Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis

Nicky J Welton, Alexandra McAleenan, Howard HZ Thom, Philippa Davies, Will Hollingworth, Julian PT Higgins, George Okoli, Jonathan AC Sterne, Gene Feder, Diane Eaton, Aroon Hingorani, Christopher Fawsitt, Trudie Lobban, Peter Bryden, Alison Richards and Reecha Sofat

NHS

National Institute for Health Research

Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis

Nicky J Welton,[1]* Alexandra McAleenan,[1] Howard HZ Thom,[1] Philippa Davies,[1] Will Hollingworth,[1] Julian PT Higgins,[1] George Okoli,[1] Jonathan AC Sterne,[1] Gene Feder,[1] Diane Eaton,[2] Aroon Hingorani,[3] Christopher Fawsitt,[1] Trudie Lobban,[4,5] Peter Bryden,[1] Alison Richards[1] and Reecha Sofat[6]

[1]School of Social and Community Medicine, Faculty of Health Sciences, University of Bristol, Bristol, UK
[2]AntiCoagulation Europe, Bromley, UK
[3]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK
[4]Atrial Fibrillation Association, Shipston on Stour, UK
[5]Arrythmia Alliance, Shipston on Stour, UK
[6]Division of Medicine, Faculty of Medical Science, University College London, London, UK

*Corresponding author

Health Technology Assessment

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

Abstract

Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis

Nicky J Welton,[1]* Alexandra McAleenan,[1] Howard HZ Thom,[1] Philippa Davies,[1] Will Hollingworth,[1] Julian PT Higgins,[1] George Okoli,[1] Jonathan AC Sterne,[1] Gene Feder,[1] Diane Eaton,[2] Aroon Hingorani,[3] Christopher Fawsitt,[1] Trudie Lobban,[4,5] Peter Bryden,[1] Alison Richards[1] and Reecha Sofat[6]

[1]School of Social and Community Medicine, Faculty of Health Sciences, University of Bristol, Bristol, UK
[2]AntiCoagulation Europe, Bromley, UK
[3]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK
[4]Atrial Fibrillation Association, Shipston on Stour, UK
[5]Arrythmia Alliance, Shipston on Stour, UK
[6]Division of Medicine, Faculty of Medical Science, University College London, London, UK

*Corresponding author Nicky.Welton@bristol.ac.uk

Background: Atrial fibrillation (AF) is a common cardiac arrhythmia that increases the risk of thromboembolic events. Anticoagulation therapy to prevent AF-related stroke has been shown to be cost-effective. A national screening programme for AF may prevent AF-related events, but would involve a substantial investment of NHS resources.

Objectives: To conduct a systematic review of the diagnostic test accuracy (DTA) of screening tests for AF, update a systematic review of comparative studies evaluating screening strategies for AF, develop an economic model to compare the cost-effectiveness of different screening strategies and review observational studies of AF screening to provide inputs to the model.

Design: Systematic review, meta-analysis and cost-effectiveness analysis.

Setting: Primary care.

Participants: Adults.

Intervention: Screening strategies, defined by screening test, age at initial and final screens, screening interval and format of screening {systematic opportunistic screening [individuals offered screening if they consult with their general practitioner (GP)] or systematic population screening (when all eligible individuals are invited to screening)}.

Main outcome measures: Sensitivity, specificity and diagnostic odds ratios; the odds ratio of detecting new AF cases compared with no screening; and the mean incremental net benefit compared with no screening.

Review methods: Two reviewers screened the search results, extracted data and assessed the risk of bias. A DTA meta-analysis was perfomed, and a decision tree and Markov model was used to evaluate the cost-effectiveness of the screening strategies.

Results: Diagnostic test accuracy depended on the screening test and how it was interpreted. In general, the screening tests identified in our review had high sensitivity (> 0.9). Systematic population and systematic opportunistic screening strategies were found to be similarly effective, with an estimated 170 individuals needed to be screened to detect one additional AF case compared with no screening. Systematic opportunistic screening was more likely to be cost-effective than systematic population screening, as long as the uptake of opportunistic screening observed in randomised controlled trials translates to practice. Modified blood pressure monitors, photoplethysmography or nurse pulse palpation were more likely to be cost-effective than other screening tests. A screening strategy with an initial screening age of 65 years and repeated screens every 5 years until age 80 years was likely to be cost-effective, provided that compliance with treatment does not decline with increasing age.

Conclusions: A national screening programme for AF is likely to represent a cost-effective use of resources. Systematic opportunistic screening is more likely to be cost-effective than systematic population screening. Nurse pulse palpation or modified blood pressure monitors would be appropriate screening tests, with confirmation by diagnostic 12-lead electrocardiography interpreted by a trained GP, with referral to a specialist in the case of an unclear diagnosis. Implementation strategies to operationalise uptake of systematic opportunistic screening in primary care should accompany any screening recommendations.

Limitations: Many inputs for the economic model relied on a single trial [the Screening for Atrial Fibrillation in the Elderly (SAFE) study] and DTA results were based on a few studies at high risk of bias/of low applicability.

Future work: Comparative studies measuring long-term outcomes of screening strategies and DTA studies for new, emerging technologies and to replicate the results for photoplethysmography and GP interpretation of 12-lead electrocardiography in a screening population.

Study registration: This study is registered as PROSPERO CRD42014013739.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

CONTENTS

List of tables

List of figures

List of abbreviations

AF	atrial fibrillation	INB	incremental net benefit
CEAC	cost-effectiveness acceptability curve	INR	international normalised ratio
		MI	myocardial infarction
CHA_2DS_2-VASc	Score for Stroke Risk in Atrial Fibrillation	NICE	National Institute for Health and Care Excellence
CHEERS	Consolidated Health Economic Evaluation Reporting Standards	NIHR	National Institute for Health Research
CI	confidence interval (Bayesian analysis)	NNS	number needed to screen
CRD	Centre for Reviews and Dissemination	NOAC	novel oral anticoagulant
		OAC	oral anticoagulation
CrI	credible interval (Bayesian analysis)	OR	odds ratio
		PICO	population, intervention, comparator, outcome
DIC	deviance information criterion		
DOAC	directly acting oral anticoagulant	PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
DOR	diagnostic odds ratio		
DTA	diagnostic test accuracy	PSSRU	Personal Social Services Research Unit
ECG	electrocardiogram		
EQ-5D	EuroQol-5 Dimensions	QALY	quality-adjusted life-year
GP	general practitioner	QOF	Quality and Outcomes Framework
HAS-BLED	Score for Bleeding Risk in Atrial Fibrillation	QUADAS-2	QUality Assessment of Diagnostic Accuracy Studies – 2
HIQA	Health Information and Quality Authority	RCT	randomised controlled trial
HR	hazard ratio	SROC	summary receiver operating characteristic
HSROC	hierarchical summary receiver operating characteristic	TIA	transient ischaemic attack
HTA	health technology assessment	TILDA	Irish Longitudinal Study on Ageing
ICD	implantable cardioverter defibrillator	WHO	World Health Organization
ICER	incremental cost-effectiveness ratio		

Plain English summary

Atrial fibrillation (AF) is the most common heart rhythm disorder. AF is a major cause of stroke and subsequent disability and mortality. Many AF patients benefit from blood-thinning drugs (anticoagulation therapies), which have been shown to reduce the risk of an AF-related stroke. AF is often diagnosed only after an AF-related stroke. Earlier detection could prevent AF-related strokes, disability and mortality. This study aimed to assess whether or not a national screening programme for AF would represent good value for money. To do this, we reviewed studies looking at how well screening tests identify individuals with AF and also studies that compare different methods of screening. We found that a national screening programme in which older individuals (aged 65–70 years) are invited to be screened during a routine appointment with their doctor is likely to be cost-effective. There was also evidence that repeated screening every 5 years until age 80 years would prove to be of value. The screening test most likely to be cost-effective involves either a nurse taking a pulse rhythm check or a specific blood pressure monitor that is also able to detect AF. Whichever screening test is used, patients with a positive result would need to have their diagnosis confirmed by a trained doctor using an electrocardiogram machine before discussing treatment options.

