, Murray ET, Holder R et al. Accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: analysis of data from screening for atrial fibrillation in the elderly (SAFE) trial. BMJ. 2007; 335(7616):380

143. Marazzi G, Iellamo F, Volterrani M, Lombardo M, Pelliccia F, Righi D et al. Comparison of Microlife BP A200 Plus and Omron M6 blood pressure monitors to detect atrial fibrillation in hypertensive patients. Advances in Therapy. 2012; 29(1):64-70

144. Martinek M, Aichinger J, Nesser HJ, Ziegler PD, Purerfellner H. New insights into long-term follow-up of atrial fibrillation ablation: full disclosure by an implantable pacemaker device. Journal of Cardiovascular Electrophysiology. 2007; 18(8):818-23

145. McManus DD, Chong JW, Soni A, Saczynski JS, Esa N, Napolitano C et al. PULSE-SMART: pulse-based arrhythmia discrimination using a novel smartphone application. Journal of Cardiovascular Electrophysiology. 2016; 27(1):51-7

146. McManus DD, Lee J, Maitas O, Esa N, Pidikiti R, Carlucci A et al. A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. Heart Rhythm. 2013; 10(3):315-319

147. Mehta DD, Nazir NT, Trohman RG, Volgman AS. Single-lead portable ECG devices: perceptions and clinical accuracy compared to conventional cardiac monitoring. Journal of Electrocardiology. 2015; 48(4):710-716

148. Miracapillo G, Addonisio L, Breschi M, F DES, Manfredini E, Corbucci G et al. Left axillary implantation of loop recorder versus the traditional left chest area: A prospective randomized study. Pacing and Clinical Electrophysiology. 2016; 39(8):830-6

149. Mittal S, Pokushalov E, Romanov A, Ferrara M, Arshad A, Musat D et al. Long-term ECG monitoring using an implantable loop recorder for the detection of atrial fibrillation after cavotricuspid isthmus ablation in patients with atrial flutter. Heart Rhythm. 2013; 10(11):1598-1604

150. Montenero AS, Quayyum A, Franciosa P, Mangiameli D, Antonelli A, Barbieri L et al. Implantable loop recorders: a novel method to judge patient perception of atrial fibrillation. Preliminary results from a pilot study. Journal of Interventional Cardiac Electrophysiology. 2004; 10(3):211-20

151. Morgan S, Mant D. Randomised trial of two approaches to screening for atrial fibrillation in UK general practice. British Journal of General Practice. 2002; 52(478):373-4, 377-80

152. Mulder AA, Wijffels MC, Wever EF, Kelder JC, Boersma LV. Arrhythmia detection after atrial fibrillation ablation: value of incremental monitoring time. Pacing and Clinical Electrophysiology.

153. Mulder BA, van Veldhuisen DJ, Crijns HJ, Bohm M, Cohen-Solal A, Babalis D et al. Effect of nebivolol on outcome in elderly patients with heart failure and atrial fibrillation: insights from SENIORS. European Journal of Heart Failure. 2012; 14(10):1171-8

154. Muller A, Scharner W, Borchardt T, Och W, Korb H. Reliability of an external loop recorder for automatic recognition and transtelephonic ECG transmission of atrial fibrillation. Journal of Telemedicine and Telecare. 2009; 15(8):391-6

155. Narasimha D, Hanna N, Beck H, Chaskes M, Glover R, Gatewood R et al. Validation of a smartphone-based event recorder for arrhythmia detection. Pacing and Clinical Electrophysiology. 2018; 41(5):487-494

156. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [Updated October 2018]. London. National Institute for Health and Care Excellence, 2014. Available from: https://www.nice.org.uk/process/pmg20/chapter/introduction-and-overview

157. Nault I, Andre P, Plourde B, Leclerc F, Sarrazin JF, Philippon F et al. Validation of a novel single lead ambulatory ECG monitor - CardiostatTM - Compared to a standard ECG Holter monitoring. Journal of Electrocardiology. 2019; 53:57-63

158. Nemati S, Ghassemi MM, Ambai V, Isakadze N, Levantsevych O, Shah A et al. Monitoring and detecting atrial fibrillation using wearable technology. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2016; 2016:3394-3397

159. Nigolian A, Dayal N, Nigolian H, Stettler C, Burri H. Diagnostic accuracy of multi-lead ECGs obtained using a pocket-sized bipolar handheld event recorder. Journal of Electrocardiology. 2018; 51(2):278-281

160. Nolker G, Mayer J, Boldt LH, Seidl K, V VAND, Massa T et al. Performance of an implantable cardiac monitor to detect atrial fibrillation: results of the DETECT AF study. Journal of Cardiovascular Electrophysiology. 2016; 27(12):1403-1410

161. Omboni S, Verberk WJ. Opportunistic screening of atrial fibrillation by automatic blood pressure measurement in the community. BMJ Open. 2016; 6(4):e010745

162. Oncu MR, Ozdemir F, Aydin SA. Evaluation of accuracy and reliability of electrocardiographs interpreted by emergency medicine assistants. Eastern Journal of Medicine. 2019; 24(4):427-433

163. Orchard J, Lowres N, Freedman SB, Ladak L, Lee W, Zwar N et al. Screening for atrial fibrillation during influenza vaccinations by primary care nurses using a smartphone electrocardiograph (iECG): A feasibility study. European Journal of Preventive Cardiology. 2016; 23(2 suppl):13-20

164. Osaka Y, Takigawa M, Takahashi A, Kuwahara T, Okubo K, Takahashi Y et al. The proportion of asymptomatic recurrence after catheter ablation of atrial fibrillation in patients with a pacemaker for sick sinus syndrome. Indian Pacing and Electrophysiology Journal. 2017; 17(5):125-131

165. Osako M, Sato T, Fujii H, Sumita T, Fujiwara H, Nakao Y et al. Reliability and efficacy of a monitoring system for an implanted pulse generator. Journal of Artificial Organs. 2002; 5(4):233-238

166. Pagola J, Juega J, Francisco-Pascual J, Moya A, Sanchis M, Bustamante A et al. Yield of atrial fibrillation detection with Textile Wearable Holter from the acute phase of stroke: pilot study of Crypto-AF registry. International Journal of Cardiology. 2018; 251:45-50

167. Park YM, Lee DI, Park HC, Shim J, Choi JI, Lim HE et al. Feasibility and accuracy of a new mobile electrocardiography device, ER-2000, in the diagnosis of arrhythmia. Journal of Arrhythmia. 2015; 31(4):201-9

168. Pastor-Perez FJ, Manzano-Fernandez S, Goya-Esteban R, Pascual-Figal DA, Barquero-Perez O, Rojo-Alvarez JL et al. Comparison of detection of arrhythmias in patients with chronic heart failure secondary to non-ischemic versus ischemic cardiomyopathy by 1 versus 7-day holter monitoring. American Journal of Cardiology. 2010; 106(5):677-81

169. Pedersen KB, Chemnitz A, Madsen C, Sandgaard NCF, Bak S, Brandes A. Low incidence of atrial fibrillation in patients with transient ischemic attack. Cerebrovascular Diseases Extra. 2016; 6(3):140-149

170. Perez-Valero J, Victoria Caballero Pintado M, Melgarejo F, Garcia-Sanchez AJ, Garcia-Haro J, Cordoba FG et al. Symbolic recurrence analysis of RR interval to detect atrial fibrillation. Journal of Clinical Medicine. 2019; 8(11):1840

171. Philippsen TJ, Christensen LS, Hansen MG, Dahl JS, Brandes A. Detection of subclinical atrial fibrillation in high-risk patients using an insertable cardiac monitor. JACC: Clinical Electrophysiology. 2017; 3(13):1557-1564

172. Plummer CJ, Henderson S, Gardener L, McComb JM. The use of permanent pacemakers in the detection of cardiac arrhythmias. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2001; 3(3):229-32

173. Plummer CJ, McComb JM, trial SA. Detection of atrial fibrillation by permanent pacemakers: observations from the STOP AF trial. Cardiac Electrophysiology Review. 2003; 7(4):333-40

174. Podd SJ, Sugihara C, Furniss SS, Sulke N. Are implantable cardiac monitors the 'gold standard' for atrial fibrillation detection? A prospective randomized trial comparing atrial fibrillation monitoring using implantable cardiac monitors and DDDRP permanent pacemakers in post atrial fibrillation ablation patients. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2016; 18(7):1000-5

175. Poh MZ, Poh YC, Chan PH, Wong CK, Pun L, Leung WW et al. Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms. Heart. 2018; 104(23):1921-1928

176. Poon K, Okin PM, Kligfield P. Diagnostic performance of a computer-based ECG rhythm algorithm. Journal of Electrocardiology. 2005; 38(3):235-8

177. Poulsen MB, Binici Z, Dominguez H, Soja AM, Kruuse C, Hornnes AH et al. Performance of short ECG recordings twice daily to detect paroxysmal atrial fibrillation in stroke and transient ischemic attack patients. International Journal of Stroke. 2017; 12(2):192-196

178. Proesmans T, Mortelmans C, Van Haelst R, Verbrugge F, Vandervoort P, Vaes B. Mobile phone-based use of the photoplethysmography technique to detect atrial

fibrillation in primary care: Diagnostic accuracy study of the FibriCheck app. JMIR MHealth and UHealth. 2019; 7(3):e12284

179. Proietti M, Farcomeni A, Goethals P, Scavee C, Vijgen J, Blankoff I et al. Cost-effectiveness and screening performance of ECG handheld machine in a population screening programme: The Belgian Heart Rhythm Week screening programme. European Journal of Preventive Cardiology. 2019; 26(9):964-972

180. Purerfellner H, Pokushalov E, Sarkar S, Koehler J, Zhou R, Urban L et al. P-wave evidence as a method for improving algorithm to detect atrial fibrillation in insertable cardiac monitors. Heart Rhythm. 2014; 11(9):1575-1583

181. Purerfellner H, Sanders P, Sarkar S, Reisfeld E, Reiland J, Koehler J et al. Adapting detection sensitivity based on evidence of irregular sinus arrhythmia to improve atrial fibrillation detection in insertable cardiac monitors. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2018; 20(FI_3):f321-f328

182. Rajakariar K, Koshy AN, Sajeev JK, Nair S, Roberts L, Teh AW. Modified positioning of a smartphone based single-lead electrocardiogram device improves detection of atrial flutter. Journal of Electrocardiology. 2018; 51(5):884-888

183. Ramkumar S, Nerlekar N, D'Souza D, Pol DJ, Kalman JM, Marwick TH. Atrial fibrillation detection using single lead portable electrocardiographic monitoring: a systematic review and meta-analysis. BMJ Open. 2018; 8(9):e024178

184. Reiffel JA, Schwarzberg R, Murry M. Comparison of autotriggered memory loop recorders versus standard loop recorders versus 24-hour Holter monitors for arrhythmia detection. American Journal of Cardiology. 2005; 95(9):1055-9

185. Reiffel JA, Verma A, Kowey PR, Halperin JL, Gersh BJ, Elkind MSV et al. Rhythm monitoring strategies in patients at high risk for atrial fibrillation and stroke: A comparative analysis from the REVEAL AF study. American Heart Journal. 2020; 219:128-136

186. Reinsch N, Ruprecht U, Buchholz J, Diehl RR, Kalsch H, Neven K. The BioMonitor 2 insertable cardiac monitor: clinical experience with a novel implantable cardiac monitor. Journal of Electrocardiology. 2018; 51(5):751-755

187. Rekhviashvili A, Baganashvili E, Tan KY, Raymakers F, Sakandelidze T. Reproducibility and diagnostic value of E100 event recorder for patients with complains on heart arrhythmias and no changes on multiple routine ECGs and 24-hour holter monitoring. Georgian Medical News. 2012; (203):29-33

188. Renier W, Geelen M, Steverlynck L, Wauters J, Aertgeerts B, Verbakel J et al. Can the heartscan be used for diagnosis and monitoring of emergencies in general practice? Acta Cardiologica. 2012; 67(5):525-531

189. Reverberi C, Rabia G, De Rosa F, Bosi D, Botti A, Benatti G. The RITMIATM smartphone app for automated detection of atrial fibrillation: accuracy in consecutive patients undergoing elective electrical cardioversion. BioMed Research International. 2019; 2019:4861951

190. Rhys GC, Azhar MF, Foster A. Screening for atrial fibrillation in patients aged 65 years or over attending annual flu vaccination clinics at a single general practice. Quality in Primary Care. 2013; 21(2):131-140

191. Ricci R, Puglisi A, Azzolini P, Spampinato A, Pignalberi C, Bellocci F et al. Reliability of a new algorithm for automatic mode switching from DDDR to DDIR pacing mode in sinus node disease patients with chronotropic incompetence and recurrent

paroxysmal atrial fibrillation. Pacing and Clinical Electrophysiology. 1996; 19(11 Pt 2):1719-23

192. Rincon F, Grassi PR, Khaled N, Atienza D, Sciuto D. Automated real-time atrial fibrillation detection on a wearable wireless sensor platform. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2012; 2012:2472-2475

193. Ritter MA, Kochhauser S, Duning T, Reinke F, Pott C, Dechering DG et al. Occult atrial fibrillation in cryptogenic stroke: detection by 7-day electrocardiogram versus implantable cardiac monitors. Stroke. 2013; 44(5):1449-1452

194. Rizos T, Rasch C, Jenetzky E, Hametner C, Kathoefer S, Reinhardt R et al. Detection of paroxysmal atrial fibrillation in acute stroke patients. Cerebrovascular Diseases. 2010; 30(4):410-7

195. Roche F, Gaspoz JM, Da Costa A, Isaaz K, Duverney D, Pichot V et al. Frequent and prolonged asymptomatic episodes of paroxysmal atrial fibrillation revealed by automatic long-term event recorders in patients with a negative 24-hour Holter. Pacing and Clinical Electrophysiology. 2002; 25(11):1587-93

196. Rojo-Martinez E, Sandin-Fuentes M, Calleja-Sanz AI, Cortijo-Garcia E, Garcia-Bermejo P, Ruiz-Pinero M et al. High performance of an implantable Holter monitor in the detection of concealed paroxysmal atrial fibrillation in patients with cryptogenic stroke and a suspected embolic mechanism. Revista de Neurología. 2013; 57(6):251-257

197. Rosenberg MA, Samuel M, Thosani A, Zimetbaum PJ. Use of a noninvasive continuous monitoring device in the management of atrial fibrillation: A pilot study. Pacing and Clinical Electrophysiology. 2013; 36(3):328-333

198. Ross LS, Bettin M, Kochhauser S, Ritter M, Minnerup J, Eckardt L et al. Sensitive detection of atrial fibrillation in acute stroke patients by short-term bedside electrocardiography monitoring software analysis. Cerebrovascular Diseases. 2018; 45(1-2):54-60

199. Roten L, Schilling M, Haberlin A, Seiler J, Schwick NG, Fuhrer J et al. Is 7-day event triggered ECG recording equivalent to 7-day Holter ECG recording for atrial fibrillation screening? Heart. 2012; 07

200. Rozen G, Vaid J, Hosseini SM, Kaadan MI, Rafael A, Roka A et al. Diagnostic accuracy of a novel mobile phone application for the detection and monitoring of atrial fibrillation. American Journal of Cardiology. 2018; 121(10):1187-1191

201. Ryabykina GV, Shokhzodaeva ZO, Sapelnikov OV, Makeev MI, Kozhemyakina ES, Shchedrina EV et al. Diagnostic utility of long-term remote ECG monitoring in compare with 24 hour Holter monitoring in patients with atrial fibrillation after catheter radiofrequency ablation in the early postoperative period. Terapevticheskii Arkhiv. 2018; 90(12):12-16

202. Sabar MI, Ara F, Henderson A, Ahmed O, Potter C, John I et al. A study to assess a novel automated electrocardiogram technology in screening for atrial fibrillation. Pacing and Clinical Electrophysiology. 2019; 03:03

203. Sack S, Mouton E, Defaye P, Dagres N, Wolfhard U, Wieneke H et al. Improved detection and analysis of sensed and paced events in dual chamber pacemakers with extended memory function. A prospective multicenter trial in 626 patients. Herz. 2001; 26(1):30-9

204. Salvatori V, Becattini C, Laureti S, Baglioni G, Germini F, Grilli P et al. Holter monitoring to detect silent atrial fibrillation in high-risk subjects: the Perugia General Practitioner Study. Internal and Emergency Medicine. 2015; 10(5):595-601

205. Samol A, Masin M, Gellner R, Otte B, Pavenstadt HJ, Ringelstein EB et al. Prevalence of unknown atrial fibrillation in patients with risk factors. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2013; 15(5):657-662

206. Sanak D, Hutyra M, Kral M, Bartkova A, Zapletalova J, Fedorco M et al. Atrial fibrillation in young ischemic stroke patients: an underestimated cause? European Neurology. 2015; 73(3-4):158-163

207. Sanak D, Hutyra M, Kral M, Bartkova A, Zapletalova J, Fedorco M et al. Paroxysmal atrial fibrillation in young cryptogenic ischemic stroke: A 3-week ECG Holter monitoring study. Biomedical Papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia. 2015; 159(2):283-287

208. Sanders P, Purerfellner H, Pokushalov E, Sarkar S, Di Bacco M, Maus B et al. Performance of a new atrial fibrillation detection algorithm in a miniaturized insertable cardiac monitor: Results from the Reveal LINQ Usability Study. Heart Rhythm. 2016; 13(7):1425-30

209. Schaefer JR, Leussler D, Rosin L, Pittrow D, Hepp T. Improved detection of paroxysmal atrial fibrillation utilizing a software-assisted electrocardiogram approach. PloS One. 2014; 9(2):e89328

210. Schuchert A, Behrens G, Meinertz T. Impact of long-term ECG recording on the detection of paroxysmal atrial fibrillation in patients after an acute ischemic stroke. Pacing and Clinical Electrophysiology. 1999; 22(7):1082-4

211. Schukraft S, Mancinetti M, Hayoz D, Faucherre Y, Cook S, Arroyo D et al. Handheld ECG Tracking of in-hOspital Atrial Fibrillation The HECTO-AF trial clinical study protocol. Trials [Electronic Resource]. 2019; 20(1):92

212. Seidl K, Meisel E, VanAgt E, Ottenhoff F, Hess M, Hauer B et al. Is the atrial high rate episode diagnostic feature reliable in detecting paroxysmal episodes of atrial tachyarrhythmias? Pacing and Clinical Electrophysiology. 1998; 21(4 Pt 1):694-700

213. Sejr MH, May O, Damgaard D, Sandal BF, Nielsen JC. External continuous ECG versus loop recording for atrial fibrillation detection in patients who had a stroke. Heart. 2019; 105(11):848-854

214. Sejr MH, Nielsen JC, Damgaard D, Sandal BF, May O. Atrial fibrillation detected by external loop recording for seven days or two-day simultaneous Holter recording: A comparison in patients with ischemic stroke or transient ischemic attack. Journal of Electrocardiology. 2017; 50(3):287-293

215. Selder JL, Breukel L, Blok S, van Rossum AC, Tulevski, II, Allaart CP. A mobile one-lead ECG device incorporated in a symptom-driven remote arrhythmia monitoring program. The first 5,982 Hartwacht ECGs. Netherlands Heart Journal. 2019; 27(1):38-45

216. Shafqat S, Kelly PJ, Furie KL. Holter monitoring in the diagnosis of stroke mechanism. Internal Medicine Journal. 2004; 34(6):305-9

217. Slocum J, Sahakian A, Swiryn S. Diagnosis of atrial fibrillation from surface electrocardiograms based on computer-detected atrial activity. Journal of Electrocardiology. 1992; 25(1):1-8

218. Solomon MD, Yang J, Sung SH, Livingston ML, Sarlas G, Lenane JC et al. Incidence and timing of potentially high-risk arrhythmias detected through long term continuous ambulatory electrocardiographic monitoring. BMC Cardiovascular Disorders. 2016; 16:35

219. Solosenko A, Petrenas A, Paliakaite B, Sornmo L, Marozas V. Detection of atrial fibrillation using a wrist-worn device. Physiological Measurement. 2019; 40(2):025003

220. Somerville S, Somerville J, Croft P, Lewis M. Atrial fibrillation: a comparison of methods to identify cases in general practice. British Journal of General Practice. 2000; 50(458):727-9

221. Sposato LA, Klein FR, Jauregui A, Ferrua M, Klin P, Zamora R et al. Newly diagnosed atrial fibrillation after acute ischemic stroke and transient ischemic attack: Importance of immediate and prolonged continuous cardiac monitoring. Journal of Stroke and Cerebrovascular Diseases. 2012; 21(3):210-216

222. Stahrenberg R, Weber-Kruger M, Seegers J, Edelmann F, Lahno R, Haase B et al. Enhanced detection of paroxysmal atrial fibrillation by early and prolonged continuous holter monitoring in patients with cerebral ischemia presenting in sinus rhythm. Stroke. 2010; 41(12):2884-8

223. Stergiou GS, Karpettas N, Protogerou A, Nasothimiou EG, Kyriakidis M. Diagnostic accuracy of a home blood pressure monitor to detect atrial fibrillation. Journal of Human Hypertension. 2009; 23(10):654-8

224. Sudlow M, Rodgers H, Kenny RA, Thomson R. Identification of patients with atrial fibrillation in general practice: a study of screening methods. BMJ. 1998; 317(7154):327-8

225. Suissa L, Lachaud S, Mahagne MH. Optimal timing and duration of continuous electrocardiographic monitoring for detecting atrial fibrillation in stroke patients. Journal of Stroke and Cerebrovascular Diseases. 2013; 22(7):991-995

226. Suissa L, Lachaud S, Mahagne MH. Continuous ECG monitoring for tracking down atrial fibrillation after stroke: Holter or automated analysis strategy? European Neurology. 2014; 72(1-2):7-12

227. Sutamnartpong P, Dharmasaroja PA, Ratanakorn D, Arunakul I. Atrial fibrillation and paroxysmal atrial fibrillation detection in patients with acute ischemic stroke. Journal of Stroke and Cerebrovascular Diseases. 2014; 23(5):1138-1141

228. Svennberg E, Stridh M, Engdahl J, Al-Khalili F, Friberg L, Frykman V et al. Safe automatic one-lead electrocardiogram analysis in screening for atrial fibrillation. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2017; 19(9):1449-1453

