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 population	At present the sub-optimal methods of detecting paroxysmal AF 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.
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.