New studies on (1) how to implement screening in general practice, (2) the performance of new tests (patches, smartphone/watch devices, iPads, hand-held devices) and (3) the effectiveness of screening with regard to the long-term risks of AF-related stroke and mortality would be useful.

Scientific summary

Background

Atrial fibrillation (AF) is the most common cardiac arrhythmia. The prevalence of AF increases with age, and AF is more prevalent in men than in women. AF increases the risk of thromboembolic events, in particular stroke, with a fivefold increase in stroke risk in individuals with AF. Treatment of AF focuses on rhythm and rate control, and on prevention of stroke using oral anticoagulation therapy in individuals in whom return to sinus rhythm is unlikely. It is estimated that one-third of those with persistent AF will not have symptoms, and therefore that a first presentation of this might be a stroke. It is these individuals who would benefit most from a national screening programme for AF and subsequent anticoagulation therapy for stroke prevention. A national screening programme for AF would, however, involve a substantial investment of NHS resources, and the cost-effectiveness of such a programme would need to be established.

Objectives

The aim of this study was to assess the cost-effectiveness of different national screening strategies for AF in older adults. To achieve this aim we:

1. conducted a systematic review of diagnostic test accuracy (DTA) studies to determine the diagnostic accuracy of screening tests for detecting AF in adults who have not sought medical attention in a primary or community care setting on account of symptoms associated with AF, and to determine the diagnostic accuracy of screening tests in systematic opportunistic, targeted and population screening settings
2. updated a previous systematic review of screening strategies for AF to answer the following questions: (i) does systematic screening increase the detection of AF compared with routine practice?; (ii) what are the characteristics of those identified with AF by screening strategy?; (iii) which combination of screening strategy, screening population and test is the most effective at detecting AF compared with routine practice?; (iv) what are the potential safety issues and adverse events associated with individual screening programmes?; (v) how acceptable is the intervention to the target population?; and (vi) what are the costs associated with systematic screening for AF?
3. developed an economic model to compare the cost-effectiveness of different national screening strategies (including no screening) based on a review of previous economic evaluations of screening for AF and a review of recent literature on the prevalence, disease progression and risk profiles of AF and screening strategies relevant to a UK population-based screening setting.

Methods

Diagnostic test accuracy review

We searched for diagnostic cohort or case–control studies in MEDLINE, PreMEDLINE, EMBASE, The Cochrane Library, the Centre for Reviews and Dissemination and the Science Citation Index. Databases were searched without language or date restrictions until January 2015. The population of interest was individuals registered in primary care and/or presenting to primary care or community centres who had not sought medical attention on account of symptoms associated with AF. The index test could be any non-invasive test for AF that could be utilised in a primary care setting or in the community. The reference standard was a 12-lead electrocardiogram (ECG) interpreted by a cardiologist. Two reviewers screened the search results, extracted and checked the data and assessed the study quality using the QUality Assessment of Diagnostic Accuracy Studies – 2 (QUADAS-2) tool. Evidence synthesis was conducted by fitting a hierarchical summary receiver operating characteristic

(HSROC) model to estimate the relationship between sensitivity and specificity of the index tests, accounting for heterogeneity between studies. Heterogeneity was explored using preplanned subgroup analyses.

Review of randomised controlled trials comparing screening strategies

We updated and adapted a recent Cochrane review of screening strategies for AF. The population of interest was adults aged ≥ 40 years in whom population-based screening could be conducted. The interventions of interest were population-based, systematic (opportunistic, targeted or population) screening programmes for AF. We relaxed the requirement for confirmation of AF with a 12-lead ECG interpreted by a general practitioner (GP), specialist or suitably trained ECG technician because of a lack of studies. During the course of our review, we identified that the Health Information and Quality Authority (HIQA) in Ireland had published a health technology assessment (HTA) in which it had updated the Cochrane review from June 2012 to June 2015. We were also informed that the Cochrane review itself was being updated (currently undergoing peer review) and, with the assistance of the review authors, we updated this version of the review while expanding the scope as described above. Because the Cochrane update was restricted to randomised controlled trials (RCTs), we also restricted our review to RCTs. In addition, we updated the update by running searches from July 2015 to December 2015. Two reviewers screened the search results, extracted and checked the data and assessed the risk of bias. The study results were pooled using meta-analysis when sufficient evidence was available.

Economic evaluation

We reviewed the literature to identify previous economic evaluations of screening strategies for AF, using the NHS Economic Evaluation (NHS EED) and HTA databases from inception until close (end of December 2015), the Cochrane screening study review searches with economic filters and the Cost-Effectiveness Analysis (CEA) Registry. On the basis of this review, together with discussions with the project team, we developed an economic model consisting of a decision tree for the screening process and outcome and a discrete-time Markov model from a linked National Institute for Health Research HTA project, to provide the expected lifetime costs and benefits of anticoagulation therapy, if appropriate, given the outcomes of the screening model. We searched MEDLINE and EMBASE between 1 January 2000 and 22 January 2016 to identify studies providing information on the natural history of AF and screening for AF relevant to our economic model. The sensitivity and specificity of different screening tests came from the results of the DTA review meta-analysis. A probabilistic analysis is presented to reflect uncertainty in the model inputs. We report the mean total incremental costs, mean total incremental quality-adjusted life-years (QALYs) and mean incremental net benefit (INB) at a willingness to pay of £20,000 per QALY. To compare different screening strategies by age at first screen and number of repeated screens, we multiplied the per-person INB by the population size that will benefit (for a given age cohort). Costs and outcomes in future years over a lifetime time horizon were discounted at an annual rate of 3.5%.

Results

Diagnostic test accuracy review

We identified 15 studies of screening tests for the detection of AF, including a 12-lead ECG, single-lead ECGs, between 1- and 12-lead ECGs, pulse palpation, modified blood pressure monitors, photoplethysmography and two-stage testing. Screening tests varied in whether they were interpreted by a cardiologist, a GP, a nurse or an automatic algorithm, although evidence was not available for every test and interpreter combination and, when present, few observations were reported, leading to a lack of statistical power to detect meaningful differences. There was a high degree of variability between studies and a high level of uncertainty in the estimates of DTA. In general, most tests had a high sensitivity, in excess of 0.9. Specificity was, in general, lower than sensitivity for all of the tests. Tests with the highest diagnostic odds ratio (DOR) were the 12-lead ECG (regardless of interpreter), between 1- and 12-lead ECG (automatic or cardiologist interpretation), two-stage tests and single-lead ECG interpreted by a GP, with all of these tests having similar DORs.

In general, for a given interpreter, the results for the single-lead ECG were less accurate and more variable than ECGs with more than one lead. Nurse interpretation of single-lead ECGs performed similarly to single-lead ECGs with other interpretation methods, but nurse interpretation of 12-lead ECGs did not perform as well as other interpretation methods. Automatic interpretation did not have a consistent impact on test accuracy, with automatic interpretation of single-lead ECGs having a high sensitivity but variable specificity. In contrast, automatic interpretation of ECGs with more leads had good specificity but variable sensitivity. The different two-stage screening strategies all had very high specificity, but sensitivity was high only when a 12-lead ECG was used as the second-stage test.

Although photoplethysmography had very high test sensitivity, this estimate was based on a single study not based in primary care, in which the index test was rated as being at high risk of bias and the applicability of the patient selection was unclear. We are therefore cautious in interpreting the results from this screening test.

The results were robust to the sensitivity analyses, but we could perform only some of these because of a lack of data.