229. Swancutt D, Hobbs R, Fitzmaurice D, Mant J, Murray E, Jowett S et al. A randomised controlled trial and cost effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in the over 65s: (SAFE). BMC Cardiovascular Disorders. 2004; 4:12

230. Swerdlow CD, Schsls W, Dijkman B, Jung W, Sheth NV, Olson WH et al. Detection of atrial fibrillation and flutter by a dual-chamber implantable cardioverter-defibrillator. For the Worldwide Jewel AF Investigators. Circulation. 2000; 101(8):878-85

231. Taggar JS, Coleman T, Lewis S, Heneghan C, Jones M. Accuracy of methods for detecting an irregular pulse and suspected atrial fibrillation: A systematic review and meta-analysis. European Journal of Preventive Cardiology. 2016; 23(12):1330-8

232. Takagi T, Miyazaki S, Kusa S, Taniguchi H, Ichihara N, Iwasawa J et al. Role of extended external auto-triggered loop recorder monitoring for atrial fibrillation. Circulation Journal. 2014; 78(11):2637-2642

233. Tang SC, Huang PW, Hung CS, Shan SM, Lin YH, Shieh JS et al. Identification of atrial fibrillation by quantitative analyses of fingertip photoplethysmogram. Scientific Reports. 2017; 7:45644

234. Tarakji KG, Wazni OM, Callahan T, Kanj M, Hakim AH, Wolski K et al. Using a novel wireless system for monitoring patients after the atrial fibrillation ablation procedure: The iTransmit study. Heart Rhythm. 2015; 12(3):554-559

235. Tarniceriu A, Harju J, Yousefi ZR, Vehkaoja A, Parak J, Yli-Hankala A et al. The accuracy of atrial fibrillation detection from wrist photoplethysmography. A study on post-operative patients. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2018; 2018:1-4

236. Tavernier R, Wolf M, Kataria V, Phlips T, Huys R, Taghji P et al. Screening for atrial fibrillation in hospitalised geriatric patients. Heart. 2018; 104(7):588-593

237. Terranova P, Valli P, Terranova P, Dell'Orto S, Greco EM. Pacemaker prevention therapy in drug-refractory paroxysmal atrial fibrillation: reliability of diagnostics and effectiveness of prevention pacing therapy in Vitatron selection device. Indian Pacing and Electrophysiology Journal. 2006; 6(2):63-74

238. Tieleman RG, Plantinga Y, Rinkes D, Bartels GL, Posma JL, Cator R et al. Validation and clinical use of a novel diagnostic device for screening of atrial fibrillation. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2014; 16(9):1291-1295

239. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ et al. Passive detection of atrial fibrillation using a commercially available smartwatch. JAMA Cardiology. 2018; 3(5):409-416

240. Towhari J, Masud N, Alanazi H. Evaluation of the diagnostic accuracy of smartphone electrocardiogram recorder compared to standard 12 lead electrocardiography in hospital settings. Saudi Medical Journal. 2019; 40(6):575-581

241. Tu HT, Chen Z, Swift C, Churilov L, Guo R, Liu X et al. Smartphone electrographic monitoring for atrial fibrillation in acute ischemic stroke and transient ischemic attack. International Journal of Stroke. 2017; 12(7):786-789

242. Tu HT, Spence S, Kalman JM, Davis SM. Twenty-eight day Holter monitoring is poorly tolerated and insensitive for paroxysmal atrial fibrillation detection in cryptogenic stroke. Internal Medicine Journal. 2014; 44(5):505-508

243. Turakhia MP, Hoang DD, Zimetbaum P, Miller JD, Froelicher VF, Kumar UN et al. Diagnostic utility of a novel leadless arrhythmia monitoring device. American Journal of Cardiology. 2013; 112(4):520-524

244. Turakhia MP, Ullal AJ, Hoang DD, Than CT, Miller JD, Friday KJ et al. Feasibility of extended ambulatory electrocardiogram monitoring to identify silent atrial fibrillation in high-risk patients: The screening study for undiagnosed atrial fibrillation (STUDY-AF). Clinical Cardiology. 2015; 38(5):285-292

245. Vaes B, Stalpaert S, Tavernier K, Thaels B, Lapeire D, Mullens W et al. The diagnostic accuracy of the MyDiagnostick to detect atrial fibrillation in primary care. BMC Family Practice. 2014; 15:113

246. Valiaho ES, Kuoppa P, Lipponen JA, Martikainen TJ, Jantti H, Rissanen TT et al. Wrist band photoplethysmography in detection of individual pulses in atrial fibrillation and algorithm-based detection of atrial fibrillation. Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology. 2019; 21(7):1031-1038

247. Veale EL, Stewart AJ, Mathie A, Lall SK, Rees-Roberts M, Savickas V et al. Pharmacists detecting atrial fibrillation (PDAF) in primary care during the influenza vaccination season: a multisite, cross-sectional screening protocol. BMJ Open. 2018; 8(3):e021121

248. Velthuis BO, Bos J, Kraaier K, Stevenhagen J, Van Opstal JM, Van Der Palen J et al. Performance of an external transtelephonic loop recorder for automated detection of paroxysmal atrial fibrillation. Annals of Noninvasive Electrocardiology. 2013; 18(6):564-570

249. Verberk WJ, De Leeuw PW. Accuracy of oscillometric blood pressure monitors for the detection of atrial fibrillation: A systematic review. Expert Review of Medical Devices. 2012; 9(6):635-640

250. Verberk WJ, Omboni S, Kollias A, Stergiou GS. Screening for atrial fibrillation with automated blood pressure measurement: Research evidence and practice recommendations. International Journal of Cardiology. 2016; 203:465-73

251. Vukajlovic D, Bojovic B, Hadzievski L, George S, Gussak I, Panescu D. Wireless remote monitoring of atrial fibrillation using reconstructed 12-lead ECGs. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2010; 2010:1113-8

252. Wachter R. Finding atrial fibrillation in stroke - evaluation of enhanced and prolonged holter monitoring. 2013. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-00985610/full Last accessed: 04/02/2020

253. Wachter R, Weber-Kruger M, Seegers J, Edelmann F, Wohlfahrt J, Wasser K et al. Age-dependent yield of screening for undetected atrial fibrillation in stroke patients: The Find-AF study. Journal of Neurology. 2013; 260(8):2042-2045

254. Wang J, Wang P, Wang S. Automated detection of atrial fibrillation in ECG signals based on wavelet packet transform and correlation function of random process. Biomedical Signal Processing and Control. 2020; https://doi.org/10.1016/j.bspc.2019.101662

255. Wasserlauf J, You C, Patel R, Valys A, Albert D, Passman R. Smartwatch performance for the detection and quantification of atrial fibrillation. Circulation: Arrhythmia and Electrophysiology. 2019; 12(6):e006834

256. Welton NJ, McAleenan A, Thom HH, Davies P, Hollingworth W, Higgins JP et al. Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis. Health Technology Assessment. 2017; 21(29)

257. Wiegand UKH, Schneider R, Bode F, Brandes A, Diederich KW, Potratz J. Long-term atrioventricular synchrony in a single-lead VDD-pacemaker system: Comparison with a DDD-pacemaker in respect of their holter data-legends. European Journal of Cardiac Pacing and Electrophysiology. 1997; 7(4):183-191

258. Wiesel J, Abraham S, Messineo FC. Screening for asymptomatic atrial fibrillation while monitoring the blood pressure at home: Trial of regular versus irregular pulse for prevention of stroke (TRIPPS 2.0). American Journal of Cardiology. 2013; 111(11):1598-1601

259. Wiesel J, Arbesfeld B, Schechter D. Comparison of the microlife blood pressure monitor with the omron blood pressure monitor for detecting atrial fibrillation. American Journal of Cardiology. 2014; 114(7):1046-1048

260. Wiesel J, Fitzig L, Herschman Y, Messineo FC. Detection of atrial fibrillation using a modified microlife blood pressure monitor. American Journal of Hypertension. 2009; 22(8):848-52

261. Wiesel J, Wiesel D, Suri R, Messineo FC. The use of a modified sphygmomanometer to detect atrial fibrillation in outpatients. Pacing and Clinical Electrophysiology. 2004; 27(5):639-43

262. Wiesel J, Wiesel DJ, Messineo FC. Home monitoring with a modified automatic sphygmomanometer to detect recurrent atrial fibrillation. Journal of Stroke and Cerebrovascular Diseases. 2007; 16(1):8-13

263. William AD, Kanbour M, Callahan T, Bhargava M, Varma N, Rickard J et al. Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: The iREAD Study. Heart Rhythm. 2018; 15(10):1561-1565

264. Williams J, Pearce K, Benett I. The effectiveness of a mobile ECG device in identifying AF: sensitivity, specificity and predictive value. British Journal of Cardiology. 2015; 22(2):70-72

265. Willits I, Keltie K, Craig J, Sims A. WatchBP Home A for opportunistically detecting atrial fibrillation during diagnosis and monitoring of hypertension: A NICE medical technology guidance. Applied Health Economics and Health Policy. 2014; 12(3):255-265

266. Winkler S, Axmann C, Schannor B, Kim S, Leuthold T, Scherf M et al. Diagnostic accuracy of a new detection algorithm for atrial fibrillation in cardiac telemonitoring with portable electrocardiogram devices. Journal of Electrocardiology. 2011; 44(4):460-464

267. Yan BP, Lai WHS, Chan CKY, Chan SC, Chan LH, Lam KM et al. Contact-free screening of atrial fibrillation by a smartphone using facial pulsatile photoplethysmographic signals. Journal of the American Heart Association. 2018; 7(8):05

268. Yang P, Pu L, Yang L, Li F, Luo Z, Guo T et al. Value of implantable loop recorders in monitoring efficacy of radiofrequency catheter ablation in atrial fibrillation. Medical Science Monitor. 2016; 22:2846-2851

269. Yang X, Li S, Zhao X, Liu L, Jiang Y, Li Z et al. Atrial fibrillation is not uncommon among patients with ischemic stroke and transient ischemic stroke in China. BMC Neurology. 2017; 17(1):207

270. Yenikomshian M, Jarvis J, Patton C, Yee C, Mortimer R, Birnbaum H et al. Cardiac arrhythmia detection outcomes among patients monitored with the Zio patch system: a systematic literature review. Current Medical Research and Opinion. 2019:1

271. Zaprutko T, Zaprutko J, Baszko A, Sawicka D, Szalek A, Dymecka M et al. Feasibility of atrial fibrillation screening with mobile health technologies at pharmacies. Journal of Cardiovascular Pharmacology and Therapeutics. 2019; 25(2):142-151

272. Ziegler PD, Koehler JL, Mehra R. Comparison of continuous versus intermittent monitoring of atrial arrhythmias. Heart Rhythm. 2006; 3(12):1445-52

273. Ziegler PD, Rogers JD, Ferreira SW, Nichols AJ, Richards M, Koehler JL et al. Long-term detection of atrial fibrillation with insertable cardiac monitors in a real-world cryptogenic stroke population. International Journal of Cardiology. 2017; 244:175-179

Appendices

Appendix A: Review protocols

Table 15: Review protocol: Diagnostic accuracy of point of care devices

ID	Field	Content
0.	PROSPERO registration number	Not registered
1.	Review title	Accuracy of methods for detecting atrial fibrillation in people with cardiovascular risk factors for AF and/or symptoms suggestive of AF
2.	Review question	What are the most accurate methods for detecting atrial fibrillation in people with cardiovascular risk factors for AF and/or symptoms suggestive of AF?
3.	Objective	To identify the most accurate methods of detecting AF in this population in the primary care clinic. A variety of tests have recently become available that claim to diagnose AF. The accuracy of these need to be tested. Although each may be used in a different way (for example, some may be used at home by patients, or some may be applied in the clinic) it is important to have data on their accuracy. Issues around two-tier testing or location of testing will not be considered in this review. This review is simply a pragmatic attempt to survey the currently available diagnostic tools and to evaluate their accuracy relative to an appropriate reference standard. Once this is known then the GC can use this information to recommend 1) the tests that can be used and 2) how or where they may be used, perhaps as part of a two-stage approach [for example a test that is found to be very sensitive but non-specific might be appropriate as a first line test to ration who goes on to more definitive (but more resource-intensive) 12 lead testing].
4.	Searches	The following databases will be searched: Cochrane Central Register of Controlled Trials (CENTRAL) Cochrane Database of Systematic Reviews (CDSR) Embase MEDLINE Searches will be restricted by: English language Other searches: None The searches may be re-run 6 weeks before final submission of the review and further studies retrieved for inclusion if relevant. The full search strategies for MEDLINE database will be published in the final review.
5.	Condition or domain	Atrial Fibrillation

ID	Field	Content
	being studied	
6.	Population	People aged over 18 with symptoms suggestive of AF (including breathlessness, palpitations, syncope/dizziness, chest discomfort) and/or with cardiovascular risk factors for AF (including TIA, stroke, Heart Failure, hypertension, valve disease).
7.	Index Test	Any point of care tests used to detect AF For example (non-exhaustive list): • Manual pulse checking • Pulse oximeters • US devices • Blood pressure monitors o Microlife BPM o Watch BP Home A • Non-portable (but non-12 lead) ECG devices • Portable ECG devices o My Diagnostick o AliveCor Kardia • Smart portable devices eg phones, watches • 12 lead ECG (when gold standard is long-term loop recording – see section below) Where the same test is used with a differing number of recordings across studies, these should be regarded as separate test strategies, and should thus be dealt with separately. Tests using differing periods of recording will also be dealt with separately. For example, pulse oximeters for 2 minutes will be in a separate category of index test to pulse oximeters used for 1 hour, and they could be compared to each other as separate index tests.
8.	Comparator/Reference standard/Confounding factors	The reference standard that is used will determine the type of AF that the measured accuracy relates to. The analyses will therefore be stratified by the reference standards used, as follows: 1. 12-lead ECG, adjudicated by an expert clinician (usually cardiologist). This will theoretically pick up all constant AF but only a small proportion of intermittent AF cases. It is therefore really only useful for determining how well an index test can pick up constant AF. 2. Ambulatory monitoring for >24 hrs [NB: OR ANY DEVICE THAT GIVES A LONG-TERM RECORDING]. These should pick up all forms of AF. It is therefore a useful way of determining how well as test can pick up any AF. Unfortunately, it is likely that studies using this reference standard will be rare. NB: The ability of the tests to pick up AF vs no AF is being evaluated in this review, not the ability to differentiate between persistent and paroxysmal.
9.	Types of study to be included	Cross-sectional/prospective/retrospective diagnostic studies, or any study containing a diagnostic accuracy analysis

ID	Field	Content	
10.	Other exclusion criteria	Studies that do not report sensitivity and specificity, or insufficient data to derive these values. Non-English language studies.	
11.	Context	N/A	
12.	Primary outcomes (critical outcomes)	• Sensitivity • Specificity • Raw data to calculate 2x2 tables to calculate sensitivity and specificity (number of true positives, true negatives, false positives and false negatives).	
13.	Secondary outcomes (important outcomes)	None	
14.	Data extraction (selection and coding)	EndNote will be used for reference management, sifting, citations and bibliographies. All references identified by the searches and from other sources will be screened for inclusion. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. The full text of these potentially eligible studies will be retrieved and assessed in line with the criteria outlined above. A standardised form will be used to extract data from the included studies (see Developing NICE guidelines: the manual section 6.4). Data extraction will be independently quality assured by a second reviewer, discrepancies will be identified and resolved through discussion (with a third party where necessary).	
15.	Risk of bias (quality) assessment	Risk of bias quality assessment will be assessed using QUADAS-2. Assessment will be independently quality assured by a second reviewer. Disagreements between the reviewers will be resolved by discussion, with involvement of a third party where necessary.	
16.	Strategy for data synthesis	Where possible data will be meta-analysed where appropriate (if at least 3 studies reporting data at the same diagnostic threshold) in WinBUGS. Summary diagnostic outcomes will be reported from the meta-analyses with their 95% confidence intervals in adapted GRADE tables. Heterogeneity will be assessed by visual inspection of the sensitivity and specificity plots and summary area under the curve (AUC) plots. Particular attention will be placed on sensitivity, determined by the committee to be the primary outcome for decision making. If meta-analysis is not possible, data will be presented as individual values in adapted GRADE profile tables and plots of un-pooled sensitivity and specificity from RevMan software.	
17.	Analysis of sub-groups	If heterogeneity is identified, where data is available, subgroup analysis will be carried out for the following subgroups: Subgroups to investigate if heterogeneity is present 1. Expertise of index test interpreter (clinician trained in the use of the index test, such as cardiologist/electrophysiologist versus non-electrophysiologically trained clinician (e.g. GP) versus patient/carer) 2. Simultaneous index and gold std vs non simultaneous	
18.	Type and method of review	☐	Intervention
		☒	Diagnostic

ID	Field	Content		
		☐	Prognostic	
		☐	Qualitative	
		☐	Epidemiologic	
		☐	Service Delivery	
		☐	Other (please specify)	
19.	Language	English		
20.	Country	England		
21.	Anticipated or actual start date			
22.	Anticipated completion date			
23.	Stage of review at time of this submission	Review stage	Started	Completed
		Preliminary searches	☐	☑
		Piloting of the study selection process	☐	☑
		Formal screening of search results against eligibility criteria	☐	☑
		Data extraction	☐	☑
		Risk of bias (quality) assessment	☐	☑
		Data analysis	☐	☑
24.	Named contact	5a. Named contact National Guideline Centre 5b Named contact e-mail 5e Organisational affiliation of the review National Institute for Health and Care Excellence (NICE) and the National Guideline Centre		
25.	Review team members	From the National Guideline Centre: Sharon Swain Mark Perry Nicole Downes Sophia Kemmis Betty Elizabeth Pearton		
26.	Funding	This systematic review is being completed by the National		

ID	Field	Content	
	sources/sponsor	Guideline Centre which receives funding from NICE.	
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.	
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: [NICE guideline webpage].	
29.	Other registration details	N/A	
30.	Reference/URL for published protocol		
31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: notifying registered stakeholders of publication publicising the guideline through NICE's newsletter and alerts issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE. [Add in any additional agree dissemination plans.]	
32.	Keywords	Diagnosis, Atrial Fibrillation	
33.	Details of existing review of same topic by same authors	N/A	
34.	Current review status	☐	Ongoing
		☒	Completed but not published
		☐	Completed and published
		☐	Completed, published and being updated
		☐	Discontinued
35..	Additional information	N/A	
36.	Details of final publication	www.nice.org.uk	

1

2 **Table 16: Health economic review protocol**

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.

Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter – see appendix B below. For questions being updated from NICE guideline CG180, the search will be run from October 2013, which was the cut-off date for the searches. For questions being updated from the NICE guideline CG36 and for new questions, the search will be run from 2003.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2003, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Studies published after 2003 that were included in the previous guideline(s) will be reassessed for inclusion and may be included or selectively excluded based on their relevance to the questions covered in this update and whether more applicable evidence is also identified. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual.[156] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion** The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below. The health economist will be guided by the following hierarchies. *Setting:* • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden).

	• OECD countries with predominantly private health insurance systems (for example, Switzerland). • Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. *Health economic study type:* • Cost–utility analysis (most applicable). • Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be. • Studies published in 2003 or later (including any such studies included in the previous guideline(s)) but that depend on unit costs and resource data entirely or predominantly from before 2003 will be rated as 'Not applicable'. • Studies published before 2003 (including any such studies included in the previous guideline(s)) will be excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the health economic analysis:* • The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

Appendix B: Literature search strategies

This literature search strategy was used for the following reviews:

- **What are the most accurate methods for detecting atrial fibrillation in people with cardiovascular risk factors for AF and/or symptoms suggestive of AF?**

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual.[156]

For more information, please see the Methods Report published as part of the accompanying documents for this guideline.

B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies for interventions as these concepts may not be well described in title, abstract or indexes and therefore difficult to retrieve. Search filters were applied to the search where appropriate.

Table 17: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 31 December 2019	Exclusions Randomised controlled trials Systematic review studies Observational studies Diagnostic tests studies

Database	Dates searched	Search filter used
Embase (OVID)	1974 – 31 December 2019	Exclusions Randomised controlled trials Systematic review studies Observational studies Diagnostic tests studies
The Cochrane Library (Wiley)	Cochrane Reviews to 2019 Issue 12 of 12 CENTRAL to 2019 Issue 12 of 12	None

Medline (Ovid) search terms

1.	exp atrial fibrillation/
2.	((atrial or atria or atrium or auricular) adj3 fibrillat*).ti,ab.
3.	AF.ti,ab.
4.	1 or 2 or 3
5.	letter/
6.	editorial/
7.	news/
8.	exp historical article/
9.	Anecdotes as Topic/
10.	comment/
11.	case report/
12.	(letter or comment*).ti.
13.	or/5-12
14.	randomized controlled trial/ or random*.ti,ab.
15.	13 not 14
16.	animals/ not humans/
17.	exp Animals, Laboratory/
18.	exp Animal Experimentation/
19.	exp Models, Animal/
20.	exp Rodentia/
21.	(rat or rats or mouse or mice).ti.
22.	or/15-21
23.	4 not 22
24.	limit 23 to English language
25.	exp "sensitivity and specificity"/
26.	(sensitivity or specificity).ti,ab.
27.	((pre test or pretest or post test) adj probability).ti,ab.
28.	(predictive value* or PPV or NPV).ti,ab.
29.	likelihood ratio*.ti,ab.
30.	likelihood function/
31.	((area under adj4 curve) or AUC).ti,ab.
32.	(receive* operat* characteristic* or receive* operat* curve* or ROC curve*).ti,ab.
33.	(diagnos* adj3 (performance* or accurac* or utilit* or value* or efficien* or effectiveness)).ti,ab.
34.	gold standard.ab.
35.	or/25-34
36.	randomized controlled trial.pt.