Review of randomised controlled trials comparing screening strategies

We identified five RCTs comparing screening strategies for AF; however, only two of these provided data that could be included in our review and only one was included for our primary outcome (the number of new AF diagnoses). The Screening for Atrial Fibrillation in the Elderly (SAFE) study therefore remains the main source of evidence on the comparative efficacy of different screening strategies for AF. Systematic population and systematic opportunistic screening strategies were found to be similarly effective, with an estimated 170 individuals needed to be screened to detect one additional AF case compared with no screening. There was no evidence that systematic screening targeted to high-risk individuals was effective compared with no screening.

Uptake of systematic population screening was typically around 50%, although uptake was as high as 70% in one study and there was variability in uptake between practices. Reasons for not attending for screening were varied, although older age and decreased mobility were commonly cited reasons. The proportion of individuals having their pulses checked under systematic opportunistic screening varied across studies (from 30% to 66%) and between practices within studies (from 8% to 93%). The proportion of individuals consulting with their GP was not reported, so it is unclear how much these uptake rates are driven by consultation rates, GPs offering pulse palpation and uptake of pulse palpation by individuals. Of those with an irregular pulse who did not have a previous diagnosis of AF, approximately 18% did not attend for an ECG test, although again this was variable across practices.

For systematic opportunistic screening, a greater proportion of the 75 newly identified cases were diagnosed outside the screening programme (44/75, 59%) than within it (31/75, 41%). This suggests that the full benefits seen in the systematic opportunistic arm may not be realised outside the context of a RCT.

There were no comparative studies of repeated screening strategies, and the majority of the evidence related to those aged 65–75 years.

Economic evaluation

Our results indicate that both systematic opportunistic and systematic population screening followed by directly acting oral anticoagulant therapy, where indicated, are likely to be cost-effective compared with no screening (current practice). Systematic opportunistic screening was more likely to be cost-effective than systematic population screening, as long as the proportion of flagged individuals who have their pulses checked observed in the SAFE study is realised in practice.

We found that photoplethysmography, modified blood pressure monitors and pulse palpation by a nurse were more likely to be cost-effective than other screening tests, because these are cheaper than other

screening tests while having adequate test sensitivity. This finding relies on the use of a 12-lead ECG diagnostic test interpreted by a GP (referred to a cardiologist when the diagnosis is unclear) in individuals with a positive screening test result.

For a single screen of a given age cohort, we found that strategies that use a higher age of screening were more likely to be cost-effective. However, when allowing for the possibility of repeated screening strategies with 5-year intervals, single screens were no longer found to be cost-effective; instead, an initial screen at age 65 followed by repeat screens every 5 years until age 80 years was found to be most likely to be cost-effective, provided that compliance with treatment does not decline with increasing age.

Although our model made many assumptions, we found that in general these were robust when tested in sensitivity analyses. If the uptake of pulse palpation in flagged patients is lower than that seen in the SAFE study, then a screening strategy in which individuals are invited by letter may be more appropriate. If the diagnosis rate of AF in the absence of screening is high, then the age at which to initiate screening is likely to be higher (70 years rather than 65 years).

Conclusions

A national screening programme for AF is likely to represent a cost-effective use of resources. Systematic opportunistic screening is more likely to be cost-effective than systematic population screening. Nurse pulse palpation or modified blood pressure monitors (if available) would be appropriate screening tests, followed by a diagnostic 12-lead ECG interpreted by a trained GP in those who screen positive, with referral to a cardiologist/specialist in cases in which the diagnosis is unclear. Implementation strategies to operationalise uptake of opportunistic screening in primary care should accompany any screening recommendations.

Research needs identified by this report are (1) the development and evaluation of strategies for the implementation of screening; (2) the evaluation of diagnostic performance against a reference standard of recent innovations (patches, smartphone/watch devices, iPads, hand-held devices) as well as devices used for other reasons that can also detect AF (pacemakers, implantable cardioverter defibrillator and implantable loop recorder devices), which may be of relevance to screening, particularly the detection of paroxysmal AF; (3) a comparative study to evaluate the long-term benefits of screening for chronic and paroxysmal AF and the yield from repeated screening tests; and (4) studies to replicate the DTA results for photoplethysmography and GP interpretation of 12-lead ECGs (with referral to a cardiologist/specialist in cases in which diagnosis is unclear) in a screening population.

Study registration

This study is registered as PROSPERO CRD42014013739.

Funding

Chapter 1 Introduction

Atrial fibrillation and its health consequences

Atrial fibrillation (AF) is the most common cardiac arrhythmia. The prevalence of AF increases with age and it is more prevalent in men than in women.[1,2] The prevalence of diagnosed AF in England has been estimated to be 1.6%,[3] although total prevalence including undiagnosed AF has been estimated at 2.4%.[1] Estimates of annual incidence range from 0.025% in the under-sixties to 0.3% in the over-eighty-fives.[4] Symptoms of AF include palpitations, breathlessness, syncope or presyncope, angina and exhaustion. AF is typically categorised into three types, and patients may remain in one category or can progress through these types of AF as their condition worsens. The categories are:

1. paroxysmal AF, which is intermittent with episodes that usually last < 48 hours and that stop without antiarrhythmic therapy
2. persistent AF, which lasts for > 1 week if untreated with antiarrhythmic therapy
3. permanent AF, which is no longer corrected with antiarrhythmic therapy.

Atrial fibrillation increases the risk of thromboembolic events, in particular stroke, because of pooling of blood in the left atrium and embolisation to the brain. The risk of stroke is increased fivefold in individuals with AF.[5] Approximately one-sixth of all strokes are attributed to AF.[6] Patients with thromboembolic stroke from AF have a higher mortality and morbidity than patients with other stroke subtypes.[7] Moreover, there appears to be a 'dose–response' relationship, with stroke risk increasing as AF progresses from paroxysmal to persistent to permanent AF.[8,9] The presence of AF is also associated with an approximately twofold higher risk of future acute myocardial infarction (MI).[10]

Treatment of AF focuses on rhythm and rate control, and the prevention of stroke using oral anticoagulation (OAC) therapy in individuals in whom return to sinus rhythm is unlikely. A strategy of treatment is recommended by the National Institute for Health and Care Excellence (NICE)[11] and similar bodies worldwide, for example, the Canadian Agency for Drugs and Technologies in Health (CADTH)[12] and the European Society of Cardiology.[13] However, this is only in individuals in whom a diagnosis of AF has already been made because they present with symptoms. It is estimated that, of those with persistent AF, one-third will not have symptoms[14] and therefore a first presentation of persistent AF might be a stroke. It is these individuals who would benefit most from anticoagulation therapy for stroke prevention, and underdetection of AF in this population represents a major gap in clinical care.[15] OAC therapies include warfarin and a class of novel directly acting oral anticoagulant therapies termed DOACs or NOACs (including dabigatran (Pradaxa®, Prazaxa®, Pradax®, Boehringer Ingelheim GmbH, Germany), rivaroxaban (Xarelto®, Bayer Health Care, Germany), apixaban (Eliquis®; Bristol-Myers Squibb, NY, USA) and edoxaban (Lixiana®, Daiichi Sankyo, Japan). Because NOACs will not remain 'novel' forever, in this report we will refer to this class of therapies as DOACs. Given the potential benefits of anticoagulation therapy and the underdiagnosis of AF, calls have been made for a national screening programme to identify patients with AF who may benefit from the prevention of AF-related stroke.[16]

A national screening programme would involve a substantial investment of NHS resources, and the cost-effectiveness of such a programme needs to be demonstrated. The benefits of screening depend on the additional number of AF cases detected and the resulting reduction in AF-related sequelae, in particular stroke, achieved by earlier therapeutic intervention in the additional detected cases.