37.	controlled clinical trial.pt.
38.	randomi#ed.ab.
39.	placebo.ab.
40.	randomly.ab.
41.	clinical trials as topic.sh.
42.	trial.ti.
43.	or/36-42
44.	Meta-Analysis/
45.	exp Meta-Analysis as Topic/
46.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
47.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
48.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
49.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
50.	(search* adj4 literature).ab.
51.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
52.	cochrane.jw.
53.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
54.	or/44-53
55.	Epidemiologic studies/
56.	Observational study/
57.	exp Cohort studies/
58.	(cohort adj (study or studies or analys* or data)).ti,ab.
59.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
60.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
61.	Controlled Before-After Studies/
62.	Historically Controlled Study/
63.	Interrupted Time Series Analysis/
64.	(before adj2 after adj2 (study or studies or data)).ti,ab.
65.	exp case control study/
66.	case control*.ti,ab.
67.	Cross-sectional studies/
68.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
69.	or/55-68
70.	24 and (35 or 43 or 54 or 69)
71.	((portable or ambulatory or monitor* or lead* or handheld or hand held or daily or long-term or short-term or strap* or device*) adj3 (ECG* or EKG* or electrocardio*)).ti,ab.
72.	((ECG* or EKG* or electrocardio*) adj2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure*)).ti,ab.
73.	(iECG* or Holter*).ti,ab.
74.	((ambulatory or event) adj monitor*).ti,ab.
75.	*electrocardiography/ or electrocardiography, ambulatory/
76.	(ILR* or loop record*).ti,ab.
77.	((heart or cardiac) adj monitor*).ti,ab.
78.	(pulse adj2 (assess* or check* or monitor* or detect* or screen* or diagnos* or

	measure* or palpation*)).ti,ab.
79.	(pulse oximetr* adj device*).ti,ab.
80.	oximetry/
81.	Pulse/
82.	((blood pressure or BP) adj2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure*)).ti,ab.
83.	Blood Pressure Monitors/ or Blood Pressure Monitoring, Ambulatory/
84.	(AliveCor or MyDiagnostic*).ti,ab.
85.	(Microlife or WatchBP or "watch BP").ti,ab.
86.	(Heartscan or Zenicor or AliveECG or Kardia*).ti,ab.
87.	(photoplethysmograph* or PPG).ti,ab.
88.	(smartwatch* or smart watch* or Applewatch* or Apple watch* or wrist watch* or wristwatch* or fitness band* or fitness tracker* or smartphone* or smart phone* or mobile phone*).ti,ab.
89.	(wearable adj2 (technology or device* or sensor* or ECG or EKG or electrocardio*)).ti,ab.
90.	(mhealth or m-health or "mobile health").ti,ab.
91.	telemedicine/
92.	point of care.ti,ab.
93.	((targeted or oppotunistic) adj2 (detect* or screen*)).ti,ab.
94.	or/71-93
95.	70 and 94

1 **Embase (Ovid) search terms**

1.	exp atrial fibrillation/
2.	((atrial or atria or atrium or auricular) adj3 fibrillat*).ti,ab.
3.	AF.ti,ab.
4.	1 or 2 or 3
5.	letter.pt. or letter/
6.	note.pt.
7.	editorial.pt.
8.	case report/ or case study/
9.	(letter or comment*).ti.
10.	or/5-9
11.	randomized controlled trial/ or random*.ti,ab.
12.	10 not 11
13.	animal/ not human/
14.	nonhuman/
15.	exp Animal Experiment/
16.	exp Experimental Animal/
17.	animal model/
18.	exp Rodent/
19.	(rat or rats or mouse or mice).ti.
20.	or/12-19
21.	4 not 20
22.	limit 21 to English language
23.	exp "sensitivity and specificity"/
24.	(sensitivity or specificity).ti,ab.
25.	((pre test or pretest or post test) adj probability).ti,ab.

26.	(predictive value* or PPV or NPV).ti,ab.
27.	likelihood ratio*.ti,ab.
28.	((area under adj4 curve) or AUC).ti,ab.
29.	(receive* operat* characteristic* or receive* operat* curve* or ROC curve*).ti,ab.
30.	(diagnos* adj3 (performance* or accurac* or utilit* or value* or efficien* or effectiveness)).ti,ab.
31.	diagnostic accuracy/
32.	diagnostic test accuracy study/
33.	gold standard.ab.
34.	or/23-33
35.	random*.ti,ab.
36.	factorial*.ti,ab.
37.	(crossover* or cross over*).ti,ab.
38.	((doubl* or singl*) adj blind*).ti,ab.
39.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
40.	crossover procedure/
41.	single blind procedure/
42.	randomized controlled trial/
43.	double blind procedure/
44.	or/35-43
45.	systematic review/
46.	Meta-Analysis/
47.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
48.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
49.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
50.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
51.	(search* adj4 literature).ab.
52.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
53.	cochrane.jw.
54.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
55.	or/45-54
56.	Clinical study/
57.	Observational study/
58.	family study/
59.	longitudinal study/
60.	retrospective study/
61.	prospective study/
62.	cohort analysis/
63.	follow-up/
64.	cohort*.ti,ab.
65.	63 and 64
66.	(cohort adj (study or studies or analys* or data)).ti,ab.
67.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
68.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or

	review or analys* or cohort* or data)).ti,ab.
69.	(before adj2 after adj2 (study or studies or data)).ti,ab.
70.	exp case control study/
71.	case control*.ti,ab.
72.	cross-sectional study/
73.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
74.	or/56-73
75.	34 or 44 or 55 or 74
76.	22 and (34 or 44 or 55 or 74)
77.	((portable or ambulatory or monitor* or lead* or handheld or hand held or daily or long-term or short-term or strap* or device*) adj3 (ECG* or EKG* or electrocardio*)).ti,ab.
78.	((ECG* or EKG* or electrocardio*) adj2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure*)).ti,ab.
79.	(iECG* or Holter*).ti,ab.
80.	((ambulatory or event) adj monitor*).ti,ab.
81.	*electrocardiography/
82.	*ambulatory electrocardiography/
83.	(ILR* or loop record*).ti,ab.
84.	((heart or cardiac) adj monitor*).ti,ab.
85.	(pulse adj2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure* or palpation*)).ti,ab.
86.	(pulse oximetr* adj device*).ti,ab.
87.	*oximetry/
88.	*pulse rate/
89.	((blood pressure or BP) adj2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure*)).ti,ab.
90.	*blood pressure monitor/
91.	(AliveCor or MyDiagnostic*).ti,ab.
92.	(Microlife or WatchBP or "watch BP").ti,ab.
93.	(Heartscan or Zenicor or AliveECG or Kardia*).ti,ab.
94.	(photoplethysmograph* or PPG).ti,ab.
95.	(smartwatch* or smart watch* or Applewatch* or Apple watch* or wrist watch* or wristwatch* or fitness band* or fitness tracker* or smartphone* or smart phone* or mobile phone*).ti,ab.
96.	(wearable adj2 (technology or device* or sensor* or ECG or EKG or electrocardio*)).ti,ab.
97.	(mhealth or m-health or "mobile health").ti,ab.
98.	*telemedicine/
99.	point of care.ti,ab.
100.	((targeted or oppotunistic) adj2 (detect* or screen*)).ti,ab.
101.	or/77-100
102.	76 and 101

1 **Cochrane Library (Wiley) search terms**

#1.	MeSH descriptor: [Atrial Fibrillation] explode all trees
#2.	((atrial or atria or atrium or auricular) near/3 fibrillat*):ti,ab
#3.	AF:ti,ab
#4.	#1 or #2 or #3
#5.	((portable or ambulatory or monitor* or lead* or handheld or hand held or daily or long-term or short-term or strap* or device*) near/3 (ECG* or EKG* or electrocardio*)):ti,ab

#6.	((ECG* or EKG* or electrocardio*) near/2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure*)):ti,ab
#7.	(iECG* or Holter*):ti,ab
#8.	((ambulatory or event) next monitor*).ti,ab
#9.	MeSH descriptor: [Electrocardiography] this term only
#10.	MeSH descriptor: [Electrocardiography, Ambulatory] this term only
#11.	(ILR* or loop record*):ti,ab
#12.	((heart or cardiac) next monitor*):ti,ab
#13.	(pulse near/2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure* or palpation*)):ti,ab
#14.	(pulse oximetr* next device*).ti,ab
#15.	MeSH descriptor: [Oximetry] this term only
#16.	MeSH descriptor: [Pulse] this term only
#17.	((blood pressure or BP) near/2 (assess* or check* or monitor* or detect* or screen* or diagnos* or measure*)).ti,ab
#18.	MeSH descriptor: [Blood Pressure Monitors] this term only
#19.	MeSH descriptor: [Blood Pressure Monitoring, Ambulatory] this term only
#20.	(AliveCor or MyDiagnostic*):ti,ab
#21.	(Microlife or WatchBP or "watch BP"):ti,ab
#22.	(Heartscan or Zenicor or AliveECG or Kardia*):ti,ab
#23.	(photoplethysmograph* or PPG):ti,ab
#24.	(smartwatch* or smart watch* or Applewatch* or Apple watch* or wrist watch* or wristwatch* or fitness band* or fitness tracker* or smartphone* or smart phone* or mobile phone*):ti,ab
#25.	(wearable near/2 (technology or device* or sensor* or ECG or EKG or electrocardio*)):ti,ab
#26.	(mhealth or m-health or "mobile health"):ti,ab
#27.	MeSH descriptor: [Telemedicine] this term only
#28.	point of care:ti,ab
#29.	((targeted or oppotunistic) near/2 (detect* or screen*)):ti,ab
#30.	(or #5-#29)
#31.	#4 and #30

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to the Atrial Fibrillation population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA). NHS EED and HTA databases are hosted by the Centre for Research and Dissemination (CRD). Additional health economics searches were run on Medline and Embase.

Table 18: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	2003– 31 December 2019	Exclusions Health economics studies
Embase	2003– 31 December 2019	Exclusions Health economics studies
Centre for Research and Dissemination (CRD)	NHSEED - 2003 to March 2015 HTA - 2003 –31 December 2019	None

1 **Medline (Ovid) search terms**

1.	exp atrial fibrillation/
2.	((atrial or atria or atrium or auricular) adj3 fibrillat*).ti,ab.
3.	AF.ti,ab.
4.	1 or 2 or 3
5.	letter/
6.	editorial/
7.	news/
8.	exp historical article/
9.	Anecdotes as Topic/
10.	comment/
11.	case report/
12.	(letter or comment*).ti.
13.	or/5-12
14.	randomized controlled trial/ or random*.ti,ab.
15.	13 not 14
16.	animals/ not humans/
17.	exp Animals, Laboratory/
18.	exp Animal Experimentation/
19.	exp Models, Animal/
20.	exp Rodentia/
21.	(rat or rats or mouse or mice).ti.
22.	or/15-21
23.	4 not 22
24.	limit 23 to English language
25.	economics/
26.	value of life/
27.	exp "costs and cost analysis"/
28.	exp Economics, Hospital/
29.	exp Economics, medical/
30.	Economics, nursing/
31.	economics, pharmaceutical/
32.	exp "Fees and Charges"/
33.	exp budgets/
34.	budget*.ti,ab.
35.	cost*.ti.
36.	(economic* or pharmaco?economic*).ti.
37.	(price* or pricing*).ti,ab.
38.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
39.	(financ* or fee or fees).ti,ab.
40.	(value adj2 (money or monetary)).ti,ab.
41.	or/25-40
42.	24 and 41

1 **Embase (Ovid) search terms**

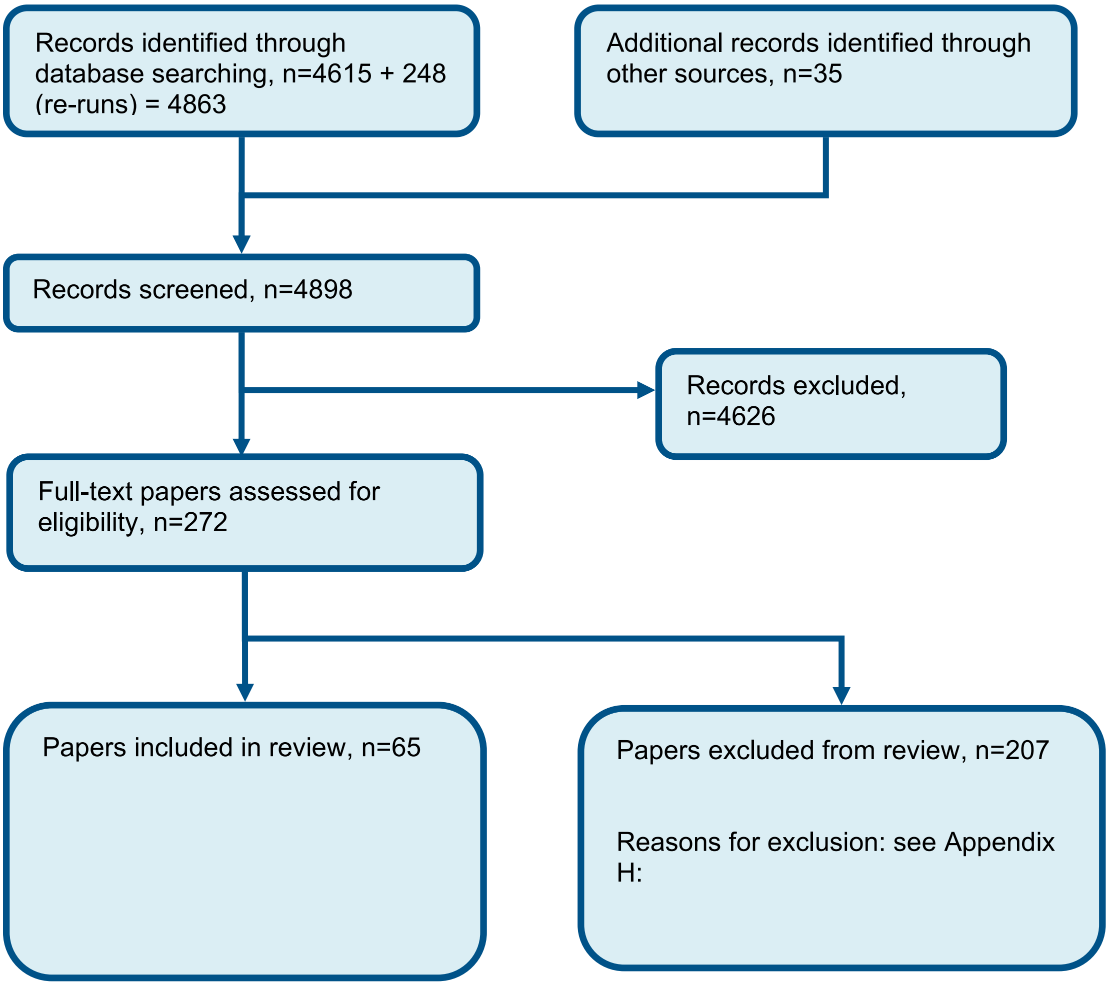
1.	exp atrial fibrillation/
2.	((atrial or atria or atrium or auricular) adj3 fibrillat*).ti,ab.
3.	AF.ti,ab.
4.	1 or 2 or 3
5.	letter.pt. or letter/
6.	note.pt.
7.	editorial.pt.
8.	case report/ or case study/
9.	(letter or comment*).ti.
10.	or/5-9
11.	randomized controlled trial/ or random*.ti,ab.
12.	10 not 11
13.	animal/ not human/
14.	nonhuman/
15.	exp Animal Experiment/
16.	exp Experimental Animal/
17.	animal model/
18.	exp Rodent/
19.	(rat or rats or mouse or mice).ti.
20.	or/12-19
21.	4 not 20
22.	limit 21 to English language
23.	health economics/
24.	exp economic evaluation/
25.	exp health care cost/
26.	exp fee/
27.	budget/
28.	funding/
29.	budget*.ti,ab.
30.	cost*.ti.
31.	(economic* or pharmaco?economic*).ti.
32.	(price* or pricing*).ti,ab.
33.	(cost* adj2 (effectiv* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
34.	(financ* or fee or fees).ti,ab.
35.	(value adj2 (money or monetary)).ti,ab.
36.	or/23-35
37.	22 and 36

2 **NHS EED and HTA (CRD) search terms**

#1.	MeSH DESCRIPTOR Atrial Fibrillation EXPLODE ALL TREES
#2.	(((atrial or atria or atrium or auricular) adj3 fibrillat*))
#3.	(AF)
#4.	(#1 or #2 or #3)

Appendix C: Clinical evidence selection

Figure 1: Flow chart of clinical study selection for the review



Appendix D: Clinical evidence tables

Table 19: Gandolfo, 2015[71]

<table>
<tr><th>Reference</th><th>Gandolfo, 2015[71]</th></tr>
<tr><td>Study type</td><td>Observational</td></tr>
<tr><td>Recruitment</td><td>consecutive</td></tr>
<tr><td>Setting</td><td>Stroke unit</td></tr>
<tr><td>Country</td><td>Italy</td></tr>
<tr><td>Sample size</td><td>207</td></tr>
<tr><td>Sample characteristics</td><td>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</td></tr>
<tr><td>Inclusion criteria</td><td>Patients admitted to stroke unit because of recent (<48 hours) TIA/stroke</td></tr>
<tr><td>Exclusion criteria</td><td>Patients with rhythm controlled by pacemakers or defibrillators</td></tr>
<tr><td>Index test(s), including number of repetitions and duration</td><td>Triple blood pressure measurement by the <u>Microlife AFib device</u> (total session time 10 minutes) usually on day of admission to SU, and <48hrs. Done by trained SU nurse</td></tr>
<tr><td>Gold standard</td><td><u>Standard 12 lead ECG</u>, interpreted by expert cardiologist (ECG performed by trained SU nurse). Normally done on day of admission to SU, and <48 hours of admission.</td></tr>
<tr><td>Expertise of index test interpreter</td><td>Not stated</td></tr>
<tr><td>Simultaneous index/gold vs non simultaneous</td><td>Not simultaneous; gold standard followed index ‘immediately after the end’ during a 10 minute evaluation session, and never>48 hours difference between them.</td></tr>
<tr><td>Results</td><td>
<table>
<tr><th></th><th>Gold standard +ve</th><th>Gold standard -ve</th><th>Total</th><td rowspan="2">Sensitivity:0.895 (0.760-0.958)
Specificity:0.988(0.958-0.997)
PPV: 0.94
NPV:0.98</td></tr>
<tr><td>Index test +ve</td><td>34</td><td>2</td><td>36</td></tr>
</table>
</td></tr>
</table>

Reference	Gandolfo, 2015[71]			
	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[104]

Reference	Kaleschke, 2009[104]
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
Expertise of index	Analysed by single blinded observer.

Reference	Kaleschke, 2009[104]				
test interpreter					
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[113]

Reference	Kearley, 2014[113]
Study type	Observational
Recruitment	consecutive
Setting	GP practices
Country	UK
Sample size	1000
Sample	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

Reference	Kearley, 2014[113]
characteristics	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[113]
Results	*If unclear on index test it was always counted as a positive test* <u>Watch BP</u> 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 <u>Omron auto-analysis</u> 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 <u>Omron ECG trace interpreted by the 4 cardiologists (pooled results using meta-analysis of the 4 cardiologists results)</u> Sensitivity: 94.4% Specificity: 94.6% <u>Merlin ECG trace interpreted by the 4 cardiologists (pooled results using meta-analysis of the 4 cardiologists results)</u> 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[116]

Reference	Kollias, 2018[116]
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
Simultaneous index/gold vs non	Simultaneous

Reference	Kollias, 2018[116]
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 [143]

Reference	Marazzi, 2012[143]
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 of repetitions and duration	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 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

Reference	Marazzi, 2012[143]
	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[117]

Reference	Koltowski, 2019[117]
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 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%

Reference	Koltowski, 2019[117]
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[122]

Reference	Kristensen, 2016[122]

Reference	Kristensen, 2016[122]
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[122]
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 [259]

Reference	Wiesel, 2014[259]
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[259]
	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[258]

Reference	Wiesel, 2013[258]
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).
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

Reference	Wiesel, 2013[258]
	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

1
2 **Table 28** Wiesel, 2009[260]

Reference	Wiesel, 2009[260]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology outpatients
Country	USA
Sample size	405
Sample	Mean age 73; male 51%; white 82%; black 8%; other 10%; CHF 6.7%; Hypertension 51.6%; DM 14.8%; CAD 37.3%

Reference	Wiesel, 2009[260]
characteristics	
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	Single readings of microlife (n=3 x 405 readings) Sensitivity: 95.3(92.8 to 97.6), Specificity 86.4 (84.3 to 88.7); TP 266, FN 13, FP 127, TN 809 3 readings (majority rule) of microlife (n=405) 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 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]

1

1 **Table 29** Vaes, 2014[245]

Reference	**Vaes, 2014[245]**
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 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

Reference	Vaes, 2014[245]
	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[220]

Reference	Somerville, 2000[220]
Study type	Observational
Recruitment	Selective case/control
Setting	One GP surgery
Country	UK
Sample size	86
Sample	30% with AF; no other details provided

Reference	Somerville, 2000[220]
characteristics	
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) 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)

Reference	Somerville, 2000[220]
	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)

1

2 **Table 31** Wiesel, 2004[261]

Reference	Wiesel, 2004[261]
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
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.