Current practice in the diagnosis and treatment of atrial fibrillation

There is currently no systematic screening strategy for AF in the UK. Current NICE guidelines[11] state that a pulse palpation should be performed to assess for AF in patients presenting with dyspnoea, palpitations, syncope or chest discomfort, or who have had a transient ischaemic attack (TIA). Once AF is identified (symptomatic, asymptomatic, paroxysmal, persistent or permanent), management strategies involve rate and/ or rhythm control. For rate control, medications including beta-blockers, digoxin or calcium channel blockers are recommended. Individuals who have new-onset AF or who are young with persistent AF may be offered electrical cardioversion, with a 4-week period of anticoagulation prior to cardioversion. Rhythm control may be achieved with medications used for rate control in some individuals. Where rhythm control is unlikely to be achieved, or there is a high risk of arrhythmia following cardioversion (electrical or by medication), it is necessary to assess stroke risk using the Score for Stroke Risk in Atrial Fibrillation (CHA_2DS_2-VASc) [constructed as a sum of the following risk factors: congestive heart failure, hypertension, aged $\geq$ 75 years (x2), diabetes, stroke (x2), vascular disease, age 65–74 years, and sex (female)].[17,18] Anticoagulation should be offered to patients with a CHA_2DS_2-VASc score of $\geq$ 2, taking bleeding risk into account [using the Score for Bleeding Risk in Atrial Fibrillation (HAS-BLED)]. Left atrial occlusion may be offered to patients who are deemed to be at high risk of anticoagulation because of comorbidities including falls and bleeding risk or intolerance to drugs. Ablative therapies may also be offered for the same indication.

Systematic screening strategies

A population-based screening programme for AF could take a variety of different formats, depending on (1) the population to be screened, (2) the screening test procedure and (3) the setting in which it is conducted. We follow Hobbs *et al.*[19] in using the following definitions of screening strategies:

- systematic population screening: general screening of a defined population, for example individuals aged $\geq$ 65 years
- systematic targeted screening: screening people at higher risk of AF, for example those with risk factors such as cardiac failure, hypertension, rheumatic heart disease or history of MI, angina, diabetes mellitus, hyperthyroidism, stroke or TIA
- systematic opportunistic screening: when a health-care professional takes the opportunity to screen an individual for AF during a consultation.

Current NICE guidelines[11] represent opportunistic case finding, rather than a screening strategy.

The current strategy recommended by NICE is pulse palpation, although any screening strategy would need to consider how this or other methods perform compared with a gold standard. For permanent or persistent AF, a 12-lead electrocardiogram (ECG) interpreted by at least one trained cardiologist/heart rhythm specialist is considered a gold standard. However, in a screening programme, quicker/cheaper methods are likely to be adopted, with confirmation of a positive result using the gold standard as necessary. Paroxysmal AF is challenging to diagnose because of its intermittent nature, and there is no accepted gold standard method for its diagnosis.[20] Continuous Holter monitoring has been shown to detect more cases than intermittent monitoring, and more cases are detected as the length of monitoring increases from 24 hours to 7 days to 30 days.[20]

For a national screening programme, we assumed that primary care, community care (e.g. pharmacies)[21] or domiciliary testing would be the most appropriate setting.[22] However, we restricted our evaluation to a primary care setting as being the most likely setting for an AF screening programme in England and Wales. Detection of AF occurs in secondary care following investigations after an event (e.g. ischaemic stroke); however, because current NHS practice following such an event is to investigate potential causes, such as AF, then in effect an opportunistic case-finding programme is already in place in this setting, and therefore this setting was not included in the analysis.

The need for an evaluation of the cost-effectiveness of a national screening programme for atrial fibrillation

The World Health Organization (WHO) sets out the Wilson and Jungner[23] criteria that should be met before the introduction of a population-based screening programme. Each of these criteria has recently been assessed for a population-based screening programme for AF.[24–26] The criteria relevant to AF are discussed in the following sections.

The condition should be an important health problem

Atrial fibrillation is a major cause of ischaemic stroke, with a fivefold increase in risk of ischaemic stroke[5] and associated increased mortality,[5,6,27–29] increased morbidity[15] and reduced quality of life.[30] AF incidence and prevalence increases with age, and a recent report with 50 years' follow-up indicates that the incidence and prevalence of AF are increasing over time, largely because of an ageing population and increased detection.[31] It has been estimated that AF is responsible for about 25,000 AF-related strokes per year in the UK.[6] AF therefore represents an important health problem that is likely to increase over time with an ageing population.

The natural history of the disease, including from latent to declared disease, should be adequately understood and there should be a detectable disease marker for the latent period/early symptomatic stage of the disease

The natural history of AF is a progression from paroxysmal to persistent to permanent AF.[32] Risk factors for AF-related stroke risk in patients are well described by the CHA_2DS_2-VASc score.[11] All types of AF may be either symptomatic or asymptomatic.[33,34] It is expected that, compared with routine practice, a population-based screening programme is likely to detect a higher proportion of asymptomatic cases that would not otherwise present until a stroke or cardiovascular event occurs. We would also expect that, unless an extended screening test is used, the majority of screen-detected cases would be persistent or permanent AF, because only a small proportion of paroxysmal cases would be in AF at the time of testing. Flaker et al.[35] found that stroke risk was higher in symptomatic patients than asymptomatic patients, but that this difference was explained by baseline characteristics.

Cost-effective primary prevention interventions should have been implemented as far as practical

Across the UK there are programmes in place to help prevent modifiable risk factors of AF (MI, angina, diabetes mellitus, hyperthyroidism, stroke or TIA[11,19]).

There should be a simple, safe, precise and validated screening test, and the distribution of the test values in the target population should be known and a suitable cut-off level defined and agreed

A 12-lead ECG interpreted by at least one cardiologist is considered a gold standard for the diagnosis of permanent or persistent AF. However, in a screening programme quicker/cheaper screening tests are likely to be adopted, with confirmation of a positive result using the gold standard diagnostic test as deemed necessary. In one arm of the Screening for Atrial Fibrillation in the Elderly (SAFE) study, Hobbs et al.[19] used pulse taking as a screening test with a follow-up ECG if an irregular pulse was found, whereas in another arm a 12-lead ECG interpreted by a cardiologist was used directly as a screening test. However, there are a variety of possible screening tests available (e.g. modified blood pressure monitors, one-lead ECG, three-lead ECG, Holter monitor), as well as devices used for other reasons that can also detect AF [pacemakers, implantable cardioverter defibrillator (ICD) and implantable loop recorder devices]. There are also newer, smaller, non-obtrusive technologies currently being trialled (patches, smartphone/watch devices, hand-held devices) that may be of relevance to a screening programme. In this report, we consider only non-invasive tests for which information on sensitivity and specificity compared with a gold standard (12-lead ECG interpreted by at least one cardiologist) is available.

A screening programme would generate patients identified with an irregular pulse on the screening test used (e.g. pulse taking), some of whom may then need a confirmatory 12-lead ECG. There is some evidence that the sensitivity and specificity of diagnostic tests depend on the interpreter [nurse, general practitioner (GP), cardiologist or computer algorithm].[19,36] Therefore, a confirmatory ECG would add an additional burden to the NHS, either through training GPs to interpret them or through referral to cardiology units.

Long-term continuous Holter monitoring is considered appropriate to detect paroxysmal AF, although there is no accepted gold standard for the diagnosis of paroxysmal AF. This may change with the introduction of newer, non-invasive devices for home monitoring. Note, however, that a 12-lead ECG interpreted by a cardiologist would detect only the small proportion of patients with paroxysmal AF who happened to be in AF at the time of testing. A population-based screening programme with only short-term monitoring is unlikely to detect many paroxysmal AF cases.

The test should be acceptable to the population
The SAFE study reported an approximately 50% uptake rate for screening and, of those identified as having an irregular pulse, 73% had an ECG. In a post-screening questionnaire, only 3.7% reported that screening was not convenient.[19] No adverse results were reported.