Reference	Wiesel, 2004[261]
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[142]

Reference	Mant, 2007[142]
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
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

Reference	Mant, 2007[142]
	• 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). 12 lead interpretive software (Biolog interpretive software) Sensitivity: 83.3(78.3-88.2) Specificity:99.1(98.7-99.5) TP 179, FN 36, FP 21, TN 2320 12 lead interpreted by GP Sensitivity:79.8(70.5-87.2) Specificity:91.6(90.1-93.1) TP 79, FN 22, FP 114, TN 1241 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)

Reference	Mant, 2007[142]
	Specificity:86.4(84.6-88.3) TP 112, FN 20, FP 180, TN 1145 <u>12 lead interpreted by practice nurse</u> Sensitivity:77.1(67.4-85) Specificity:85.1(83-86.9) TP 74, FN 22, FP 198, TN 1132 <u>Limb lead ECG interpreted by practice nurse</u> Sensitivity:72.0(63.9-80.1) Specificity:83.4(81.4-85.4) TP 85, FN 33, FP 220, TN 1107 <u>Chest lead ECG interpreted by practice nurse</u> Sensitivity:68.7(60.1-76.4) Specificity:82.8(80.7-84.8) TP 92, FN 42, FP 22, TN 1066 <u>12 lead interpretive software combined with GP interpretation (positive if either or both is positive)</u> 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

1

2 **Table 33** Lown, 2018[138]

Reference	Lown, 2018[138]
Study type	Observational
Recruitment	Selective case/control

Reference	Lown, 2018[138]
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
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	Unclear, but unlikely

Reference	Lown, 2018[138]
index/gold vs non simultaneous	
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 [267]

Reference	Yan, 2018[267]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology inpatients
Country	Hong Kong
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

Reference	Yan, 2018[267]
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	<u>Fingertip plethysmography</u> Sensitivity 94.7(87.1-97.9); Specificity 93(87.5-96.1); TP 71, FN 4, TN 132, FP 10 <u>Facial photoplethysmography</u> 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 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[238]

Reference	Tieleman, 2014[238]
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
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

Reference	Tieleman, 2014[238]
	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[223]

Reference	Stergiou, 2009[223]
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.
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

Reference	Stergiou, 2009[223]
	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	Using just the first reading per patient (thus modelling the accuracy if just one BP measure is done): Sensitivity:0.93 (0.74-0.99); specificity 0.89(0.76-0.96); TP 25, FN 2, FP 5, TN 40 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]: Sensitivity:1.00 (0.84-1); specificity 0.76(0.60-0.87); TP 27, FN 0, FP 11, TN 34 Using all 3 readings per patient (thus modelling the accuracy if 3 BP measures are done) [AF diagnosis if just one is positive]: Sensitivity:1.00 (0.84-1); specificity 0.69(0.53-0.81); TP 27, FN 0, FP 14, TN 31 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']: 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[24]

Reference	Bumgarner, 2018[24]
Study type	Observational case control.
Recruitment	Selective case/control

Reference	Bumgarner, 2018[24]
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
Results	<u>KB algorithm automatic reading (this is the most relevant as this will be the most likely way it is used clinically)</u> *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

Reference	Bumgarner, 2018[24]
	<u>KB algorithm reading interpreted by electrophysiologists</u> *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[27]

Reference	Caldwell, 2012[27]
Study type	Case-control observational
Recruitment	Selective case/control
Setting	Anticoagulation outpatient clinic
Country	UK
Sample size	157
Sample characteristics	Not reported
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	• 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

Reference	Caldwell, 2012[27]
duration	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 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

Reference	Caldwell, 2012[27]
	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[69]

Reference	Fan, 2019[69]
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 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

Reference	Fan, 2019[69]
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 [6]

Reference	Arevalo-Manso, 2016[6]
Study type	Consecutive, observational
Recruitment	consecutive
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

Reference	Arevalo-Manso, 2016[6]
	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
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.

Reference	Arevalo-Manso, 2016[6]
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[53]

Reference	Desteghe, 2017[53]
Study type	Obervational
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 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.

Reference	Desteghe, 2017[53]
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** <u>Automated with implanted device [PM/ICD] patients included</u> Sensitivity: 60.5% Specificity: 93.3% (TP 23, FN 15, FP 19, TN 263) <u>Automated with PM/ICD patients excluded</u> Sensitivity: 81.8% Specificity: 94.2% (TP 18, FN 4, FP 14, TN 229)

Reference	Desteghe, 2017[53]
	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) 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

Reference	Desteghe, 2017[53]
	Sensitivity: 90.9% Specificity: 94.7% (TP 20, FN 0, FP 5, TN 230) 10 illegible – these are taken into account when calculating accuracy <u>Electrophysiologist 2 with PM/ICD patients included</u> Sensitivity: 63.2% Specificity: 95.7% (TP 24, FN 14, FP 4, TN 270) 8 illegible – these are taken into account when calculating accuracy <u>Electrophysiologist 2 with PM/ICD patients excluded</u> 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 [84]

Reference	Haverkamp, 2019[84]
Study type	Observational
Recruitment	consecutive
Setting	Cardiac inpatients
Country	Norway
Sample size	94
Sample	37% female; mean age 58;

Reference	Haverkamp, 2019[84]
characteristics	
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[145]

Reference	McManus, 2016[145]
Study type	Observational
Recruitment	People before and after a cardioversion – thus very much a case-control situation
Settings	Cardiac inpatients
Country	USA
Sample size	128

Reference	McManus, 2016[145]
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 labeled 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 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

Reference	McManus, 2016[145]
	AF (other than just age) and/or symptoms suggestive of AF]

Table 44 Muller, 2009[154]

Reference	Muller, 2009[154]
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
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[167]

Reference	Park, 2015[167]
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.
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

Reference	Park, 2015[167]
	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[199]

Reference	Roten, 2012[199]
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 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.

Reference	Roten, 2012[199]
	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[178]

Reference	Proesmans, 2019[178]
Study type	Observational
Recruitment	Selective case/control
Setting	GP centres

Reference	Proesmans, 2019[178]
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.
Expertise of index test interpreter	Researcher so likely to have high expertise
Simultaneous index/gold vs non simultaneous	Unclear – no mention of synchronicity

Reference	Proesmans, 2019[178]
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[200]

Reference	Rozen, 2018[200]
Study type	Observational – case control
Recruitment	Selective case/control

Reference	Rozen, 2018[200]
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
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

Reference	Rozen, 2018[200]
	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[213]

Reference	Sejr, 2019[213]
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 of repetitions and duration	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 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

Reference	Sejr, 2019[213]
	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[153]

Reference	Mulder, 2012[153]
Study type	Observational
Recruitment	consecutive
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

Reference	Mulder, 2012[153]
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 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[110]

Reference	Kao, 2018[110]
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 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.

Reference	Kao, 2018[110]
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[146]

Reference	McManus, 2013[146]
Study type	Observational
Recruitment	Selective case/control (paired)
Setting	Cardioversion patients
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),	iPhone 4S camera. Placed directly on right index or second finger for 2 minutes while AF detection application was run. Pulse signal

Reference	McManus, 2013[146]
including number of repetitions and duration	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% 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]

1 **Table 53** Williams, 2015[264]

Reference	**Williams, 2015[264]**
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.
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

Reference	Williams, 2015[264]
	AF (other than just age) and/or symptoms suggestive of AF]

1
2 **Table 54** Brito, 2018[21]

Reference	Brito, 2018[21]
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
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	<u>Results for detection of AF/flutter</u> **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

Reference	Brito, 2018[21]
	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[56]

Reference	Doliwa, 2009[56]
Study type	Observational
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	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.

Reference	Doliwa, 2009[56]
of repetitions and duration	
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[159]

Reference	Nigolian, 2018[159]
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),	Beurer ME 80 device – a pocket sized (reconstructing 9 lead) ECG device that had electrodes mounted on each end. Can be used

Reference	Nigolian, 2018[159]
including number of repetitions and duration	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 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]

1
2 **Table 57** Winkler, 2011[266]

Reference	Winkler, 2011[266]
Study type	Observational
Recruitment	consecutive
Setting	Cardiology inpatients
Country	Germany

Reference	Winkler, 2011[266]
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.
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[263]

Reference	William, 2018[263]
Study type	Observational
Recruitment	Consecutive, but paired analysis in that each patient was medically CV or not

Reference	William, 2018[263]
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
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

Reference	William, 2018[263]
	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[248]

Reference	Velthuis, 2013[248]
Study type	Observational
Recruitment	consecutive
Setting	Stroke Unit
Country	Netherlands
Sample size	153
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	Yes, same devices used and the gold standard was simply the use of physicians rather than automated readings.

Reference	Velthuis, 2013[248]
simultaneous	
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[78]

Reference	Haberman, 2015[78]
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
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

Reference	Haberman, 2015[78]
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[177]

Reference	Poulsen, 2017[177]
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
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

Reference	Poulsen, 2017[177]
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[188]

Reference	Renier, 2012[188]
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.
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)**

Reference	Renier, 2012[188]
	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[194]

Reference	Rizos, 2010[194]
Study type	Observational
Recruitment	consecutive
Settings	Tertiary care stroke unit
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

Reference	Rizos, 2010[194]
	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.
Limitations	Risk of bias (QUADAS 2 – risk of bias): Very serious Indirectness (QUADAS 2 - applicability): none

Table 64. Vukajlovic, 2010[251]

Reference	Vukajlovic, 2010[251]
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;

Reference	Vukajlovic, 2010[251]
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
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[198]

Reference	Ross, 2018[198]
Study type	Observational
Recruitment	consecutive
Setting	Stroke Unit
Country	Germany

Reference	Ross, 2018[198]
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
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[135]

Reference	Lin, 2010[135]
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 of repetitions and duration	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. 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)

Reference	Lin, 2010[135]
	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[68]

Reference	Fallet, 2019[68]
Study type	Observational
Recruitment	Consecutive
Setting	Patients referred for catheter ablation
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**

Reference	Fallet, 2019[68]
	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[124]

Reference	Kvist, 2019[124]
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

Reference	Kvist, 2019[124]
	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.
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[5]

Reference	Antonicelli, 2012[5]
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
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

Reference	Antonicelli, 2012[5]
	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[132]

Reference	Lewis, 2011[132]
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

Reference	Lewis, 2011[132]
	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 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[176]

Reference	Poon, 2005[176]
Study type	Observational

Reference	Poon, 2005[176]
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.
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[73]

Reference	Greg, 2008[73]
Study type	Observational

Reference	Greg, 2008[73]
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
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[92]

Reference	Hobbs, 2005[92]
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
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

Reference	Hobbs, 2005[92]
	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)** 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

Reference	Hobbs, 2005[92]
	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[127]

Reference	Langley, 2012[127]
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

Reference	Langley, 2012[127]
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** 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[190]

Reference	Rhys, 2013[190]
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.

Reference	Rhys, 2013[190]
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
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[217]

Reference	Slocum, 1992[217]

Reference	Slocum, 1992[217]
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 of repetitions and duration	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 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[79]

Reference	Hald, 2017[79]
Study type	Observational

Reference	Hald, 2017[79]
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
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

Reference	Hald, 2017[79]
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[89]

Reference	Himmelreich, 2019[89]
Study type	Observational
Recruitment	Consecutive
Setting	Primary care
Country	Holland
Sample size	219
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

Reference	Himmelreich, 2019[89]
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
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[189]

Reference	Reverberi, 2019[189]

Reference	Reverberi, 2019[189]
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 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

1

1 **Table 80**. Sabar, 2019[202]

Reference	Sabar, 2019[202]
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
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%

Reference	Sabar, 2019[202]
	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[255]

Reference	Wasserlauf, 2019[255]
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

Reference	Wasserlauf, 2019[255]
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[46]

Reference	Cunha, 2019[46]
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.

Reference	Cunha, 2019[46]
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 Indirectness (QUADAS 2 - applicability): Serious indirectness

Table 83. Brown, 2019[22]

Reference	Brown, 2019[22]
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	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

Reference	Brown, 2019[22]
of repetitions and duration	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 Indirectness (QUADAS 2 - applicability): no indirectness

Appendix E: Coupled sensitivity and specificity forest plots and sROC curves

E.1 Coupled sensitivity and specificity forest plots

Note that Forest Plots are only available for results where raw data (i.e. TP, FP, FN, TN) were provided. It was not possible to include data in forest plots or pooled analyses where no raw data were available, even if the 95% CIs were provided. Hence some forest plots may not be present, or some forest plots may lack studies that are included in sections 1.5.5 and 1.5.6.

STRATUM 1: 12 lead ECG as gold standard

Mobile devices

Figure 2: AliveCor (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Cunha (auto), 2019	20	2	2	76	0.91 [0.71, 0.99]	0.97 [0.91, 1.00]
Desteghe (auto), 2017	12	6	10	237	0.55 [0.32, 0.76]	0.98 [0.95, 0.99]
Desteghe (expert), 2017	20	5	0	230	1.00 [0.83, 1.00]	0.98 [0.95, 0.99]
Haberman (expert), 2015	17	1	1	111	0.94 [0.73, 1.00]	0.99 [0.95, 1.00]
Himmelreich (auto), 2019	20	4	3	187	0.87 [0.66, 0.97]	0.98 [0.95, 0.99]
Himmelreich (expert), 2019	23	0	0	191	1.00 [0.85, 1.00]	1.00 [0.98, 1.00]
Lown (auto), 2018	72	4	10	332	0.88 [0.79, 0.94]	0.99 [0.97, 1.00]
William (auto), 2018	57	47	23	96	0.71 [0.60, 0.81]	0.67 [0.59, 0.75]
William (expert), 2018	75	19	5	124	0.94 [0.86, 0.98]	0.87 [0.80, 0.92]
Williams (clin), 2015	26	9	3	57	0.90 [0.73, 0.98]	0.86 [0.76, 0.94]
Williams (expert), 2015	27	16	2	50	0.93 [0.77, 0.99]	0.76 [0.64, 0.85]

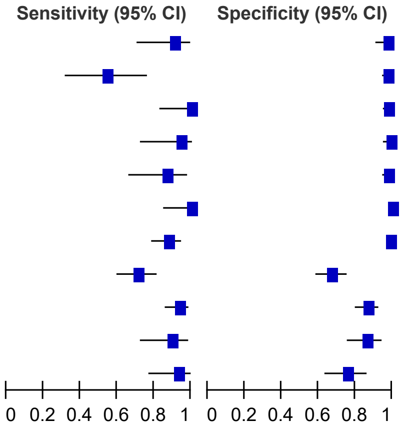


Figure 3: Kardia band (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bumgarner (auto), 2018	80	11	11	67	0.88 [0.79, 0.94]	0.86 [0.76, 0.93]
Bumgarner (expert), 2018	63	7	28	71	0.69 [0.59, 0.78]	0.91 [0.82, 0.96]

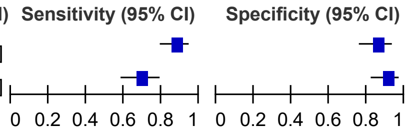


Figure 4: Beurer ME90 device – lead I (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brito (auto), 2018	16	40	2	65	0.89 [0.65, 0.99]	0.62 [0.52, 0.71]
Brito (expert), 2018	16	0	3	107	0.84 [0.60, 0.97]	1.00 [0.97, 1.00]

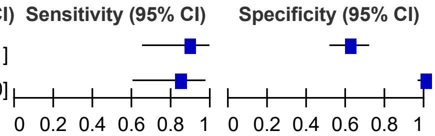


Figure 5: Beurer ME90 device – lead I and mv4 lead (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brito (auto), 2018	15	16	2	86	0.88 [0.64, 0.99]	0.84 [0.76, 0.91]

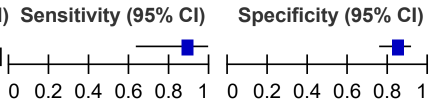


Figure 6: Beurer ME90 device – mv4 lead (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brito (auto), 2018	16	24	1	78	0.94 [0.71, 1.00]	0.76 [0.67, 0.84]
Brito (expert), 2018	16	0	3	107	0.84 [0.60, 0.97]	1.00 [0.97, 1.00]

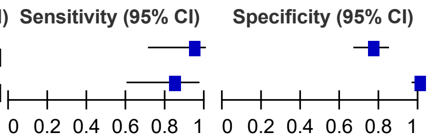


Figure 7: Beurer ME80 device (GS = 12 lead ECG) (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Nigolian (clin), 2018	12	4	4	32	0.75 [0.48, 0.93]	0.89 [0.74, 0.97]
Nigolian (expert), 2018	16	2	0	34	1.00 [0.79, 1.00]	0.94 [0.81, 0.99]

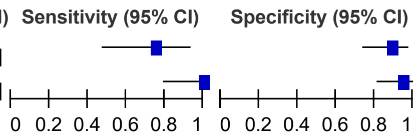


Figure 8: ECG check (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Haverkamp (auto), 2019	11	5	0	78	1.00 [0.72, 1.00]	0.94 [0.86, 0.98]

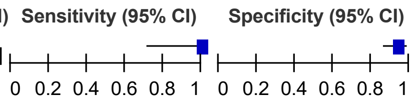


Figure 9: Merlin (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 10: MyDiagnostik (1 measure) (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Desteghe (auto), 2017	18	14	4	229	0.82 [0.60, 0.95]	0.94 [0.91, 0.97]
Desteghe (expert), 2017	17	11	3	226	0.85 [0.62, 0.97]	0.95 [0.92, 0.98]
Tieleman (auto), 2014	108	12	0	748	1.00 [0.97, 1.00]	0.98 [0.97, 0.99]

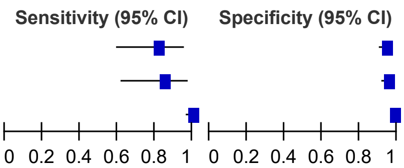


Figure 11: MyDiagnostik (3 measures, majority rule) (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Vaes (auto), 2014	90	6	6	79	0.94 [0.87, 0.98]	0.93 [0.85, 0.97]

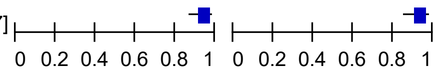


Figure 12: Omron Heartscan (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kearley (auto), 2014	78	219	1	701	0.99 [0.93, 1.00]	0.76 [0.73, 0.79]
Renier (auto), 2012	12	0	1	164	0.92 [0.64, 1.00]	1.00 [0.98, 1.00]
Renier (clin), 2012	9	9	4	155	0.69 [0.39, 0.91]	0.95 [0.90, 0.97]

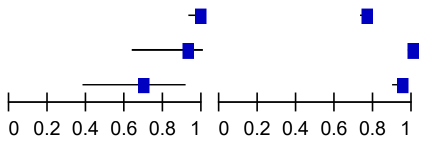


Figure 13: ECG bone (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 14: Zenecor ECG thumb (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Doliwa (expert), 2009	47	4	2	47	0.96 [0.86, 1.00]	0.92 [0.81, 0.98]

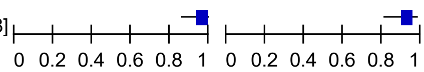


Figure 15: Polar H7 (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lown (auto), 2018	79	6	3	330	0.96 [0.90, 0.99]	0.98 [0.96, 0.99]

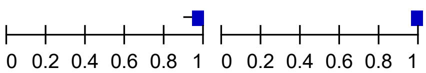
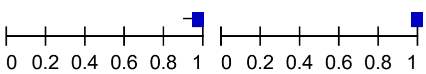


Figure 16: Firstbeat Bodyguard 2 (GS = 12 lead ECG)

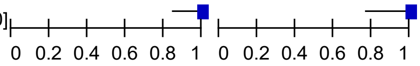
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lown (auto), 2018	79	5	3	331	0.96 [0.90, 0.99]	0.99 [0.97, 1.00]



1

Figure 17: Cardiobip (GS = 12 lead ECG)

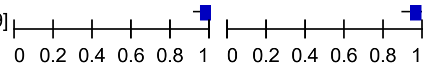
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Vukajlovic (expert), 2010	22	0	0	14	1.00 [0.85, 1.00]	1.00 [0.77, 1.00]



2

Figure 16: RITMIA (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Riverberi (auto), 2019	96	4	3	87	0.97 [0.91, 0.99]	0.96 [0.89, 0.99]



3

Figure 18: Mobile ECG device ER-2000s. Mode 1 (GS = 12 lead ECG)

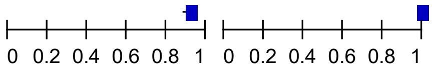
Forest plot not possible to generate as no raw data available

Figure 19: Mobile ECG device ER-2000s. Mode 2 (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 20: Huawei band 2 smartband (GS = 12 lead ECG)

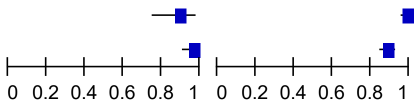
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Fan (auto), 2019	267	1	23	333	0.92 [0.88, 0.95]	1.00 [0.98, 1.00]



<u>BP devices</u>

Figure 21: Microlife BP3MQ1-2D (3 readings, majority rule) (GS = 12 lead ECG)

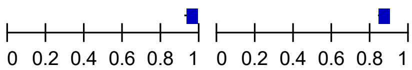
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Gandolfo (auto), 2015	34	2	4	167	0.89 [0.75, 0.97]	0.99 [0.96, 1.00]
Wiesel (auto), 2009	90	35	3	277	0.97 [0.91, 0.99]	0.89 [0.85, 0.92]



4

Figure 22: Microlife BP3MQ1-2D (1 reading) (GS = 12 lead ECG)

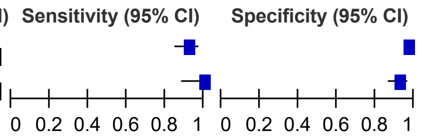
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Wiesel (auto), 2009	266	127	13	809	0.95 [0.92, 0.97]	0.86 [0.84, 0.89]



1

Figure 23: Microlife BPA 200 (3 readings, majority rule) (GS = 12 lead ECG)

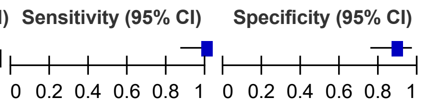
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Marazzi (auto), 2012	93	12	8	390	0.92 [0.85, 0.97]	0.97 [0.95, 0.98]
Wiesel (expert), 2014	30	12	0	141	1.00 [0.88, 1.00]	0.92 [0.87, 0.96]



2

Figure 24: Microlife BPA 100 Plus (3 readings, majority rule) (GS = 12 lead ECG)

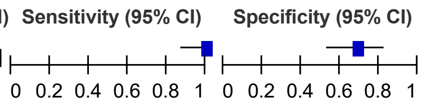
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Stergiou (auto), 2009	27	5	0	40	1.00 [0.87, 1.00]	0.89 [0.76, 0.96]



3

Figure 25: Microlife BPA 100 Plus (3 readings, majority rule) (GS = 12 lead ECG)

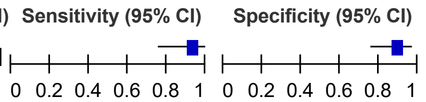
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Stergiou (auto), 2009	27	14	0	31	1.00 [0.87, 1.00]	0.69 [0.53, 0.82]



4

Figure 26: Microlife BPA 100 Plus (1st reading) (GS = 12 lead ECG)

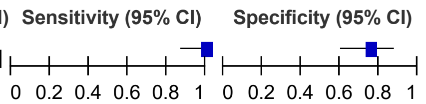
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Stergiou (auto), 2009	25	5	2	40	0.93 [0.76, 0.99]	0.89 [0.76, 0.96]