There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals
We assumed that AF detected using a screening test would be confirmed using a 12-lead ECG, as set out in the NICE clinical guidelines.[11] The gold standard is a 12-lead ECG interpreted by a cardiologist; however, increasingly GPs are being trained to interpret ECGs, with referral to a cardiologist only in unclear cases. For the purposes of our economic model we assumed that, following a screening test, there would be a diagnostic 12-lead ECG interpreted by a GP with referral to a specialist to help with interpretation if necessary. Following a confirmed diagnosis, treatment and care options are set out in the NICE clinical guidelines.[11]

There should be an accepted treatment/therapy for patients identified through early detection, with evidence of early treatment/therapy leading to better outcomes than later treatment/therapy
As the morbidity and mortality associated with AF-related stroke are substantial, therapy to reduce AF-related stroke risk earlier rather than later is likely to lead to better outcomes; however, earlier treatment also implies more adverse events (bleeding events). The cost-effectiveness of anticoagulant therapy (warfarin or DOACs) to prevent stroke in eligible AF patients has been established,[37] indicating that it is an efficient use of NHS resources. However, the trials informing these assessments were not conducted in screen-detected cases. Mant *et al.*[38] found that screen-detected AF patients had a lower risk of stroke than diagnosed AF patients, but they found no evidence that the relative efficacy of warfarin compared with aspirin depends on whether AF is screen detected or not. However, the low event rates mean that this result is very uncertain, and the data were consistent with both a beneficial and a harmful effect of warfarin.

Other treatment options include rate control and rhythm control therapies. As screen-detected AF is likely to be asymptomatic, we would assume that a small proportion would be offered rate control until their heart rhythm disorder progresses to be symptomatic. Cardioversion may be offered to younger patients to avoid long-term anticoagulation; however, in an elderly population-based screening programme this would be a small proportion of screen-detected cases.

There should be agreed evidence-based policies covering which individuals should be offered therapy and the appropriate therapy to be offered
The NICE clinical guideline[11] clearly sets out who is eligible for anticoagulation and rate and rhythm control according to stroke and bleeding risk scores (see *Current practice in the diagnosis and treatment of atrial fibrillation*). Screen-detected patients would be eligible for anticoagulation only if the CHA_2DS_2-VASc score is ≥ 2, taking bleeding risk into account using the HAS-BLED score.

Clinical management of the condition and patient outcomes should be optimised in all health-care providers prior to participation in a screening programme

It has been estimated that around 35% of patients with diagnosed AF who are eligible and who do not have a contraindication are not receiving anticoagulation.[39] In addition, the introduction of DOACs has not led to uptake as expected. The potential therapy benefits from screening may, therefore, not be realised.

There should be evidence from high-quality randomised controlled trials that the screening programme is effective at reducing morbidity or mortality

There are currently no randomised controlled trials (RCTs) evaluating a population-based screening programme that have collected morbidity or mortality outcomes. The STROKESTOP trial[26,40] that is currently under way aims to fill this evidence gap, but is not due to report until 2019. RCT evidence on long-term outcomes is limited to the follow-up period available and so may not be able to fully answer this question. There is RCT evidence available on the prevalence of screen-detected cases under different screening strategies;[19] however, it is currently an assumption that early detection and treatment leads to morbidity and mortality benefits.

There should be evidence that the complete screening programme (test/diagnostic procedures/therapy/treatment/intervention) is clinically, socially and ethically acceptable to health professionals and the public

There have been several publications in the clinical literature calling for a national screening programme for AF,[16,41,42] including a consensus statement from the Royal College of Physicians of Edinburgh in 2012.[43] This suggests that such a programme would be acceptable to clinicians. A recent survey[44] of health-care professionals found that general practices have access to resources for AF screening and there is enthusiasm among non-GP health-care professionals to train and play a role in AF screening. However, there were concerns about a lack of staff, time and capacity.

Individuals invited to screening may choose to attend or not and so it is reasonable to assume that screening is acceptable in those that take up the screening opportunity. There may be a variety of reasons why some individuals do not attend for screening. It is important that this is an informed choice and that sufficient information is provided on the screening test and what positive and negative results mean, including the need for subsequent diagnostic tests and the treatment and therapy options that are available.

Patients diagnosed with AF are fearful of the consequences of a debilitating AF-related stroke. It is therefore important that they understand their risk of AF-related stroke and are able to make an informed choice about anticoagulation therapy after consulting with their health-care professionals. The options should be fully explained to them based on their eligibility and suitability and the risks and benefits of treatment, and should be reviewed periodically. NICE has produced Patient Decision Aids for this purpose.[45]

Warfarin needs to be dose adjusted with regular monitoring to ensure that the patient stays within the recommended therapeutic international normalised ratio (INR) range, making warfarin unacceptable to some patients.[46,47] Warfarin also has interactions with other medications including over-the-counter preparations, foods high in vitamin K and alcohol, which can impact on quality of life, meaning that some patients are reluctant to take warfarin. Recent NICE guidance DG14[48] has recommended coagulometers, devices that allow patients to self-monitor the INR, which is an alternative to regular INR monitoring, although some patients have experienced difficulties gaining access and support for these devices [Diane Eaton, AntiCoagulation Europe (ACE), 2016, personal communication]. DOACs do not require regular INR monitoring and have fewer drug and food interactions, making them more acceptable to many patients, although there are some concerns around lack of antidotes if there is a bleed (although dabigatran now has an antidote and others are in development).

Patients need to understand that, although anticoagulation therapy reduces the risk of AF-related stroke, it does not delay or reduce symptoms of AF. The importance of adhering to the medication in the absence of relief from symptoms therefore needs to be stressed repeatedly.

The proportion of eligible patients being prescribed DOACs is increasing, but is still relatively low.[39] If a population-based screening programme were to be put in place, the benefits of the programme would be realised only if screen-detected AF patients were put on appropriate anticoagulation therapy (e.g. not aspirin monotherapy, which is no longer recommended by NICE[11]). Therefore, there needs to be support at a local clinical commissioning group level to provide access to anticoagulation therapies recommended by NICE to eligible patients detected through screening.

The benefit from the screening programme should outweigh the physical and psychological harm (caused by test, diagnostic procedures and treatment)

The screening tests for AF investigated in this report are non-invasive and unlikely to cause physical harm. The SAFE study recorded the acceptability of screening to individuals and measured anxiety and quality of life before and after screening.[19] The authors found screening to be acceptable to the majority of individuals, but anxiety was higher and quality of life was lower in those who received a positive diagnosis of AF following screening. A 12-lead ECG, used to confirm the diagnosis of AF, is also non-invasive, although it requires the placement of pads on the upper body under clothing. The benefits in terms of reducing the risk of AF-related stroke are likely to outweigh adverse events (bleeding risk) and risk factors for bleeding on anticoagulant therapy are also well described by the HAS-BLED score.[11] However, as described earlier, there is not yet robust evidence of the impact of anticoagulation therapy in a screen-detected AF population.

The opportunity cost of the screening programme (including testing, diagnosis, therapy and treatment, administration, training and quality assurance) should represent value for money

Hobbs *et al.*[19] found that a systematic opportunistic screening programme is likely to be cost-effective in men and women from the age of 65 years, but their model assumed that benefits in screen-detected and non-screen-detected populations are the same and it was unclear whether or not lack of compliance to therapy, if indicated, was modelled. The cost-effectiveness of OACs in non-screen-detected AF has recently been demonstrated.[37]

A recent report for the National Screening Committee[24] concluded that there was a need to formally assess the cost-effectiveness of a screening programme for AF, and that such a cost-effectiveness analysis should include both costs of detecting undiagnosed cases of AF and also allow for stroke risk in screen-detected AF to be lower than that in clinically diagnosed AF.

Key questions arising from the World Health Organization criteria for the assessment of a screening programme

Many of the WHO criteria are clearly met, although some questions remain. Key inputs to assess the value of a screening programme are:

- the relative stroke risk for asymptomatic compared with symptomatic AF
- the proportion of these asymptomatic cases that would be eligible to benefit from anticoagulation therapy
- the split between AF type (paroxysmal, persistent, permanent) in screen-detected cases and those missed by screening (false negatives)
- the relative sensitivity and specificity of the range of possible screening tests
- the resources required for performing the screening tests and subsequent diagnostic tests
- the risk factor profiles of screen-detected AF cases
- the proportion of people with AF detected by screening who are eligible for anticoagulation who would actually receive it
- the impact of screening on long-term morbidity and mortality outcomes.

Research questions

Aim

The aim of this project was to assess the cost-effectiveness of different national screening strategies for AF in older adults. To achieve this aim, we conducted a series of reviews and developed an economic model.

Objectives

- To conduct a systematic review of diagnostic test accuracy (DTA) studies to determine:

  - the diagnostic accuracy of screening tests for detecting AF in adults ($\geq$ 18 years) who have not sought medical attention in a primary or community care setting on account of symptoms associated with AF
  - the diagnostic accuracy in systematic opportunistic, targeted and population screening settings.

- To update a previous systematic review[49] of screening strategies for AF to answer the following questions.

  - Does systematic screening increase the detection of AF compared with routine practice?
  - What are the characteristics of those identified with AF by screening strategy?
  - Which combination of screening strategy, screening population and test is the most effective at detecting AF compared with routine practice?
  - What are the potential safety issues and adverse events associated with individual screening programmes?
  - How acceptable is the intervention to the target population?
  - What are the costs associated with systematic screening for AF?

- To develop an economic model to compare the cost-effectiveness of different national screening strategies (including no screening) in a primary care setting in England and Wales, with the following objectives:

  - to review previous economic evaluations of screening for AF
  - to review recent literature on the prevalence, disease progression and risk profiles of AF and screening strategies relevant to a UK population-based screening setting
  - to evaluate the cost-effectiveness of different population-based screening strategies in England and Wales.

Chapter 2 Methods for the systematic reviews of diagnostic test accuracy studies and randomised controlled trials of screening for atrial fibrillation

Introduction

In this chapter we describe the methods for (1) the systematic review of DTA studies for the diagnosis of AF and (2) the systematic review of RCTs of screening for AF. The methods for the economic evaluation, incorporating a review of economic evaluations of mass screening strategies for AF and a review of natural history parameters relevant to a UK mass screening population, are provided in *Chapter 5*.

Diagnostic test accuracy systematic review

Objectives

The primary objective of this review was to determine the DTA of screening tests for detecting AF in adults (aged ≥ 18 years) who have not sought medical attention in a primary or community care setting on account of symptoms associated with AF.

The secondary objective of this review was to determine the DTA of index tests in various (i.e. systematic opportunistic, targeted and population) screening settings.

Review question

- *Study design*. Diagnostic cohort or case–control study.
- *Population*. People who had not sought medical attention on account of symptoms associated with AF.
- *Presentation*. Individuals registered in primary care and/or presenting to primary care or a community centre. Individuals may be invited to screening regardless of medical history (this may be done on the basis of age – systematic screening); on presenting to the GP for issues unrelated to AF symptoms (e.g. flu vaccination – systematic opportunistic screening); or based on their medical history/the presence of risk factors that are associated with AF (targeted screening).
- *Prior tests.* No prior testing for AF was required for the inclusion of individuals.
- *Index test.* Any non-invasive test that could be utilised in a primary care setting or in the community.
- *Purpose.* A screening test to identify people with AF who have not sought medical attention on account of symptoms associated with AF.
- *Target disorder.* AF.
- *Reference standard.* 12-lead ECG interpreted by a cardiologist.

Search strategy and selection criteria

MEDLINE, PreMEDLINE, EMBASE, The Cochrane Library, the Centre for Reviews and Dissemination (CRD) and the Science Citation Index were searched without language or date restrictions until January 2015. The search strategy included terms for atrial fibrillation and DTA or atrial fibrillation and diagnosis and the names of specific tests [pulse palpation, finger probe, ECG (single-lead, 12-lead, ambulatory, serial, continuous), Holter, cardiac event recorder, modified blood pressure monitor] (see *Appendix 1* for the search strategy run in MEDLINE). Grey literature sources were also searched [Google, National Guideline Clearing House, Guidelines International Network (GIN), Current Controlled Trials, WHO International Clinical Trials Registry Platform, ClinicalTrials.gov, UK Clinical Trials Gateway, OpenGrey, NICE website, King's Fund, Department of Health, National Institute for Health Research (NIHR) UK Clinical Research

Network (UKCRN), PROSPERO, European Society of Cardiology, American College of Cardiology, British Heart Foundation, Atrial Fibrillation Association, Arrhythmia Alliance] with search terms including AF and diagnosis. Additionally, the reference lists of identified systematic reviews and eligible studies were hand-searched to identify potentially relevant studies.

Titles and abstracts were screened independently by two reviewers (AM and GO) to identify potentially relevant studies. At title and abstract screening stage, studies were excluded if the target condition was not AF, the study was not a DTA study, the study was performed exclusively in people with diagnosed or treated AF, or if the index test was invasive, not possible in primary care or did not detect arrhythmia. Full texts of the remaining studies were then obtained and assessed for study eligibility using the full set of inclusion and exclusion criteria. This was also performed independently by two reviewers (AM and GO). Conflicts at each stage were resolved through discussion with a third reviewer (PD).

Inclusion/exclusion criteria

Study design/participants
We included cross-sectional, case–control, cohort studies, and RCTs that recruited at least 40 adults (≥ 18 years) who had not sought medical attention on account of symptoms associated with AF (with the exception of cases in case–control studies). Case–control studies had to recruit at least 40 adults (≥ 18 years). Case–control studies that had a control group made up exclusively of people with another diagnosed arrhythmia were excluded. Studies in stroke inpatients and outpatients, cardiology inpatients and outpatients, anticoagulant outpatients and intensive care patients were excluded (as case and non-case mix were likely to be different and/or these were inappropriate settings for screening) as were studies in patients with a diagnosis of AF who had had treatment such as ablation or cardioversion and studies in people with pacemakers/paced rhythms.

Tests
The index test must have been non-invasive and suitable for screening to detect arrhythmia and could be administered in a primary or community care setting and the presence or absence of AF must have been confirmed with a 12-lead ECG interpreted by a cardiologist (the reference standard). Although the index test must have been able to be performed in a community or primary care setting, index tests interpreted by a cardiologist were eligible because it is feasible that screening could take place in primary care but that the test results could be sent to a cardiologist for interpretation. Studies in which the reference standard was not a 12-lead ECG interpreted by a cardiologist were excluded. Studies that investigated invasive or echocardiographic methods of identifying AF were also excluded.

Setting
Studies were eligible for inclusion if they were performed in secondary or tertiary care settings, in addition to primary care or community care settings, as long as the case and non-case mix was judged to be likely to be similar to that seen during screening (e.g. outpatient day surgery) and the tests used could be administered in a primary or community care setting. The requirement for a study population similar to that seen during screening was to minimise spectrum bias.[50]

Outcomes
Only studies that reported the results per person and from which either diagnostic two-by-two contingency tables could be generated (i.e. studies that reported data from which the number of true positives, true negatives, false positives and false negatives could be extracted or calculated) or direct estimates of sensitivity (proportion of true positives) and specificity (proportion of true negatives) together with standard errors could be obtained were included. Studies in which the unit of analysis was not the person (e.g. the unit of analysis was the reading or the segment of a reading) and studies in which DTA information for AF only could not be extracted (e.g. if AF and atrial flutter were combined) were also excluded.