5

Figure 27: Microlife BPA 100 Plus (1st 2 readings) (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Stergiou (auto), 2009	27	11	0	34	1.00 [0.87, 1.00]	0.76 [0.60, 0.87]



6

Figure 28: Microlife Watch BP (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kearley (expert), 2014	75	95	4	825	0.95 [0.88, 0.99]	0.90 [0.88, 0.92]
Lown (auto), 2018	79	22	3	314	0.96 [0.90, 0.99]	0.93 [0.90, 0.96]

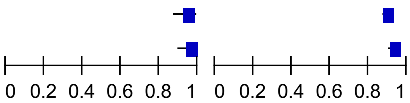


Figure 29: Heart Spectrum BP monitor algorithm 1 (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kao (expert), 2018	28	1	1	32	0.97 [0.82, 1.00]	0.97 [0.84, 1.00]

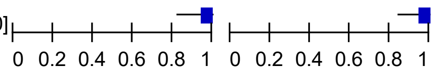


Figure 30: Heart Spectrum BP monitor algorithm 2 (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kao (expert), 2018	26	0	3	33	0.90 [0.73, 0.98]	1.00 [0.89, 1.00]

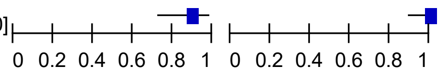


Figure 31: Heart Spectrum BP monitor algorithm 3 (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kao (expert), 2018	29	2	0	31	1.00 [0.88, 1.00]	0.94 [0.80, 0.99]

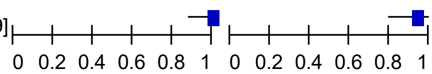


Figure 32: Omron 712 (2 readings) (GS = 12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Wiesel (auto), 2004	54	36	0	360	1.00 [0.93, 1.00]	0.91 [0.88, 0.94]

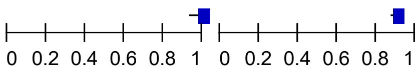
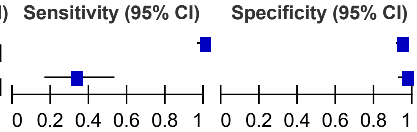


Figure 33: Omron M6 Comfort (1 reading) (GS = 12 lead ECG)

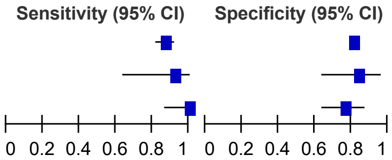
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Marazzi (auto), 2012	101	23	0	379	1.00 [0.96, 1.00]	0.94 [0.92, 0.96]
Wiesel (expert), 2014	10	5	20	148	0.33 [0.17, 0.53]	0.97 [0.93, 0.99]



PULSE PALPATION

Figure 34: Pulse palpation (GS=12 lead ECG)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Hobbs (clin), 2005	190	441	28	1919	0.87 (0.82, 0.91]	0.81 [0.80, 0.83]
Somerville (clin), 2000	12	4	1	21	0.92 [0.64, 1.00]	0.84 [0.64, 0.95]
Somerville (expert), 2000	26	14	0	46	1.00 [0.87, 1.00]	0.77 [0.64, 0.87]



PHOTOPLETHYSMOGRAPHY

Figure 35: iPhone 4s app - 2 minute pulse waveforms with PULSESMART app (using RMSSD, ShE and Poincare thresholds) from fingertip pulse recordings (1 reading)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
McManus (auto), 2016	95	6	3	85	0.97 [0.91, 0.99]	0.93 [0.86, 0.98]

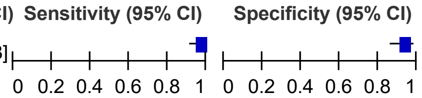


Figure 36: iPhone 4s app - 2 minute pulse waveforms with PULSESMART app (using RMSSD and ShE thresholds) from fingertip pulse recordings (1 reading) (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 37: iPhone 4s app - 2 minute pulse waveforms with PULSESMART app (using RMSSD threshold) from fingertip pulse recordings (1 reading) (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 38: iPhone 4s app - 2 minute pulse waveforms with PULSESMART app (using ShE threshold) from fingertip pulse recordings (1 reading) (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 39: Fingertip CardioRhythm 3 readings, majority rule

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rozen (auto), 2018	94	8	7	80	0.93 [0.86, 0.97]	0.91 [0.83, 0.96]

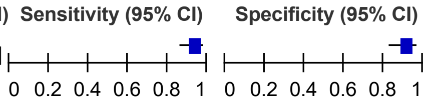


Figure 40: Fingertip CardioRhythm 3 readings, minority rule

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Yan (auto), 2018	71	10	4	132	0.95 [0.87, 0.99]	0.93 [0.87, 0.97]

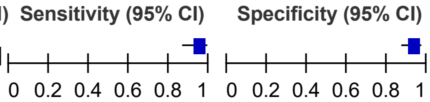


Figure 41: Facial CardioRhythm 3 readings, minority rule

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Yan (auto), 2018	71	6	4	136	0.95 [0.87, 0.99]	0.96 [0.91, 0.98]

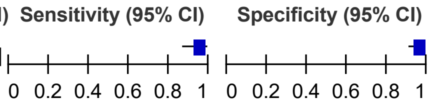


Figure 42: Fibricheck app 3 readings

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Proesmans (expert), 2019	87	4	13	119	0.87 [0.79, 0.93]	0.97 [0.92, 0.99]

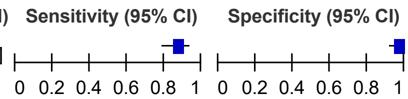


Figure 43: Huawei Honor 7A fingertip/LED device (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 44: Huawei Mate 9 fingertip/LED device (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 45: The screening technique involves a finger-probe instrument (as used in pulse oximetry) that utilises the principle of photoplethysmography. (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 46: Wrist-type photoplethysmographic (PPG) device. Using inter-beat interval (IBI) features (mean, SD, median, IQR, min, max and RMSSD. (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 47: Wrist-type photoplethysmographic (PPG) device. Using 'wave' features (Adaptive organisation Index, variance of the slope of the phase difference, permutation entropy, fractional spectral radius and spectral purity index) (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 48: Wrist-type photoplethysmographic (PPG) device. Using BOTH IBI and wave features (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

<u>3 LEAD TELE ECG</u>

Figure 49: CG 7100 3 lead Tele-ECG

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antonicelli (expert), 2012	1	0	0	106	1.00 [0.03, 1.00]	1.00 [0.97, 1.00]

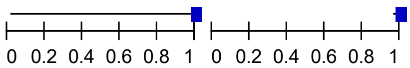


Figure 50: Handheld tele ECG device with dry electrodes that records 3 lead ECG. (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 51: Portable ECG monitor (PEM) – 3 lead ECG, 1 reading

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kristensen (clin), 2016	13	1	2	73	0.87 [0.60, 0.98]	0.99 [0.93, 1.00]

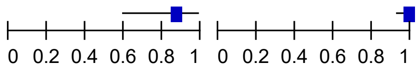


Figure 52: Medi-Trace 3 lead ECG algorithm 1, 1 reading

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lin (auto), 2010	1114	8	78	0	0.93 [0.92, 0.95]	0.00 [0.00, 0.37]

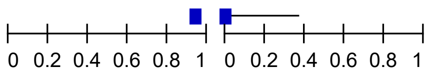


Figure 53: Medi-Trace 3 lead ECG algorithm 2, 1 reading

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lin (auto), 2010	1135	7	58	0	0.95 [0.94, 0.96]	0.00 [0.00, 0.41]

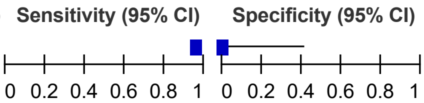


Figure 54: 6 lead ECG with prototype recorder placed on thorax/abdomen in sitting, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Caldwell (expert), 2012	76	0	2	79	0.97 [0.91, 1.00]	1.00 [0.95, 1.00]

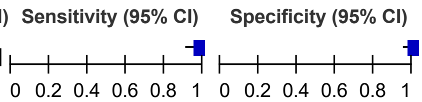


Figure 55: 6 lead ECG with prototype recorder placed on thorax/abdomen in supine, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Caldwell (expert), 2012	76	0	2	79	0.97 [0.91, 1.00]	1.00 [0.95, 1.00]

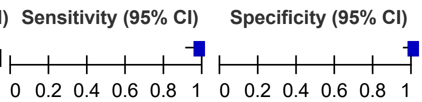


Figure 56: 6 lead ECG with prototype recorder placed on standard positions, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Caldwell (expert), 2012	76	0	2	79	0.97 [0.91, 1.00]	1.00 [0.95, 1.00]

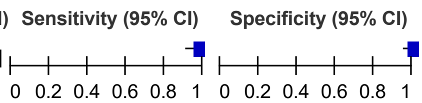
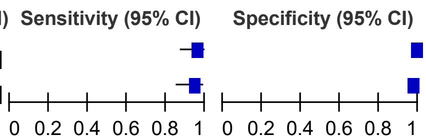


Figure 57: 6 lead ECG Rhythm pad

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Sabar (auto), 2019	63	7	3	566	0.95 [0.87, 0.99]	0.99 [0.97, 1.00]
Sabar (expert), 2019	62	18	4	555	0.94 [0.85, 0.98]	0.97 [0.95, 0.98]



<u>OTHER non- 12 LEAD ECG</u>

Figure 58: Limb lead ECG, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mant (Clin GP), 2007	104	156	22	1202	0.83 [0.75, 0.89]	0.89 [0.87, 0.90]
Mant (clin nurse), 2007	85	220	33	1107	0.72 [0.63, 0.80]	0.83 [0.81, 0.85]

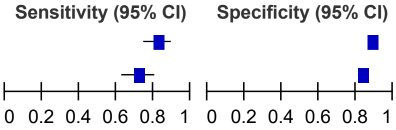


Figure 59: Chest lead ECG, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mant (Clin GP), 2007	112	180	20	1145	0.85 [0.78, 0.90]	0.86 [0.84, 0.88]
Mant (clin nurse), 2007	92	22	42	1066	0.69 [0.60, 0.76]	0.98 [0.97, 0.99]

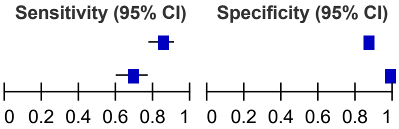


Figure 60: V1, V4 leads, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Greg (auto), 2008	96	17	13	1659	0.88 [0.80, 0.93]	0.99 [0.98, 0.99]

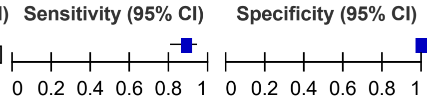


Figure 61: V2, V5 leads, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Greg (auto), 2008	92	17	17	1659	0.84 [0.76, 0.91]	0.99 [0.98, 0.99]

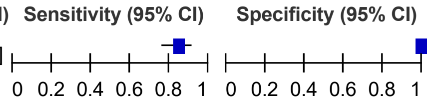
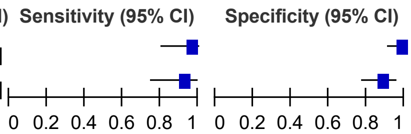


Figure 62: Bipolar lead ECG, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Somerville (clin), 2000	25	1	1	59	0.96 [0.80, 1.00]	0.98 [0.91, 1.00]
Somerville (expert), 2000	24	7	2	53	0.92 [0.75, 0.99]	0.88 [0.77, 0.95]



12 LEAD ECG (non expert)

Figure 63: 12 lead ECG interpreted by non-expert interpreter, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Hald (clin nurse), 2017	10	3	0	74	1.00 [0.69, 1.00]	0.96 [0.89, 0.99]
Kvist (clin nurse), 2019	66	0	2	1270	0.97 [0.90, 1.00]	1.00 [1.00, 1.00]
Mant (Clin GP), 2007	79	114	22	1241	0.78 [0.69, 0.86]	0.92 [0.90, 0.93]
Mant (clin nurse), 2007	74	198	22	1132	0.77 [0.67, 0.85]	0.85 [0.83, 0.87]
Rhys (clin GP), 2013	2	0	0	30	1.00 [0.16, 1.00]	1.00 [0.88, 1.00]
Somerville (clin GP), 2000	26	1	0	59	1.00 [0.87, 1.00]	0.98 [0.91, 1.00]
Somerville (clin nurse), 2000	13	6	0	19	1.00 [0.75, 1.00]	0.76 [0.55, 0.91]

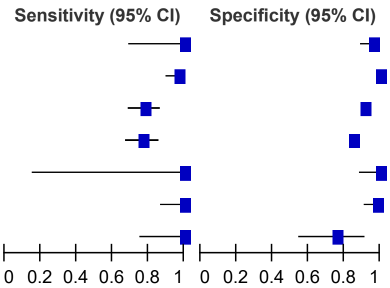


Figure 64: 12 lead ECG interpreted by non-expert interpreter combined with algorithm interpretation, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mant (clin + auto), 2007	91	121	8	1234	0.92 [0.85, 0.96]	0.91 [0.89, 0.93]

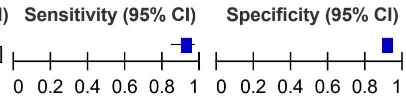


Figure 65: 12 lead ECG detection algorithm based on a co-efficient of variation of the beat intervals (CV). Threshold set at 0.12 (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 66: 12 lead ECG detection algorithm based on the co-efficient of sample entropy (COSEn). Threshold set at -1.19 (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 67: 12 lead ECG detection algorithm 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 (GS = 12 lead ECG)

Forest plot not possible to generate as no raw data available

Figure 68: 12 lead ECG algorithm interpreted by Cardioview, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rhys (auto), 2013	2	0	0	30	1.00 [0.16, 1.00]	1.00 [0.88, 1.00]

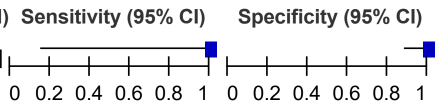


Figure 69: 12 lead ECG algorithm interpreted by MUSE software, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Poon (auto), 2005	227	41	23	3663	0.91 [0.87, 0.94]	0.99 [0.99, 0.99]

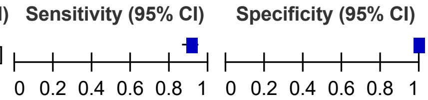


Figure 70: 12 lead ECG algorithm interpreted by Mant algorithm, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mant (auto), 2007	179	21	36	2320	0.83 [0.78, 0.88]	0.99 [0.99, 0.99]

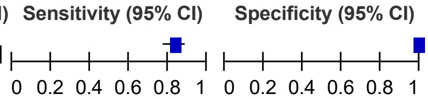


Figure 71: 12 lead ECG algorithm interpreted by Slocum algorithm, 1 measure

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Slocum (auto), 1992	28	5	13	36	0.68 [0.52, 0.82]	0.88 [0.74, 0.96]

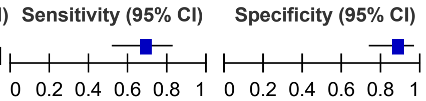
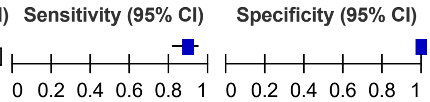


Figure 72: Computer interpretation of full 12 lead ECG V1-V6

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Greg (auto), 2008	97	17	12	1659	0.89 [0.82, 0.94]	0.99 [0.98, 0.99]



STRATUM 2: >24 hour ambulatory monitoring [such as Holter] as gold standard

BP MONITORS

Figure 73: 24 hour ambulatory Microlife Afib Watch BP

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kollias (auto), 2018	1013	78	78	4609	0.93 [0.91, 0.94]	0.98 [0.98, 0.99]

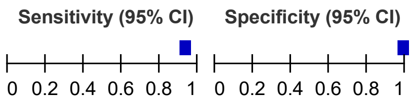
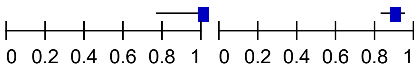


Figure 74: AF-BP monitor device (daily use for 30 days)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Wiesel (auto), 2013	14	13	0	112	1.00 [0.77, 1.00]	0.90 [0.83, 0.94]



<u>HOLTER <7 DAYS</u>

Figure 75: Holter 1 day

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mulder (auto), 2012	11	0	10	75	0.52 [0.30, 0.74]	1.00 [0.95, 1.00]

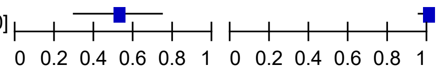


Figure 76: Holter 2 day

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mulder (auto), 2012	14	0	7	75	0.67 [0.43, 0.85]	1.00 [0.95, 1.00]

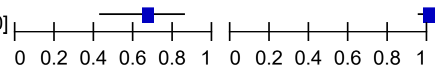


Figure 77: Holter 3 day

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mulder (auto), 2012	17	0	4	75	0.81 [0.58, 0.95]	1.00 [0.95, 1.00]

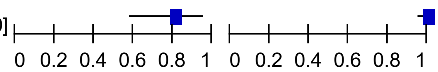


Figure 78: Holter 4 day

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mulder (auto), 2012	18	0	3	75	0.86 [0.64, 0.97]	1.00 [0.95, 1.00]

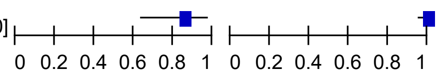


Figure 79: Holter 5 day

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mulder (auto), 2012	19	0	2	75	0.90 [0.70, 0.99]	1.00 [0.95, 1.00]

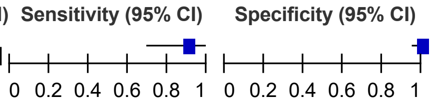
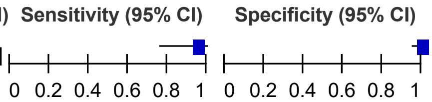


Figure 80: Holter 6 day

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mulder (auto), 2012	20	0	1	75	0.95 [0.76, 1.00]	1.00 [0.95, 1.00]



OTHER LONGER TERM DEVICES

Figure 81: R test evolution 3 triggered ECG (48 hrs)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Roten (expert), 2012	37	0	5	58	0.88 [0.74, 0.96]	1.00 [0.94, 1.00]

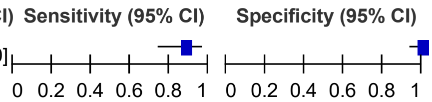


Figure 82: R test evolution 3 triggered ECG (24 hrs)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Sejr (auto), 2019	35	179	3	1195	0.92 [0.79, 0.98]	0.87 [0.85, 0.89]
sejr (expert), 2019	32	27	6	1347	0.84 [0.69, 0.94]	0.98 [0.97, 0.99]

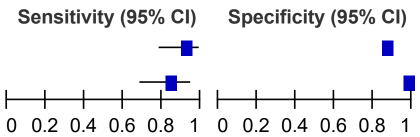


Figure 83: Vitaphone 3100 BT external loop recorder (24 hrs)

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Muller (auto), 2009	24	12	0	12	1.00 [0.86, 1.00]	0.50 [0.29, 0.71]
Velthius (auto), 2013	56	1134	3	1162	0.95 [0.86, 0.99]	0.51 [0.49, 0.53]

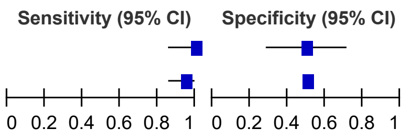


Figure 84: SRAclinic, Apoplex Medical Technologies. Stroke Risk Analysis (SRA) – software analysis of every hourly ECG snippet of continuous (non 12 lead) ECG monitoring, and report sent daily to stroke unit.(automated) threshold of 0-1=SR and 2 or more =AF

Forest plot not possible to generate as no raw data available

Figure 85: 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.(automated) threshold of 0-2=SR and 3 or more =AF

Forest plot not possible to generate as no raw data available

Figure 86: 48 hr ECG without AFRS

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Arevalo-Manso (auto/clin), 2016	1	0	12	46	0.08 [0.00, 0.36]	1.00 [0.92, 1.00]

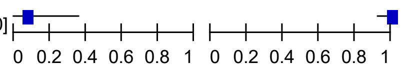


Figure 87: 48 hrs AGC with AFRS

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Arevalo-Manso (auto/clin), 2016	4	0	3	10	0.57 [0.18, 0.90]	1.00 [0.69, 1.00]

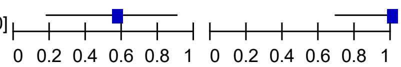


Figure 88: 12 bit resolution ECG 1-2 hrs

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rizos (auto), 2010	21	40	8	67	0.72 [0.53, 0.87]	0.63 [0.53, 0.72]

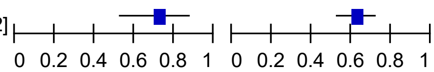


Figure 89: 6 Channel Holter

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rizos (auto), 2010	3	0	10	107	0.23 [0.05, 0.54]	1.00 [0.97, 1.00]

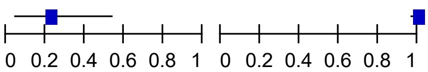


Figure 90: Zenecor thumb ECG twice daily for 30days

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Poulsen (expert), 2017	10	10	7	68	0.59 [0.33, 0.82]	0.87 [0.78, 0.94]

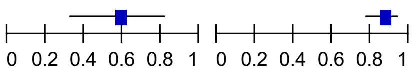
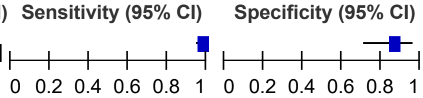


Figure 91: Kardia-Band

Forest plot not possible to generate as no raw data available

Figure 92: Cardiomatrix with telemetry for median 46 hours

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brown (auto), 2019	218	5	5	32	0.98 [0.95, 0.99]	0.86 [0.71, 0.95]



E.2 ROC curves

Figure 93: Meta-analysis for AliveCor handheld lead I ECG

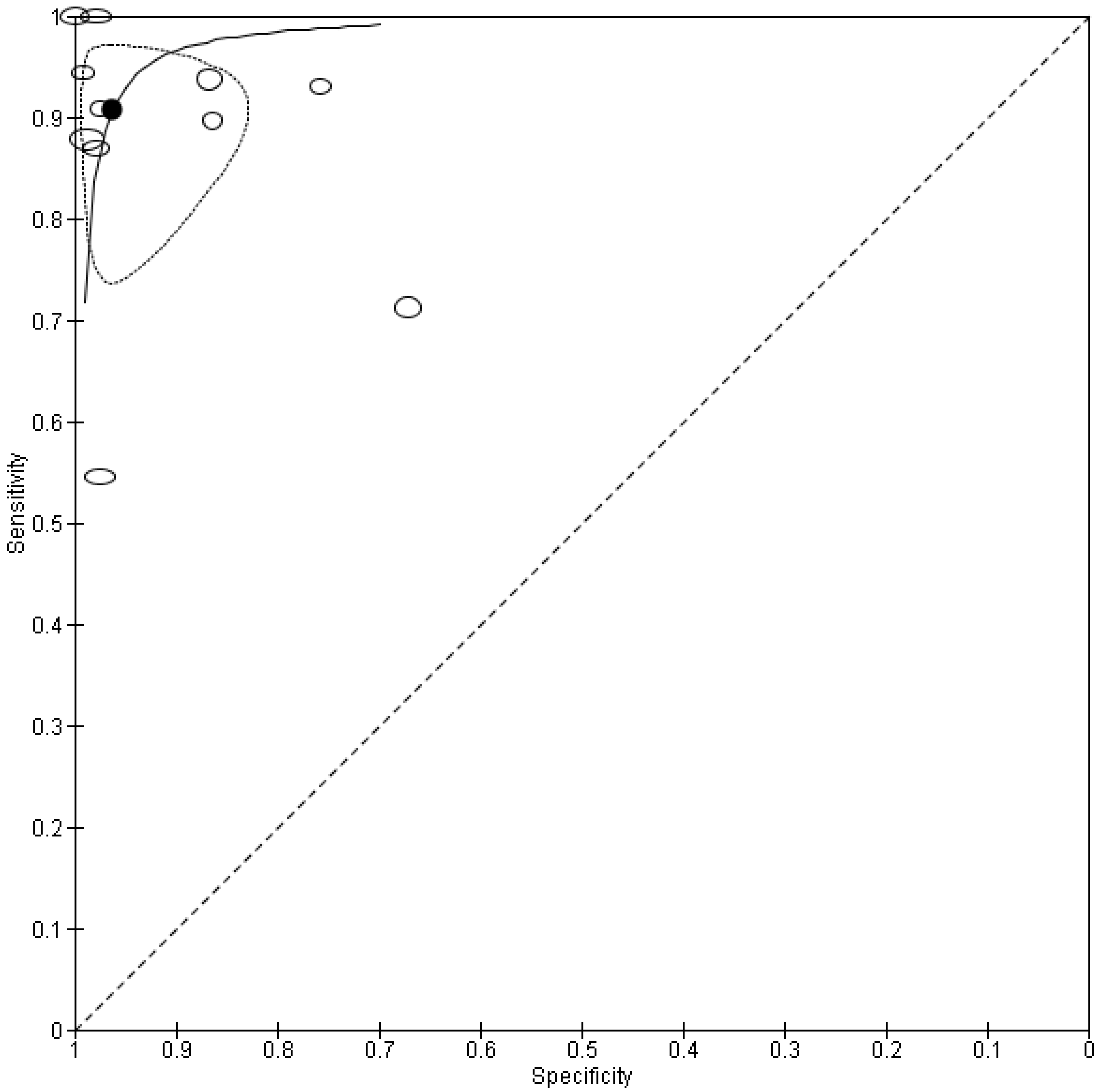
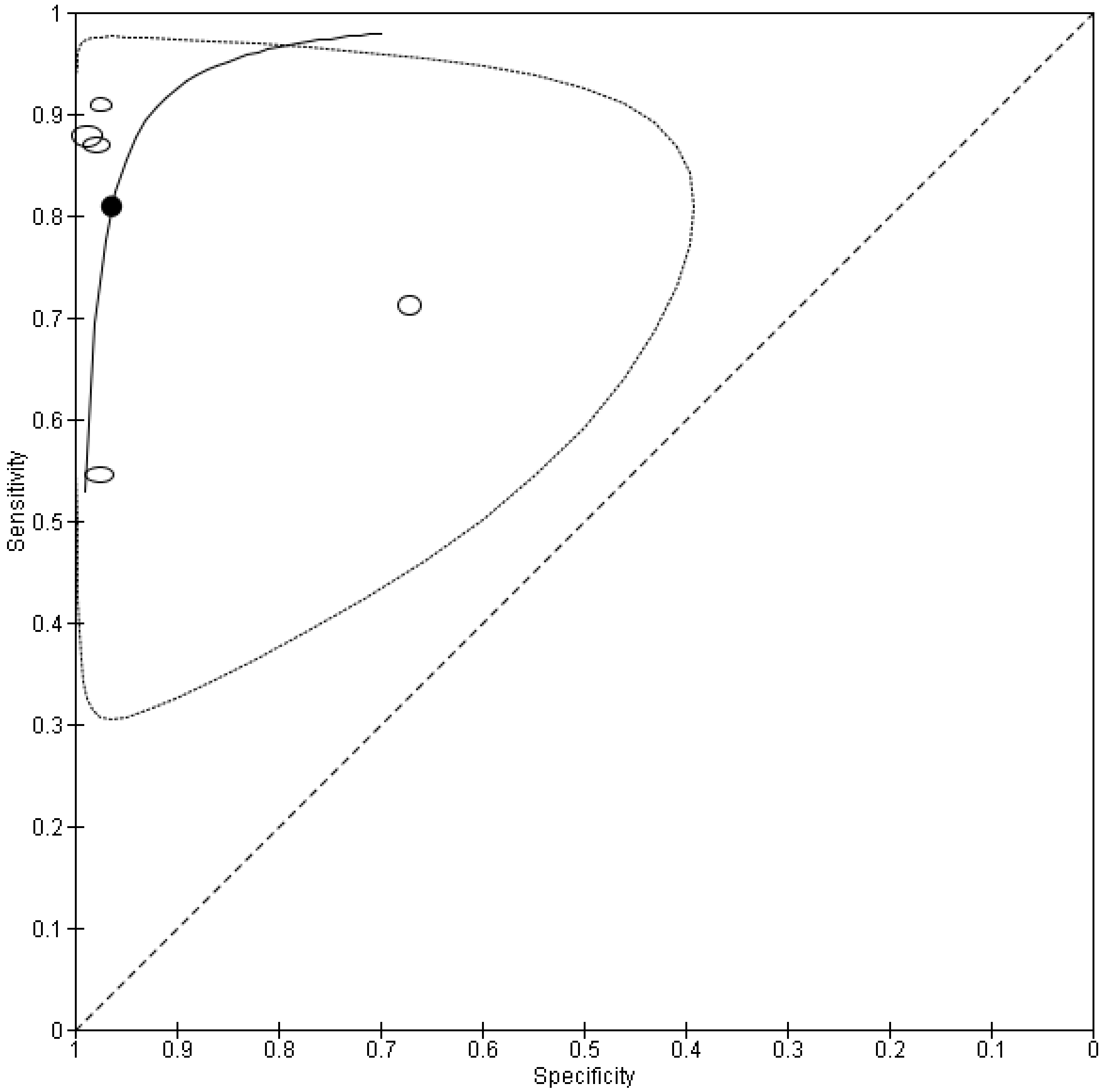
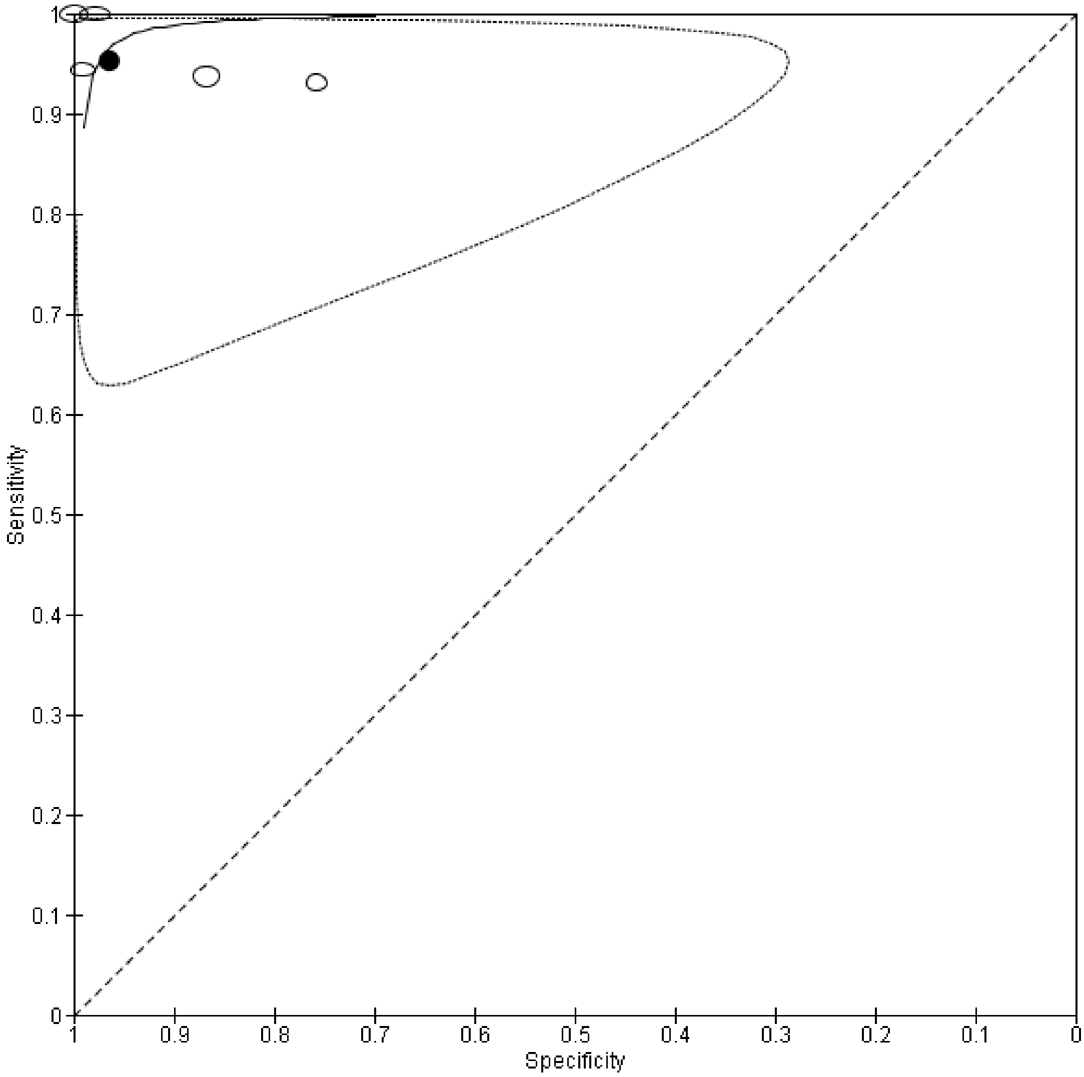


Figure 94: Meta-analysis for AliveCor handheld lead I ECG (automated subgroup)



1 **Figure 95: Meta-analysis for AliveCor handheld lead I ECG (expert subgroup)**



2

Figure 96: Meta-analysis for MyDiagnostik (no sub-grouping as<3 in any sub-group)

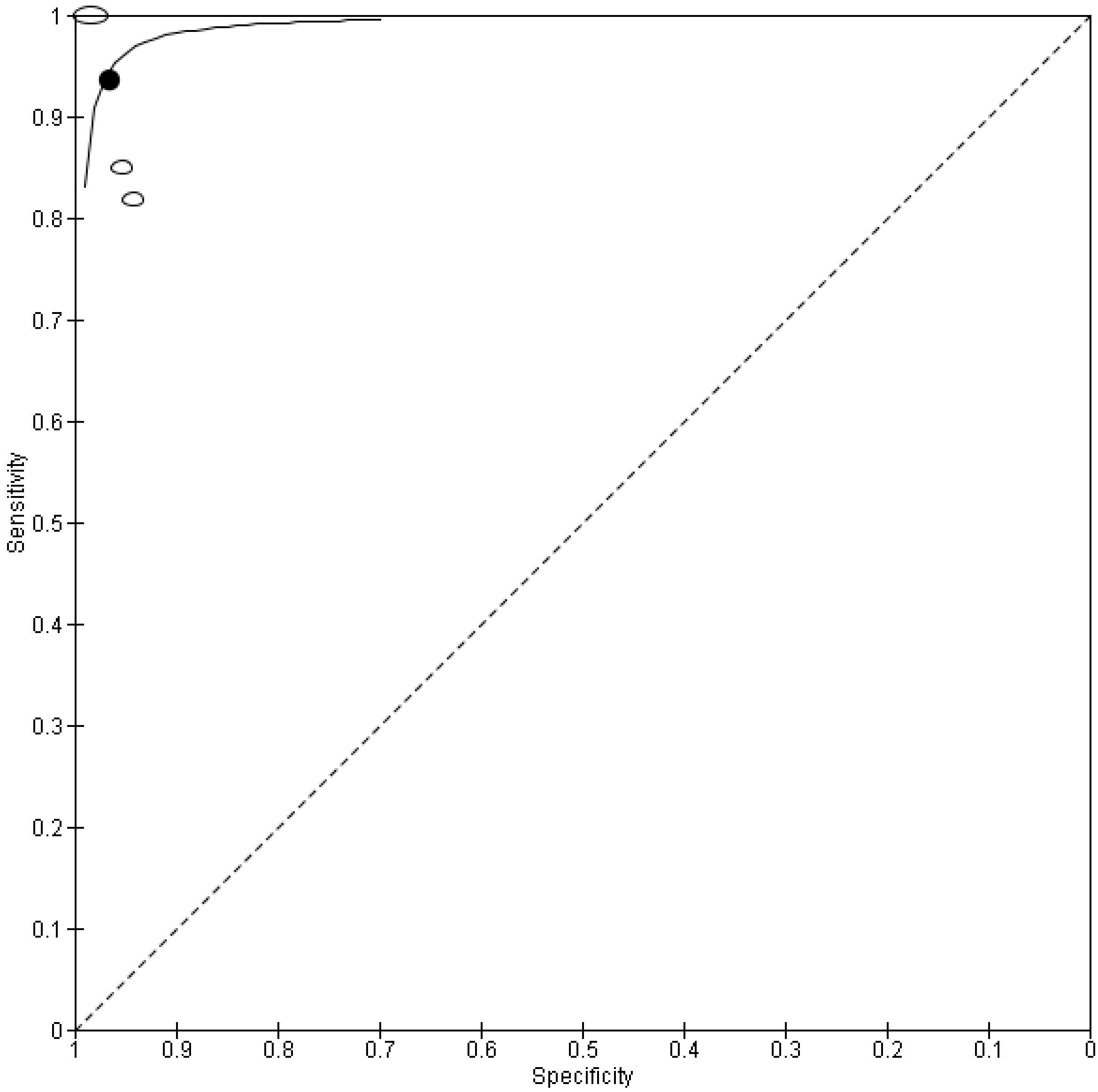


Figure 97: Meta-analysis for Omron Heartscan (no sub-grouping as<3 in any sub-group)