Data extraction

Data extraction was performed by one reviewer (AM) and checked by a second reviewer (GO). Any disagreements were resolved through discussion with a third reviewer (PD).

The following data were extracted:

- authors, publication year, journal
- study design
- characteristics of study participants including age, sex and ethnicity
- study inclusion and exclusion criteria
- setting
- prevalence of comorbidities (e.g. hypertension, diabetes, renal failure, heart failure, heart disease, previous stroke or MI) or risk of stroke CHA_2DS_2-VASc score
- prevalence of AF and prevalence of AF by type (paroxysmal, persistent, permanent)
- method of participant recruitment
- index test, including how and when the test was performed, the frequency of screening and the length of monitoring, and who performed the test and interpreted the test results
- definition of a positive index test result/cut-off value
- reference test, including definition for positive disease
- whether the readers of the index test and reference standard were blind to the results of the other test and other clinical information available to readers
- number of missing or unavailable test results
- numbers of true positives, true negatives, false positives and false negatives.

When studies reported the DTA of multiple thresholds for the same test or the DTA of more than one type of the same class of device, data for all thresholds/devices were extracted.

In studies in which three groups were included (cases with AF, control subjects with another arrhythmia and sinus rhythm/healthy control subjects), DTA information was extracted for AF cases compared with sinus rhythm/healthy control subjects when possible. When this was not possible (i.e. the analyses pooled the two control groups), data were extracted but the study was not included in the meta-analysis.

Study quality

Study quality was assessed using the QUality Assessment of Diagnostic Accuracy Studies – 2 (QUADAS-2) tool.[51] The QUADAS-2 tool consists of four key domains: patient selection, index test, reference standard and flow of patients through the study and timing of index tests, and reference standard (flow and timing). Each domain is assessed in terms of risk of bias, and the first three are assessed in terms of concerns regarding applicability. The tool was tailored to the review by modifying signalling questions, which help to inform judgements, and developing review-specific guidance. The signalling questions and guidance for this review are detailed in *Appendix 2*.

Statistical methods typically used to gauge risk of reporting bias may be misleading if applied to meta-analyses of test accuracy.[52] Therefore, reporting bias was not assessed statistically using methods such as funnel plots.

Statistical analysis

Data inputs for the statistical analysis

If results were reported for more than one threshold for a study, then we included data for each reported threshold in the statistical analysis. Some studies reported results broken down by interpreter. For example, in some studies the same set of readings was independently interpreted by different health-care professionals. These 'replications' provided information on the variability in sensitivity and specificity that is likely to be achieved in practice across different interpreters. Other studies set out to compare the sensitivity and

specificity depending on interpreter type (e.g. nurse, GP, cardiologist). We explored whether or not results for diagnostic accuracy depended on the interpreter by further classifying the tests according to whether they were interpreted by a nurse, a GP or a cardiologist, or automatically/using an algorithm.

Descriptive analyses

Summary sensitivity and specificity values for each index test (and threshold) compared with the reference standard were calculated together with confidence intervals (CIs) and were displayed using forest plots.

Evidence synthesis

Formal evidence synthesis was conducted by fitting a hierarchical summary receiver operating characteristic (HSROC) model[53] to estimate the relationship between sensitivity and specificity of the index tests. The HSROC model estimates a summary receiver operating characteristic (SROC) curve allowing for correlations between sensitivity and specificity and heterogeneity between studies. The SROC curve depends on the threshold (cut-off point) for the test, the accuracy of the test and the shape of the curve (*Figure 1*).

We attempted to fit both symmetrical and asymmetrical HSROC models; however, there were too few data to identify the shape parameter (regardless of whether this was constant, exchangeable or independent across studies) and so results for symmetrical HSROC models only are reported. We adapted the HSROC model to allow an extra level of variability between interpreters in studies with replicates; however, we again found that there were too few data to estimate this extra level of variability. We therefore reported results from a model that does not estimate between-interpreter variability. We allowed for multiple thresholds by treating these as independent observations but with the same study-specific accuracy. We fitted two models, one with and one without classifying the tests according to the interpreter.

The models were fitted using a Bayesian approach computed in OpenBUGS version 3.2.3 rev 1012[54] (MRC Biostatistics Unit, Cambridge, UK). Model fit was assessed using the posterior mean deviance,[55] and models were compared using the deviance information criterion (DIC).[55] Full details of the model fitted, including computer code, are provided in *Appendix 3*.

The results of the HSROC models were presented as SROC curves with confidence regions and prediction regions for sensitivity and specificity.
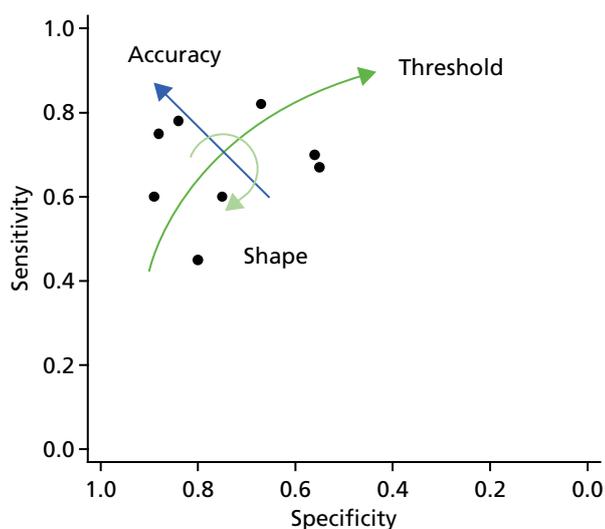
FIGURE 1 Illustration of the SROC curve.

Investigations of heterogeneity and subgroup and sensitivity analyses

We inspected between-study variance for evidence of heterogeneity and compared the fit of a HSROC curve with that of a fixed-effects SROC curve. When sufficient evidence was available, we explored possible explanations for heterogeneity through prespecified subgroup analyses: study design, study year, setting, method of recruitment into the study (and whether it would mimic a targeted, systematic opportunistic or systematic screening programme), age of participants, prevalence of AF (and AF subtypes), prevalence of comorbidities and stroke risk score, frequency of screening (if multiple moment-in-time screening tests were identified), length of monitoring (if ambulatory tests were identified) and cut-off value. We also conducted sensitivity analysis for risk-of-bias indicators according to the QUADAS-2 domains.

Systematic review of randomised controlled trials comparing screening strategies

Objectives

The objective of this systematic review was to update the Cochrane review of screening strategies for AF.[49] During the course of our review we identified that the Health Information and Quality Authority (HIQA) in Ireland had published a health technology assessment (HTA) of a national screening programme for AF in primary care.[25] The HIQA HTA project[25] included a review of the effectiveness of screening, which updated the Cochrane review[49] with the results of a literature search covering June 2012 to June 2015. In addition, we were also informed that the Cochrane review itself was being updated (currently undergoing peer review, Patrick Moran, Health and Quality Association, Dublin, Ireland, 2015, personal communication) and, with the assistance of the review authors, we updated this version of the review while expanding the scope to include head-to-head trials and trials that had not confirmed the presence or absence of AF with a 12-lead ECG interpreted by a GP, a specialist or a suitably trained ECG technician.

Review questions

Our review questions were as follows.

- Does systematic screening increase the detection of AF compared with current routine practice?
- What are the characteristics of those identified with AF by screening strategy?
- Which combination of screening strategy, screening population and test is the most effective at detecting AF compared with routine practice?
- What are the potential safety issues and adverse events associated with individual screening programmes?
- How acceptable is the intervention to the target population?
- What are the costs associated with systematic screening for AF?

The protocol for the systematic review was registered with the NIHR international prospective register of scientific reviews (PROSPERO) prior to executing the literature search strategy (registration no. CRD42014013739).[56] In line with the Cochrane review, which we aimed to update, our PICO (population, intervention, comparator, outcome) framework was as follows.