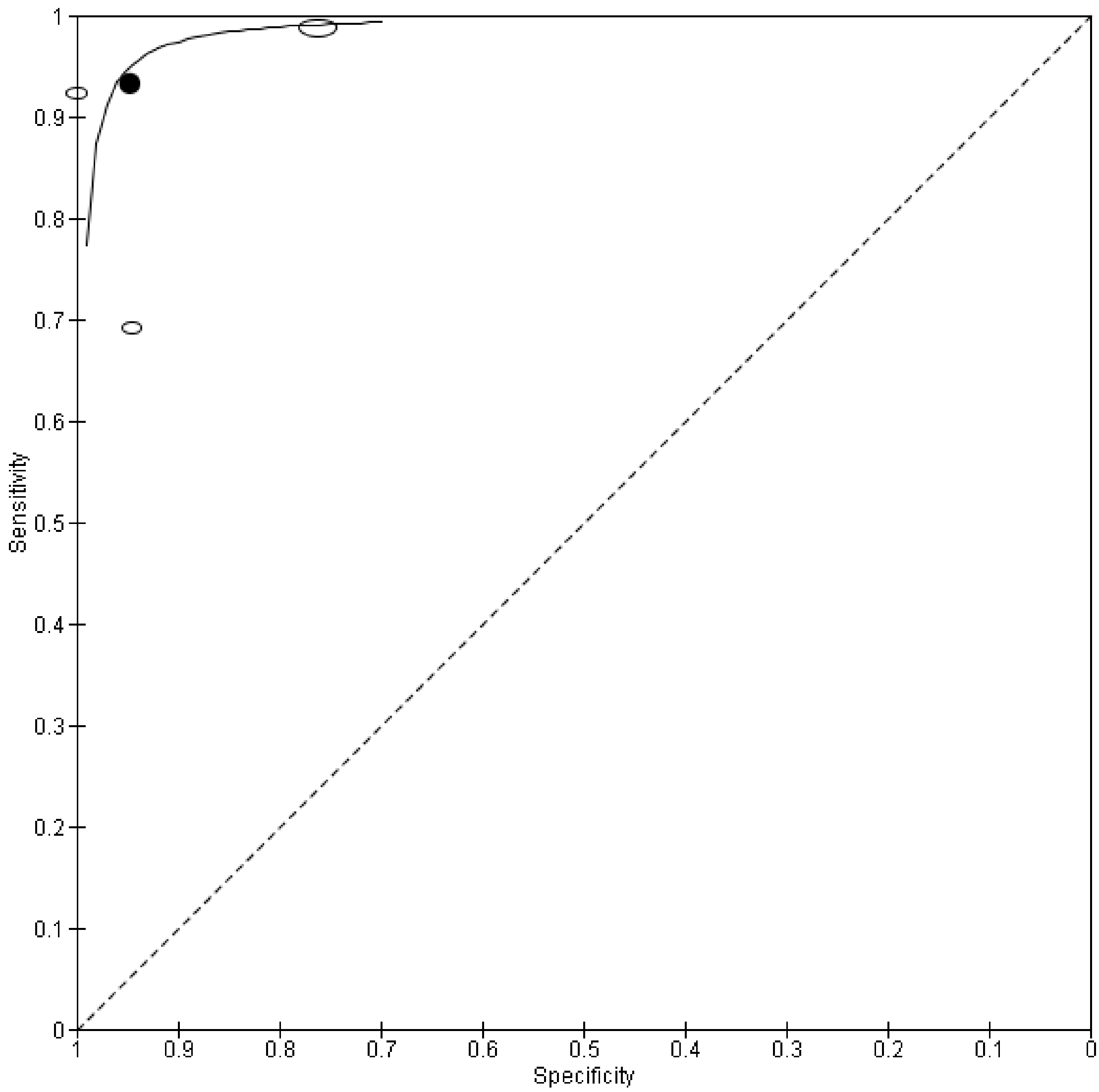


Figure 98: Meta-analysis for pulse palpation (no sub-grouping as<3 in any sub-
`

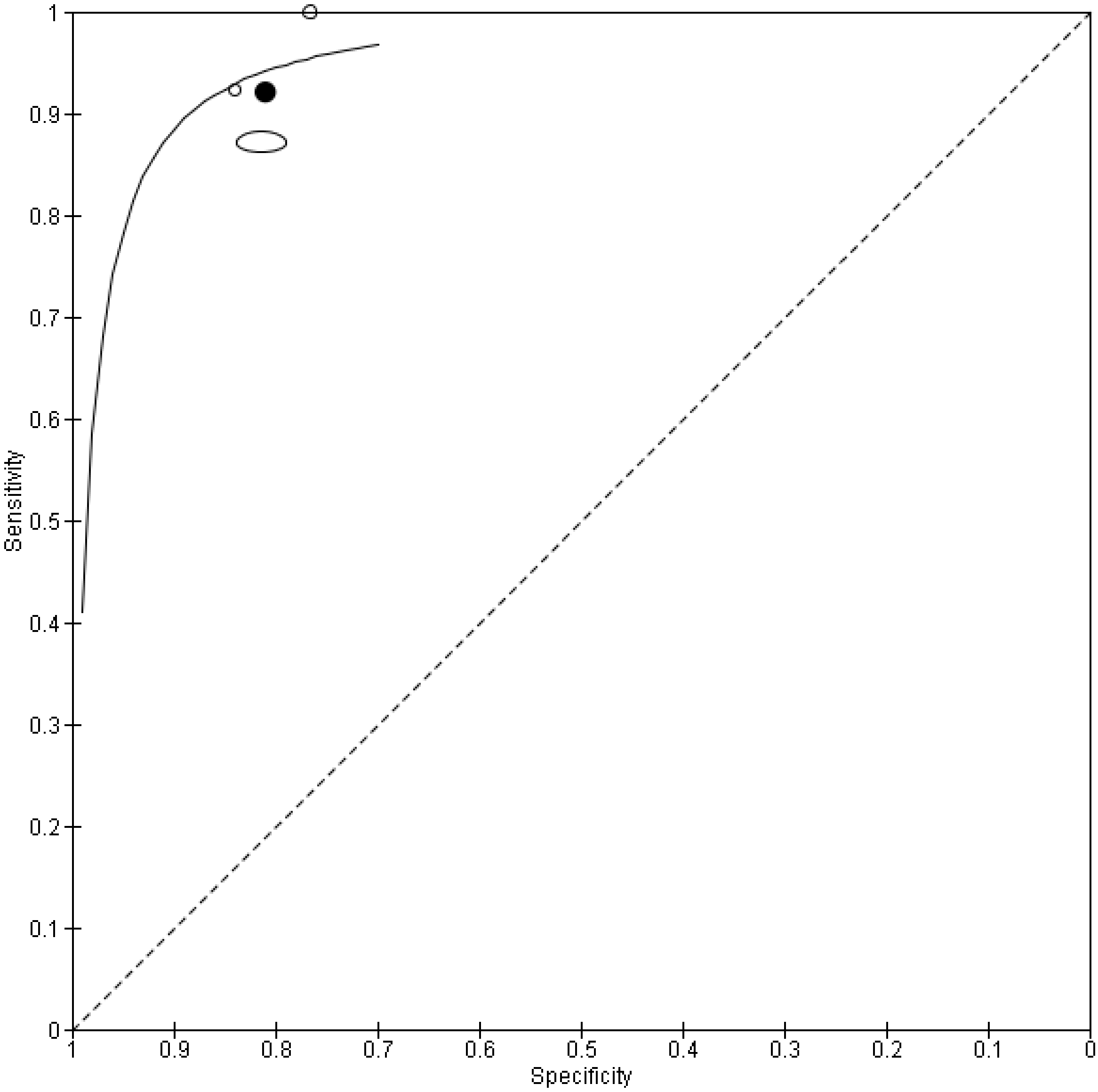
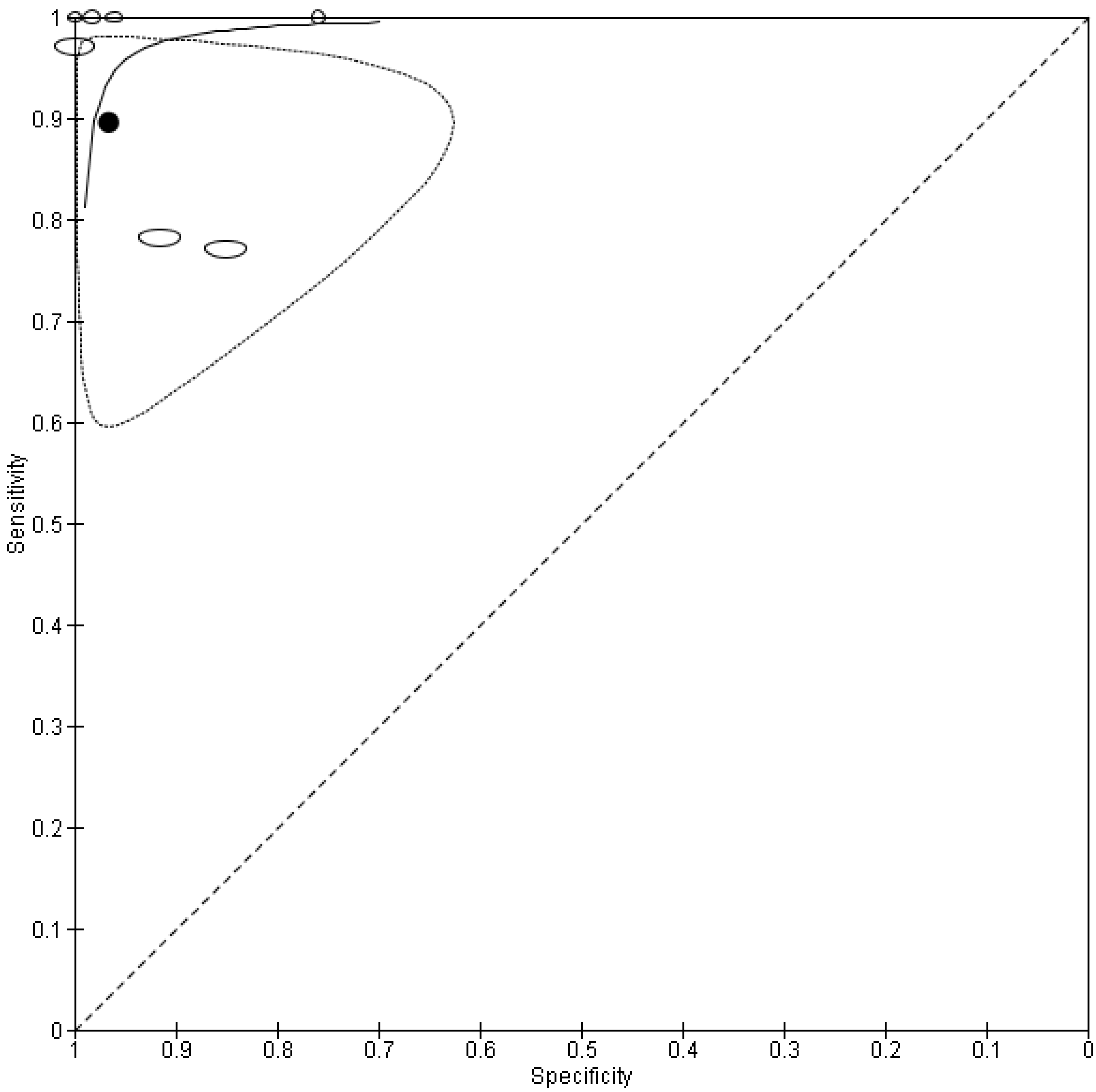
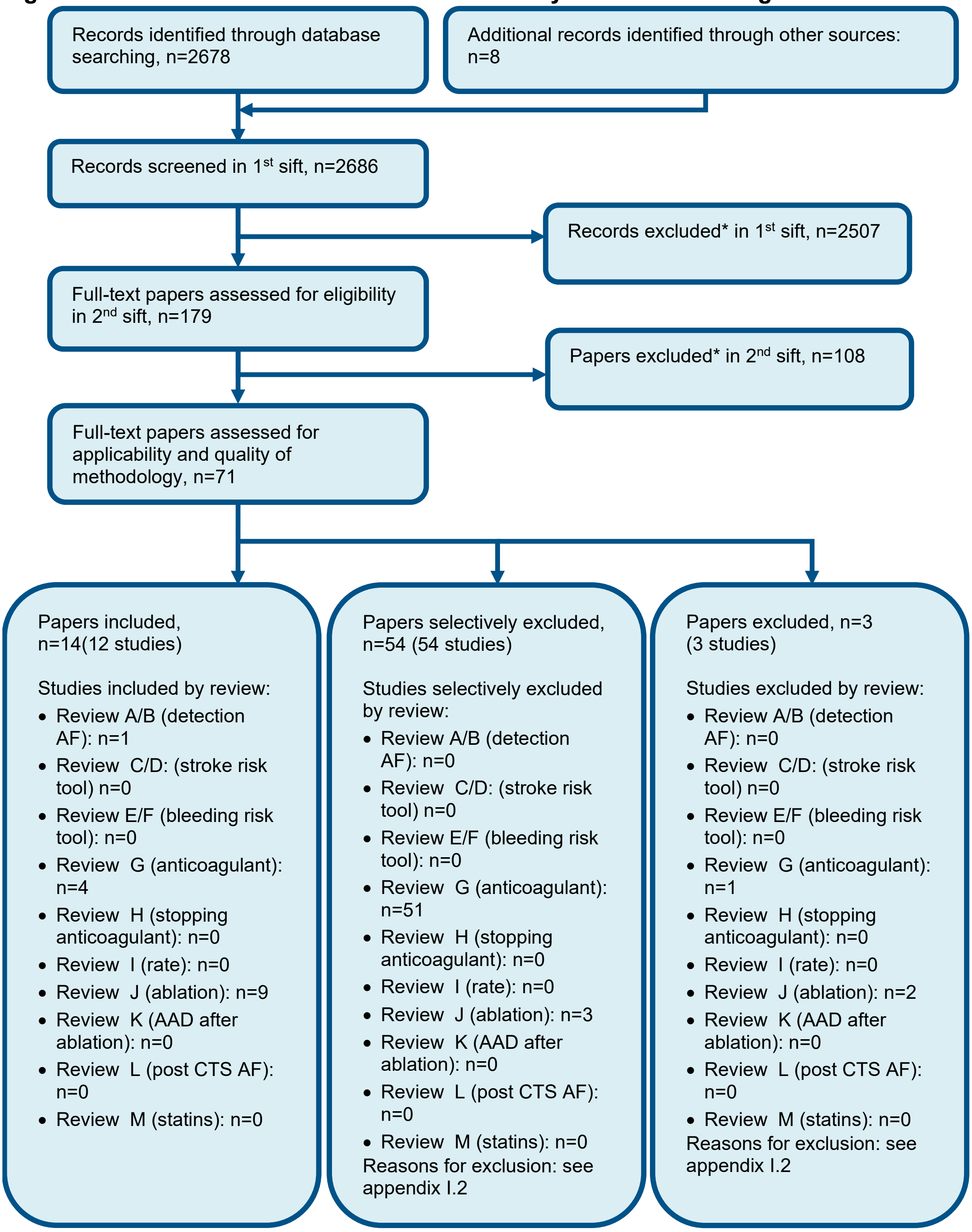


Figure 99: Meta-analysis for 12 lead ECG by non-expert clinicians



Appendix F: Health economic evidence selection

Figure 100: Flow chart of health economic study selection for the guideline



* Non-relevant population, intervention, comparison, design or setting; non-English language

Appendix G: Health economic evidence tables

Please see evidence review A.

Appendix H: QUADAS2 risk of bias assessment

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
Antonicelli, 2012[5]	Random	yes	Yes	Within 1 day	None reported	Serious risk of bias
Arevalo-Manso, 2016[6]	Consecutive	Unclear	unclear	unclear	None reported	Very serious risk of bias
Brito, 2018[21]	consecutive	yes	unclear	Not simultaneous	None reported	Very serious risk of bias
Bumgarner, 2018[24]	Case-control	Yes	Yes	Not simultaneous, but a very short interval between	169 simultaneous 12-lead ECG and KB recordings obtained from study participants, and of these 57 KB recordings were determined as unclassified by the KB algorithm. Of the 57 unclassified KB tracings, 16(28%) were due to baseline artifact and low amplitude of the recording, 12 (21%) were due to a recording of <30 s in duration, 6(10%) were due to a heartrate of <50 beats/min, 5 (9%) were due to a heart rate of >100 beats/min, and the	Serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
					remaining 18 (32%) were unclassified due to an unclear reason. However these represent a drawback of the FB and so these should have been designated as negative findings rather than excluded. The authors presented the calculated accuracies using only the interpretable KB values. However they did present the raw data including the missing/unclassified data, which has been used by the systematic reviewer to calculate more pragmatic accuracy values (with designation of missing data as a negative result).	
Caldwell, 2012[27]	Case-control	partial	partial	Not simultaneous but same session	None reported	Very serious risk of bias
Desteghe, 2017[53]	Consecutive	NA for automated	Yes, blinded	GS done	Yes – 24/344 lost	Serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
		measurements. For manual interpretation readings unclear		immediately before IT	from analysis because they could not hold device properly. Had they been included a less accurate result may have ensued. But <10% so not a serious risk of bias	for automatic readings and very serious for manual readings
Doliwa, 2009[56]	consecutive	yes	Yes	Not simultaneous	None reported	Serious risk of bias
Fallet, 2019[68]	consecutive	unclear	unclear	Appears to be simultaneous: 'temporally aligned'	None reported	Very serious risk of bias
Fan, 2019[69]	Unclear but appears to be random	NA as algorithm is automatic	Unclear (only states blinded from baseline characteristics)	simultaneous	4/112 as ECG data unclear – unlikely to pose a significant risk of bias	Serious risk of bias
Gandolfo 2015[71]	Unselected consecutive patients admitted with stroke	NA as automated	Yes, cardiologist no knowledge of index test results	<48 hours but usually less than 1 day	None reported	Serious risk of bias
Greg, 2008[73]	Random	NA - automated	No	Simultaneous	None	Very serious risk of bias
Haberman, 2015[78]	consecutive	unclear	unclear	Not simultaneous	None reported	Very serious risk of bias
Hald, 2017[79]	Random	The gold standard interpretations were performed 'post-study'	Yes, blinded	Simultaneous	No loss of data	No serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
		so likely that index test interpretations were made prior to any gold standard interpretations. Thus effectively blinded.				
Haverkamp, 2019[84]	Consecutive	Not applicable as automated	Unclear. Appears possible it was unblinded as the 'reports' of previous 12 lead ECG seems to imply interpretation had already been made. Reports that data were analyses independently by 2 observers, but unclear if this relates to index tests and GS tests.	unclear	Not reported	Very serious risk of bias
Hobbs, 2005[92]	Random	Blinding not stated. Anonymised traces but does not necessarily imply blinding. For automatic measures, NA.	Blinding not stated. Anonymised traces but does not necessarily imply blinding	Simultaneous	Varied between index tests but all involved high attrition at >10%. Possible that the GPs and nurses not returning interpretations were the less accurate participants	Very serious risk of bias
Kaleschke, 2009[104]	Consecutive	Yes – 'all ECG analyses were blinded to the analysis result of the other ECG	Yes- blinded	12 lead ECG 'immediately' before index test. Estimated	3/508 lost due to technical quality issues (n=2) and insufficient clinical	Serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
		modality and to clinical information of the patient'.		to be a 5-10 second delay	data (n=1). Not a serious risk of attrition bias.	
Kao, 2018[110]	Unclear – possibly case-control	No blinding	No blinding	simultaneous	1 person lost from analysis but due to ineligibility. Therefore no risk of bias	Very serious risk of bias
Kearley, 2014[113]	Consecutive	NA as automated for Watch BP and Omron. For cardiologist analysed data for Omron and merlin blinded.	Yes, the cardiologists were blinded to results of index tests and clinical data.	Gold standard done at the end of the same day after the index tests but exact timing unclear	Watch BP: 1 lost; Omron auto analysis: 2 lost Omron ECG trace: 4 lost;Merlin: 20 lost; All <10% so not regarded as significant	Serious risk of bias
Kollias, 2018[116]	Consecutive	NA as fully automated	Unclear (not reported)	simultaneous	None	Serious risk of bias
Koltowski, 2019[117]	consecutive	Unclear – no report of blinding.	Carried out first in all cases but this does not ensure blinding as interpretation could have occurred after index tests. Therefore unclear	Short but not simultaneous	1 lost because of tremors due to Parkinson's disease – no serious risk of attrition bias	Very serious risk of bias
Kristensen, 2016[122]	Case control	Yes	Yes	Simultaneous	4 lost due to poor ECG quality. But <10%	No serious risk of bias
Kvist, 2019[124]	consecutive	Unclear – although index tests done first possible that interpretation could have occurred after gold standard tests	Yes	1 hour delay maximum	2 lost due to leaving laboratory before 12 lead ECG completed. <0.2% and so would not affect results	Very serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results completed	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
Langley, 2012[127]	Random	Yes – algorithm used	Yes – gold standard assignments of status made in past, long before study inception (and index test evaluation)	simultaneous	None reported - based on pre-existing database	No serious risk of bias
Lewis, 2011[132]	Random	NA as automated	Yes	Immediately afterwards	None reported	Serious risk of bias
Lin, 2010[135]	Case-control. AF and non-AF (defined by gold standard) tested under different conditions and so results cannot be superimposed.	NA as automated	unclear	No - unclear	None reported	Very serious risk of bias
Lown, 2018[138]	Described as case-control and likely to be as prevalence of AF in study is 57%, way above the expected value	yes	Yes	Not simultaneous but in same session	Zero	Very serious risk of bias
Mant, 2007[142]	random	Yes, blinded	Yes, blinded	Simultaneous	3 ECGs lost which is very small compared to total number.	No serious risk of bias
Marazzi, 2012[143]	consecutive	NA – fully automated	Cardiologists blinded to index test results	Simultaneous	52 missing. 29 excluded because of willingness to be	Serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
					studied. Other 23 unclear.	
McManus, 2013[146]	Case-control (paired)	Unclear	Unclear	unclear	None reported	Very serious risk of bias
McManus, 2016[145]	People before and after a cardioversion – thus very much a case-control situation	NA as automated	Unclear	simultaneous	None reported	Serious risk of bias
Mulder, 2012[153]	consecutive	NA as automated	unclear	simultaneous	Not reported	Serious risk of bias
Muller 2009[154]	24 with AF and 24 without – thus appears to be case control but described as consecutive	automated	Unclear	Simultaneous	None reported	Very serious risk of bias
Nigolian, 2018[159]	consecutive	yes	Yes	Not simultaneous	None reported	Serious risk of bias
Park, 2015[167]	Consecutive	Blinded to identity and history of patient but not reported if blinded to GS results	Blinded to identity and history of patient but not reported if blinded to IT results	simultaneous	None reported	Very serious risk of bias
Poon, 2005[176]	Random	NA - automated	No	Simultaneous	None	Serious risk of bias
Poulsen, 2017[177]	consecutive	unclear	unclear	Simultaneous (concurrent)	5 lost – 2 withdrew consent before initiation and 3 had diagnosis changed. So not a threat to	Very serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
					validity.	
Proesmans, 2019[178]	Case-control	NA as automated for PPG device; unclear for 1 lead device	Yes	Probably not	Some data lost due to poor quality, but sensitivity analyses done	Serious risk of bias
Renier, 2012[188]	consecutive	yes	Yes	Not simultaneous	67 lost – 40 because of no 12 lead ECG, 12 because heartscan could not be put on chest, 15 refused consent, 3 because of problems with right index position and 7 below 18 years. Only 15 of these relate to outcome, which is <10%.	Serious risk of bias
Rhys, 2013[190]	Random	Yes - done prior to any gold standard interpretation	Not blinded to algorithm result but blinded to GPST2's interpretation	simultaneous	7 excluded – 5 because cardiologists unable to read faxed transmission and 2 because of poor quality ECGs..>10% so potential bias	Very serious risk of bias
Rizos, 2010[194]	consecutive	Unclear, but for automatic measures NA.	Unclear	concurrent	none	Very serious risk of bias for manual measures and serious for automatic measures
Ross, 2018[198]	consecutive	NA as automated	unclear	concurrrent	Significant losses of 21%. 32 due to etiology being	Very serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
					pathologic findings, 161 due to incomplete data.	
Roten, 2012[199]	consecutive	unclear	unclear	simultaneous	None reported (any transient loss of data included in accuracy analysis)	Very serious risk of bias
Rozen, 2018[200]	Case-control	NA as automated	yes	Not clear, but probably not simultaneous	Minor losses (n=2) pre-CV due to inappropriate inclusion (n=1), technical issues with CRMA (n=1). 5 missing from post-CV measurements because of normal sinus rhythm at baseline (n=1), contraindication to procedure (n=3), drop-out (n=1). Unlikely to have affected overall results as <10%	Serious risk of bias
Sejr, 2019[213]	Consecutive	yes	Yes	yes	excluded 95 patients, in whom ELR recording was not started correctly, but this is <<10% so not a concern	No serious risk of bias
Slocum, 1992[217]	Case control	Yes - automated	Unclear	simultaneous	No loss of data	Serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
Somerville, 2000[220]	Case control	Unclear	Unclear	Unclear but in same session	86 attended out of 154 invited. However if we can assume 86 were enrolled data loss is zero.	Very serious risk of bias
Stergiou, 2009[223]	Appears to be case/control	NA as automated	Not reported	Simultaneous	None	Very serious risk of bias
Tieleman, 2014[238]	Random	NA as fully automated	Yes	Short but not simultaneous	None reported	Serious risk of bias
Vaes, 2014[245]	Selective case/control	NA as fully automated	Yes, blinded	Short but not simultaneous	None reported	Very serious risk of bias
Velthuis, 2013[248]	Consecutive	NA as automated	yes	yes	26 people excluded due to detected AF prior to ELR monitoring, 13 excluded as discharged during monitoring or unco-operative and 6 signal quality insufficient. Apart from latter 6, most of these not lost for reasons related to outcome so not a risk of bias	No risk of bias
Vukajlovic, 2010[251]	consecutive	yes	Yes	Not simultaneous	none	Serious risk of bias
Wiesel, 2004[261]	NA as automated	Unclear	Unclear	Within 5 minutes	Unclear but 446/464 possible paired readings analysed. The loss of 18	Very serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
					readings probably does not constitute a risk of attrition bias	
Wiesel, 2009[260]	consecutive	NA as automated	Yes, blinded	Not simultaneous	None reported	Serious risk of bias
Wiesel, 2013[258]	consecutive	Effectively yes, as automated	Yes	ECGs done prior to BP measures so not simultaneous. However short interval of time.	21 lost – 10 withdrew before any readings, 1 did not record any ECG readings, 1 with a pacemaker erroneously registered and 9 did not record logs of AF-BP monitor readings. These relatively high losses may have removed the least compliant from the analysis thus biasing the analysis. However the logistic regression analysis adjusts for this, removing bias.	Serious risk of bias
Wiesel, 2014[259]	consecutive	unclear	Yes	Not simultaneous. ECG done just before index tests but time interval not reported	None	Very serious risk of bias
William, 2018[263]	Consecutive, but paired	yes	yes	Not simultaneous	62 non-interpretable readings, which were	Very serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
	analysis in that each patient was medically CV or not				not accounted for by paper's own analyses. These could indicate high risk of bias (could be argued that a non-interpretable reading would just prompt a further attempt and so just taking the interpretable readings is probably sensible, but the lack of interpretability may not be random and may be systematic and related to a specific person's waveform)	
Williams, 2015[264]	Case-control but not clear	yes	Yes	simultaneous	4 data points lost due to artefacts in the ECG recordings (or illegible). This does not reflect any issue with the index test and so the exclusion is appropriate and will not cause bias.	Serious risk of bias
Winkler, 2011[266]	consecutive	NA – automated using algorithm	unclear	Not simultaneous	2/60 data points lost due to problems with quality – but unclear if this was in index or gold standard ECG	Very serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
					readings. Nevertheless <10% so not a serious risk of bias	
Yan, 2018[267]	Consecutive	NA as automated	Yes	Not simultaneous but same session	16; presence of pacemaker (n=12), declined to complete all measurements (n=4)	Serious risk of bias
Himmelreich, 2019[89]	Consecutive	Yes	Yes	simultaneous	5 missing – 2 for missing 1 lead or 12 lead recordings and 3 for non-overlapping recordings. <10% so not a cause for concern	No serious risk of bias
Reverberi, 2019[189]	Consecutive	NA as automated	Yes	Not simultaneous	5 missing – due to spontaneous restoration of normal rhythm the day before the CV procedure. <10% so not a cause for concern	Serious risk of bias
Sabar, 2019[202]	Consecutive	Yes	Yes	Not simultaneous	103 missing – due to use for initial refining of algorithm. Not clear if this was part of the pre-hoc design of the study.	Very serious risk of bias
Wasserlauf, 2019[255]	Consecutive	Unclear	Unclear	simultaneous	None	Very serious risk of bias

Study	Random selection or case control	Index test with blinding of gold standard test results	Gold standard test with blinding of index test results	Time interval between index and gold standard	Loss of data from analysis	Overall risk of bias
Cunha, 2019[46]	Consecutive	Unclear	Unclear	Unclear	Unclear	Very serious risk of bias
Brown, 2019[22]	Consecutive	Unclear	No blinding	Simultaneous	5/265 lost due to no index test. Unlikely to be a risk of attrition bias.	Very serious risk of bias

Appendix I: Excluded studies

I.1 Excluded clinical studies

Table 84: Studies excluded from the clinical review

Reference	Reason for exclusion
Afzal 2015[2]	Systematic review
Alshraideh 2015[3]	Review
Alves 2019[4]	Does not evaluate accuracy for detecting AF
Athif 2018[7]	Unable to obtain
Barrett 2014[9]	Does not evaluate accuracy for detecting AF specifically (looked at several arrythmias together)
Barthelemy 2003[10]	Does not evaluate accuracy for detecting AF
Bell 2000[11]	Review
Berge 2018[12]	Does not evaluate accuracy for detecting AF
Berge, 2017[13]	Conference abstract
Bettin 2019[14]	not a point of care device
Beukema 2009[15]	No diagnostic accuracy data
Bonomi 2018[16]	Healthy controls had no reference standard measurment - assumed to be free from AF
Botto 2009[17]	No detection of specificity; simulation study
Bourdillon, 1978[18]	inappropriate gold standard - clinician decision without ECG
Brasier 2019[19]	Inappropriate reference standard - 1 lead internet enabled ECG
Brembilla-Perrot 2011[20]	myotonic dystrophy population
Buechi 2017[23]	SYSTEMATIC REVIEW
Burkowitz 2016[25]	SYSTEMATIC REVIEW
Busch 2017[26]	Not a diagnostic accuracy study; gold standard not defined
Callizo 2017[28]	AF prevalence study in the retinal vascular occlusion population. No diagnostic accuracy outcomes
Censi 2013[29]	Simulation study
Chan 2016[30]	reference standard is lead 1
Chan 2017[31]	reference standard is lead 1
Chan 2017[32]	reference standard is lead 1
Charitos 2012[33]	Detection or recurrence of AF; only sensitivity measured; simulation study
Chen, 2017[34]	manual interpretation of 1 lead ECG was gold standard
Choe 2015[35]	simulation study; no specificities reported; sensitivities reported mostly on a low resolution graph
Chong 2015[37]	unclear reference standard
Chong 2018[36]	Unclear gold standard
Chovancik 2019[38]	Unclear gold standard
Christensen 2014[39]	No diagnostic accuracy evaluation
Ciconte 2017[40]	not a point of care device - implantable cardiac monitor
Conroy 2017[41]	unclear reference standard - appears to be the pre-existing status of patients (healthy/ AF), but the method of original

Reference	Reason for exclusion
	diagnosis not described
Cooke 2006[42]	Review - check the 3 studies
Couderc 2015[43]	No diagnostic accuracy analysis for detection devices
Coutts 2014[44]	Review - check the Higgins study referred to
Dagres, 2010[47]	Letter to the editor
Damiano 2016[48]	Post-surgical ablation - not the population of interest; no diagnostic accuracy data
de Voogt 2006[49]	No diagnostic accuracy evaluation
DeBoard 2018[50]	not a point of care device
Defaye 1998[51]	pacemaker data
Derkac 2017[52]	No diagnostic accuracy evaluation
Diamantopoulos 2016[54]	cost effectiveness study
Dimarco 2018[55]	Does not evaluate accuracy for detecting AF
Dorr 2019[57]	Reference stanadard was I lead ECG
Dussault 2015[58]	SYSTEMATIC REVIEW
Edgerton 2011[59]	post-ablation population; no diagnostic accuracy evaluation
Eitel 2011[60]	not point of care device; no diagnostic accuracy evaluation
Elijovich 2009[61]	Does not evaluate accuracy for detecting AF
Engdahl 2013[62]	No diagnostic accuracy evaluation
Engdahl, 2018[63]	No diagnostic accuracy evaluation
Ermini 2013[64]	No diagnostic accuracy evaluation
Evans 2017[66]	Does not evaluate accuracy for detecting AF
Gaillard 2010[70]	no useful diagnostic accuracy data
Grond 2013[74]	No diagnostic accuracy evaluation
Gunalp 2006[75]	Does not evaluate accuracy for detecting AF
Hanke 2009[80]	Not evaluating a point of care test (implantable device)
Harju 2018[81]	Unclear if reference ECG was 12 lead
Harris 2012[82]	Systematic review
Hartikainen 2019[83]	Gold standard Holter << 24 hours (appears to be 5 mins)
Hendrikx 2014[85]	No gold standard specified
Higgins 2010[88]	trial website page only available
Higgins 2014[87]	Not a diagnostic study - investigating predictive value of early AF detection post stroke for later 90 day AF.
Hindricks 2010[90]	not a point of care device - ICM
Hochstadt 2019[93]	Unclear if ECG gold standard was 12 lead
Ip 2012[95]	Not evaluating a point of care test (implantable device)
Ip 2019[94]	opinion piece
Israel 2001[97]	Not a point of care device
Israel 2017[96]	No diagnostic accuracy evaluation
Jabaudon 2004[98]	No diagnostic accuracy evaluation
Jacobs, 2018[99]	health economic analysis paper
Jiang 2012[100]	Not testing a device but an algorithm to automate standard ECG or Holter detection.
K 2017[86]	No diagnostic accuracy evaluation

Reference	Reason for exclusion
Kaasenbrood, 2016[101]	cardiologist interpretation of my diagnostik ECG was gold standard, not full 12 lead
Kabutoya 2017[102]	unclear gold standard: details of the type of ECG not given
Kallmunzer 2012[107]	Gold standard was 'history of documented and verified AF' - not the protocol definition
Kallmunzer 2014[106]	Reference standard was 6 lead ECG
Kane 2016[108]	SYSTEMATIC REVIEW
Kang 2018[109]	Evaluation of heart sounds
Karaoguz 2019[111]	Does not evaluate accuracy for detecting AF
Kashiwa 2019[112]	Gold standard Holter but duration unclear (appears to be approximately 3hrs, which is <<24hrs minimum)
Kircher 2012[114]	Review
Kishore 2014[115]	SYSTEMATIC REVIEW
Korompoki 2017[119]	SR check refs
Koshy 2018[120]	Does not evaluate accuracy for detecting AF
Koshy 2018[121]	Does not evaluate accuracy for detecting AF
Kristensen 2016[122]	sick sinus syndrome patients
Krivoshei 2017[123]	unclear gold standard: details of the type of ECG not given
Lahdenoja 2018[126]	unclear gold standard: details of the type of ECG not given
Lau, 2013[128]	Conference abstract
Lauschke 2017[129]	not a point of care device
Lee 2013[130]	Unclear gold standard
Levin 2014[131]	cost-effectiveness study
Li 2019[133]	Review
Liao 2007[134]	SYSTEMATIC REVIEW
Liu 2010[136]	Does not evaluate accuracy for detecting AF
Lowe 2018[137]	simulation study, investigating modifiers to accuracy
Lowres, 2014[140]	Gold standard was predominantly lead 1 not lead 12
Lowres, 2016 [139]	No information on diagnostic accuracy
Martinek 2007[144]	not a point of care device in ablation population
Mehta 2015[147]	Does not evaluate accuracy for detecting AF
Miracapillo 2016[148]	Not a point of care device
Mittal 2013[149]	Not a point of care device
Montenero 2004[150]	not a point of care device
Morgan, 2002[151]	Reference standard was lead II rhythm strip
Narasimha 2018[155]	Does not evaluate accuracy for detecting AF
Nault 2019[157]	Does not evaluate accuracy for detecting AF
Nemati 2016[158]	Unclear gold standard
Nolker 2016[160]	Not a point of care device
Omboni, 2016[161]	interpretation of 1 lead ECG was gold standard
Orchard, 2016[163]	Gold standard was 2 cardiologists interpretation of the 1 lead iECG - not 12 lead ECG
Osaka 2017[164]	sick sinus syndrome patients
Osako 2002[165]	sick sinus syndrome patients; not a point of care device

Reference	Reason for exclusion
Pagola 2018[166]	Does not evaluate accuracy for detecting AF
Pastor-Perez 2010[168]	no evaluation of diagnostic accuracy for AF detection
Pedersen 2016[169]	no evaluation of diagnostic accuracy for AF detection
Philippsen 2017[171]	not a point of care device
Plummer 2001[172]	not a point of care device
Plummer 2003[173]	not a point of care device
Podd 2016[174]	not point of care devices
Poh 2018[175]	Unclear gold standard – appears to be lead I ECG
Purerfellner 2014[180]	not a point of care device
Purerfellner 2018[181]	not a point of care device
Rajakariar, 2018 [182]	Detection of atrial flutter only
Ramkumar 2018[183]	SR - check references
Reiffel 2005[184]	Does not evaluate accuracy for detecting AF
Reinsch 2018[186]	Not a point of care device
Rekhviashvili 2012[187]	no evaluation of diagnostic accuracy for AF detection
Ricci 1996[191]	sinus node disease patients
Rincon 2012[192]	Evaluated the accuracy of an algorithm by evaluating how well it picked up arrhythmias compared to 'manual annotations' using 10 hr ECG recordings from the MIT-BIH database. However these readings were not derived from the wearable wireless sensor platform under investigation - thus only the algorithm, not the device + algorithm, were evaluated. In addition unclear if 12 lead ECG and not >24hrs.
Ritter 2013[193]	not a point of care device
Roche 2002[195]	Not a diagnostic accuracy study; gold standard not defined
Rojo-Martinez 2013[196]	not in English
Rosenberg 2013[197]	Insufficient data to estimate specificity during period that gold standard was applied; no diagnostic accuracy analysis; no gold standard defined.
Ryabykina 2018[201]	not in English
Sack 2001[203]	not a point of care device
Salvatori 2015[204]	no evaluation of diagnostic accuracy
Samol 2013[205]	no evaluation of diagnostic accuracy
Sanak 2015[206]	no evaluation of diagnostic accuracy
Sanak 2015[207]	no evaluation of diagnostic accuracy
Sanders 2016[208]	not a point of care device
Schaefer 2014[209]	Unclear gold standard – although it seems 12 lead ECG was used, the accuracy results do not show results using this gold standard
Schuchert 1999[210]	Does not evaluate accuracy for detecting AF
Seidl 1998[212]	Not a point of care device
Sejr 2017[214]	Does not evaluate accuracy for detecting AF
Selder 2019[215]	Inappropriate gold standard - 1 lead ECG interpreted by cardiologist
Shafqat 2004[216]	no evaluation of diagnostic accuracy
Solomon 2016[218]	no evaluation of diagnostic accuracy
Solosenko 2019[219]	Unclear gold standard
Sposato 2012[221]	Does not evaluate accuracy for detecting AF

Reference	Reason for exclusion
Stahrenberg 2010[222]	no evaluation of diagnostic accuracy
Sudlow, 1998 [224]	limb lead ECG was gold standard
Suissa 2013[225]	Does not evaluate accuracy for detecting AF
Suissa 2014[226]	No diagnostic accuracy analysis for detection devices
Sutamnartpong 2014[227]	Does not evaluate accuracy for detecting AF
Svennberg, 2017 [228]	manual interpretation of 1 lead ECG was gold standard
Swerdlow 2000[230]	not a point of care device
Taggar 2016[231]	Systematic review
Takagi 2014[232]	unclear gold standard.
Tang 2017[233]	Unclear if gold standard 12 lead ECG (<<24 hours, normally only 10 minutes)
Tarakji 2015[234]	Gold standard was transtelephonic monitoring. Duration unclear and not Holter.
Tarniceriu 2018[235]	Unable to obtain
Tavernier, 2018 [236]	gold standard by expert consensus which did not necessarily involve 12 lead ECG
Terranova 2006[237]	no evaluation of diagnostic accuracy
Tison 2018[239]	all those in validation cohort had AF
Tu 2014[242]	no evaluation of diagnostic accuracy
Tu 2017[241]	protocol
Turakhia 2013[243]	no evaluation of diagnostic accuracy
Turakhia 2015[244]	no evaluation of diagnostic accuracy
Veale, [247]	protocol paper
Verberk 2012[249]	SR - checks refs
Verberk 2016[250]	review
Wachter 2013[253]	no evaluation of diagnostic accuracy
Wachter 2013[252]	trial website page only available
Welton 2017[256]	SR - checks refs
Wiegand 1997[257]	Data from pacemaker patients
Wiesel 2007[262]	Unclear if gold standard is 12 lead ECG
Willits 2014[265]	SR - checks refs
Yang 2016[268]	not a point of care device
Yang 2017[269]	no evaluation of diagnostic accuracy
Yenikomshian 2019[270]	SR - checks refs
Ziegler 2006[272]	Data from pacemaker patients
Ziegler 2017[273]	not a point of care device
Lumikari, 2019[141]	Non randomised; no diagnostic accuracy outcomes or data
Wang, 2020[254]	Gold standard not 12 lead
Etiwy, 2019[65]	Not aimed at detecting AF
Cuadrado-Godia, 2019[45]	Non randomised; no diagnostic accuracy outcomes or data
Eysenck, 2019[67]	Gold standard was intra-arterial BP assessment
Kwon, 2019[125]	Gold standard was lead I recording only
Kabutoya, 2019[103]	Gold standard was lead I recording only
Towhari, 2019[240]	Abstract only
Adami, 2019[1]	Prediction rather than detection of PAF

Reference	Reason for exclusion
Proietti, 2019[179]	modelling study; of HE relevance
Schukraft, 2019[211]	protocol only
Godin, 2019[72]	Non randomised; no diagnostic accuracy outcomes or data
Kong, 2019[118]	No diagnostic accuracy evaluation
Kalidas, 2019[105]	Algorithm tested on a database where gold standard is 2 lead
Valiaho, 2019[246]	Gold standard was 3 lead ECG
Hisazaki, 2019[91]	Accuracy of discriminating between sources of atrial arrythmias rather than existence of AF itself
Guo, 2019[76]	Non randomised; only those positive on index were given gold standard test
Attia, 2019[8]	Prediction rather than detection of PAF
Swancutt, 2004[229]	protocol only
Zaprutko, 2019[271]	Gold standard was index tests (lead I) interpreted by cardiologist
Guo, 2019[77]	No gold standard given to all; only those with a positive result on index test were given the 'gold standard', which could be clinical evaluation, ECG or 24hr Holter.
Perez-Valero, 2019[170]	Gold standard not specified as 12 lead ECG
Reiffel, 2020[185]	No diagnostic accuracy analysis; simulations of shorter point of care tests based on varying durations of ICM data
Oncu, 2019[162]	Gold standard not specified as 12 lead ECG

I.2 Excluded health economic studies

None.

Appendix J: Research recommendations

J.1 Detection of persistent AF

Research question: What is the diagnostic accuracy of key index tests (such as Alive Cor, MyDiagnostik, Microlife BP monitors, iphone plethysmography, pulse palpation) against the gold standard of 12 lead ECG, in people with risk factors for AF/symptoms of AF?

Why this is important:

In an ideal world every patient suspected of persistent AF would be given 12 lead ECG interpreted by a cardiologist, as this is the gold standard for AF diagnosis. Unfortunately, such 12 lead ECG is not always feasible to arrange in the primary care setting, as it is expensive, impractical and time-consuming. The ideal scenario would be the discovery of an alternative test that has comparable sensitivity and specificity to 12 lead ECG, but that is also cheap, simple and automated. The primary aim of this research question is therefore to evaluate if any currently available non-12 lead tests have sufficient accuracy to be used as a stand-alone diagnostic tool. The evidence to date is equivocal: although some devices appear to have excellent accuracy they are based on isolated, small or occasionally flawed studies, and further high-quality evidence is required.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People with risk factors for AF/symptoms of AF. Index tests(s): Key index tests such as the Alive Cor, MyDiagnostik, Microlife BP monitors, iphone plethysmography, pulse palpation Gold standard: 12 lead ECG interpreted by a cardiologist Outcome(s): sensitivity and specificity
Importance to patients or the population	At present the sub-optimal sensitivity of pulse palpation may lead to some patients with AF remaining undiagnosed, and therefore untreated, for a longer period of time. This may lead to avoidable strokes and other morbidity. More accurate initial tests would reduce these problems.
Relevance to NICE guidance	Good quality research in this area might allow NICE to recommend devices with more accurate detection of AF.
Relevance to the NHS	More accurate AF testing would lead to reductions in the costs of stroke.
National priorities	This is not relevant to a National priority area.
Current evidence base	In the guideline review, high accuracy was observed for several lead I devices, blood pressure monitors and plethysmographic tools. In mobile ECG devices, for example, sensitivity/ specificity values of 1.0/0.94 were found for the ECG check, 0.94/0.97 for my Diagnostik, 0.96/0.92 for the Zenecor thumb device and 1.0/1.0 for the Cardiobip. Similarly, the heart spectrum blood pressure monitor had sensitivity/sensitivity of 0.97/0.97, and iPhone plethysmographic devices had values of 0.97/0.93. However there was often uncertainty of the true accuracy because of a lack of statistical power. For example, the ECG check, Cardiobip, Zenecor and heart spectrum evidence were based on very small single studies (n=36 to n=100). In addition studies were limited by methodological limitations such as poor blinding of tests. It is hoped that this research recommendation will lead to high quality research that will provide precise and robust evidence to add to the current knowledge base.
Equality	This research recommendation does not address equality issues.
Study design	Cross-sectional diagnostic study. Ideally all index tests would be evaluated on each participant, with a separate 12 lead ECG done

	simultaneously for each test.
Feasibility	There are no ethical issues, and the proposed research can be carried out on a realistic timescale and at a reasonable cost. One issue will be the use of several tests on the same person with a separate 12 lead ECG done concurrently with each. This will lead to the inconvenience and possible discomfort of participants, and may interfere with the patient's clinical care. There are no known harms of AF testing and so it is not envisaged that multiple testing will increase the risk of adverse effects.
Other comments	None
Importance	• High: the research is essential to inform future updates of key recommendations in the guideline.

Detection of paroxysmal AF

Research question: A.1 What is the diagnostic accuracy of key index tests (to be specified) against the absolute gold standard (to be determined) of prolonged ambulatory monitoring, in people suspected of having paroxysmal AF?

Why this is important:

Detection of paroxysmal AF is difficult. Due to the episodic nature of paroxysmal AF, it may not be detected by a single point-in-time test. It is therefore important to be able to accurately detect paroxysmal AF using a strategy that takes account of this, possibly by allowing multiple measurements over days or weeks. An accurate test for paroxysmal AF will reduce the number of undetected cases, and therefore reduce the number of strokes and other adverse events.

The current evidence base suggests that some ambulatory tests using mobile technology may be useful to detect paroxysmal AF. However the estimates of accuracy are uncertain and the quality of data is poor. Many studies were small-scale and a major limitation was the quality of the reference standard used in the studies. Although the reference standard should be the 'gold' standard (i.e., the reference standard should provide a 'true' diagnosis, or the closest possible approximation to it) there does not seem to be an established reference standard used for paroxysmal AF. For example, in many studies a 24 hour Holter monitor was used as the reference standard. Such a reference standard may tend to over-estimate the sensitivity of the test devices because other studies have shown that a 24 hour Holter monitor to only pick up a small fraction of cases.

This research study aims to compare current devices to establish their accuracy. This study will attempt to avoid the drawbacks of previous work, using large numbers, and a robust reference (gold) standard.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People with suspected paroxysmal AF. Suspicion is most likely to relate to symptoms that suggest AF episodes. Index tests(s): Key index tests such as mobile lead I devices, mobile BP monitors, i-phone plethysmography, or skin patches used on a repeated basis over a time period that matches the patients' patterns of symptoms Gold standard: To be determined. 24 hour Holter should not be used as it has not been shown to be a true gold standard. Outcome(s): sensitivity and specificity
Importance to patients or the	At present the sub-optimal methods of detecting paroxysmal AF may lead to some patients with AF remaining undiagnosed, and therefore untreated,

population	for a longer period of time. This may lead to avoidable strokes and other morbidity.
Relevance to NICE guidance	Good quality research in this area might allow NICE to recommend devices/strategies with more accurate detection of AF.
Relevance to the NHS	New guidance that recommends a particular investigation to detect potential paroxysmal AF could lead to an increase in the number of investigations in the community, possibly increased number of referrals to secondary care and also an increase in the number of new diagnoses of AF. This would have some resource implications. These patients would then presumably be anti-coagulated which has a cost. However, that cost is very likely to be less that the costs associated with them not being diagnosed and having a stroke with the associated morbidity and mortality. More accurate tests would reduce these problems
National priorities	This is relevant to a National priority area. In the new Primary Care Network DES for 2020 there is a section on 'Anticipatory Care'. This asks GPs in networks (groups of GP practices) to "identify priority patients at risk of unwarranted health outcomes". This would certainly include those with undiagnosed AF at risk of stroke. Please see: https://www.engage.england.nhs.uk/survey/primary-care-networks-service-specifications/supporting_documents/Draft%20PCN%20Service%20Specifications%20December%202019.pdf
Current evidence base	The current evidence base is uncertain, as many studies were small-scale and the gold standards were frequently not appropriate. For example, the Kardia-band had an excellent sensitivity/specificity of 0.98/0.99 but this was based on a single study of just 26 people. Uncertainty of the true population effect was thus very high. As another example, the Microlife Watch BP device used at 20 minute intervals over 24 hours had a good sensitivity/specificity of 0.93/0.98, based on a large study of 5778 people. However, the gold standard was a 24 hour Holter device, which has been shown to be insensitive compared to other gold standards. Thus further high quality research is required.
Equality	This research recommendation does not address equality issues. We did not identify specific ethnicities or other groups that should be investigated in a different way, or prioritised, but we are not aware of there being apparent or implied discrimination in the recommendation as it stands. People with learning disabilities have worse cardiovascular morbidity and mortality, as do those with severe and enduring mental health problems. The reasons for this are multi-factorial.
Study design	Cross-sectional diagnostic study. Ideally all index tests would be evaluated on each participant.
Feasibility	The proposed research can be carried out on a realistic timescale and at a reasonable cost. We are not aware of specific ethical issues though technical issues are a possibility depending upon the type of technology used.
Other comments	
Importance	• High: the research is essential to inform future updates of key recommendations in the guideline.

1