Study design

We had originally planned to include all of the study designs that were considered by the original Cochrane review (RCTs, cluster RCTs, controlled before-and-after studies and interrupted time series). However, the Cochrane update was restricted to RCTs only (Patrick Moran, personal communication). We therefore changed our inclusion criteria to focus solely on RCTs.

Population

The population of interest was adults aged ≥ 40 years of either sex. This age cut-off point was chosen based on epidemiological data which suggest that AF is extremely uncommon below the age of 40 years.[57] We restricted the population to those in whom screening could occur (and therefore, for example, excluded testing/monitoring for AF in stroke patients.

Interventions and comparator

The interventions of interest were population-based, systematic (opportunistic, targeted or population) screening programmes for AF (see *Chapter 1*, *Systematic screening strategies*). In contrast to the Cochrane review, in which studies that compared one or more screening strategies with no screening (routine practice) were included, we additionally included studies that compared two or more screening strategies (without a 'no screening' arm).

Our primary interest was in screening strategies conducted within a primary care setting, but we did not exclude studies on the basis of setting as long as the population screened was generalisable to the general population.

Primary outcome

The primary outcome was the difference in the detection of new AF cases associated with screening compared with usual practice. To be eligible for the Cochrane review, cases of AF had to be confirmed using a 12-lead or continuous ambulatory ECG interpreted by a GP, a specialist or a suitably trained ECG technician. However, because of a lack of studies, we expanded the scope and did not exclude studies on the basis of the method used to confirm AF.

Secondary outcomes

As in the original Cochrane review,[49] other outcomes of interest, including for our economic evaluation, were:

- change in diagnosed AF (after screening compared with before screening)
- the acceptability of systematic screening programmes
- adverse events associated with systematic screening
- costs associated with systematic screening programmes for AF.

In addition, we were also interested in the patient characteristics of those detected with AF, for example age, sex, clinical history, CHA_2DS_2-VASc score and AF type.

Search strategy and selection criteria

The Cochrane review team confirmed that all head-to-head RCTs of systematic screening would have been identified by their search strategy and would have been examined at full text, although those without a no screening control arm or without AF diagnosis confirmed by a 12-lead or continuous ambulatory ECG interpreted by a GP, specialist or suitably trained ECG technician would then have been excluded. The Cochrane review team kindly provided us with a list of excluded studies that they thought may be relevant to this review. We examined this list to identify studies meeting our inclusion criteria. In addition, we updated the update by searching MEDLINE and PreMEDLINE, EMBASE, The Cochrane Library's CENTRAL and CINAHL from July 2015 to December 2015.

An information specialist searched health-care databases and other literature sources to identify published and unpublished literature on human subjects, in any language. Our searches were developed for MEDLINE and then adapted for other literature sources when necessary (see *Appendix 4* for details). Reference and citation tracking were undertaken to identify further relevant studies. When necessary, we contacted lead authors for more information on published and unpublished studies that might be relevant. We also sought information on studies in progress.

After removal of duplicates, articles were screened in parallel by two reviewers (GO, AM) applying our eligibility criteria and using a two-stage sifting approach to review article title and abstract and then full text. As the Cochrane review and the ongoing update review excluded comparative studies that did not have a no screening arm (control), two reviewers (GO, AM) reassessed the list of excluded studies provided by the Cochrane review authors for inclusion. In all cases, disagreements were discussed between the two reviewers and, if not resolved, resolution was sought through the involvement of a third reviewer (PD).

Data extraction

We developed data extraction forms and piloted them on two studies. Data were extracted from included studies according to the PICO framework using the piloted extraction forms. Data extraction was conducted by one reviewer and then reviewed by a second reviewer. Disagreements were discussed between the two reviewers and, if not resolved, resolution was sought through the involvement of a third reviewer. We extracted data on the following:

- study details (identifier, study design, location, year, length of follow-up, study funder)
- participant details (number of participants, age, sex)
- intervention details (population screening type, screening test, method of AF diagnosis, timing)
- comparator details
- details relevant to the risk-of-bias assessment
- effect modifiers (demography of the study population, study settings).

Assessment of risk of bias

Risk-of-bias assessments were conducted by one reviewer and then reviewed by a second reviewer. Disagreements were discussed between the two reviewers and, if not resolved, resolution was sought through involvement of a third reviewer. Risk of bias was assessed using the Cochrane tool for assessing risk of bias in randomised trials.[58] This tool assigns the judgement of low, high or unclear risk of bias for each of the following sources of bias: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data and selective reporting. Additional items for assessing the risk of bias in cluster randomised trials were derived from the Cochrane-suggested risk-of-bias assessment tool for Effective Practice and Organisation of Care (EPOC) reviews.[59]

Abstracts that were assessed to meet the study eligibility criteria for inclusion were subjected to data extraction only and not to assessment for risk of bias because of a lack of information to enable judgement on some of the criteria.

We planned to use funnel plots to assess reporting bias; however, insufficient numbers of studies were identified for this to be possible.

Data analysis

We have tabulated study-specific characteristics, outcome measures and risk-of-bias assessments, and we provide a narrative overview of the results from the included studies.

We report results for binary outcomes (e.g. proportion of new AF diagnoses) only for studies in which there is a clear denominator. We report results for the proportion of new AF diagnoses among those without a previous diagnosis of AF and summarise relative effects between study arms using the odds ratio (OR). The number needed to screen (NNS) to detect an additional newly diagnosed AF case was calculated as the reciprocal of the risk difference (difference in proportion of new cases detected between study arms). We computed the prevalence of previously diagnosed AF prior to screening by pooling across arms within a study. We calculated the prevalence of AF diagnoses before and after screening to compute the percentage change in prevalence as a result of screening. We examined the acceptability of systematic screening programmes within the screened population in three ways: from the level of uptake achieved, feedback elicited from the participants and health professionals involved and a description of any direct costs associated with screening that were borne by the person to whom the screening programme was offered. Patient characteristics of those diagnosed with AF (age, sex, clinical history, CHA_2DS_2-VASc score, AF type) are described when reported. We report any adverse events associated with screening reported in the studies. We also report information on the incremental cost to the NHS per additional newly diagnosed AF case identified for a given screening strategy compared with routine care when this was reported in RCTs conducted in the UK. We planned the following subgroup analyses when enough data were available: aged ≥ 65 years compared with 65–75 years compared with > 75 years; men compared with

women; different ethnic and socioeconomic groups; and community compared with specialist settings. We also planned a sensitivity analysis excluding trials with a high risk of bias if there were enough data.

Meta-analysis was conducted when feasible (enough data on outcomes defined in the same way) to obtain pooled ORs with 95% CIs. We did not anticipate there being a sufficient number of studies to estimate random-effects meta-analysis models and so reported results from fixed-effects meta-analysis only.

Chapter 3 Results: diagnostic test accuracy review

Study selection

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart for the review is shown in *Figure 2*. After the removal of duplicate records, we identified 4084 potential citations, of which 3809 studies were excluded at title and abstract level as it was clear that the target condition was not AF, the study was not a DTA study, the study was performed exclusively in people with diagnosed or treated AF, or that the index test was invasive, not possible in primary care or did not detect arrhythmia. After a full-text review of the remaining 270 studies, 253 studies were excluded because they did not meet the full set of inclusion criteria. The list of studies excluded and the reasons for their exclusion are detailed in *Appendix 5*. Although there was consensus between reviewers that these studies should be excluded, there were often multiple reasons for exclusion and consensus was not always obtained. Reasons for exclusion were that the study was not a DTA study; that the target condition was not AF (this included studies in which DTA data for AF could not be extracted, for example when AF and atrial flutter were combined); that the information was not per person (i.e. the unit of analysis was not the person, for example the unit of analysis was the reading or the segment of a reading); population reasons (people
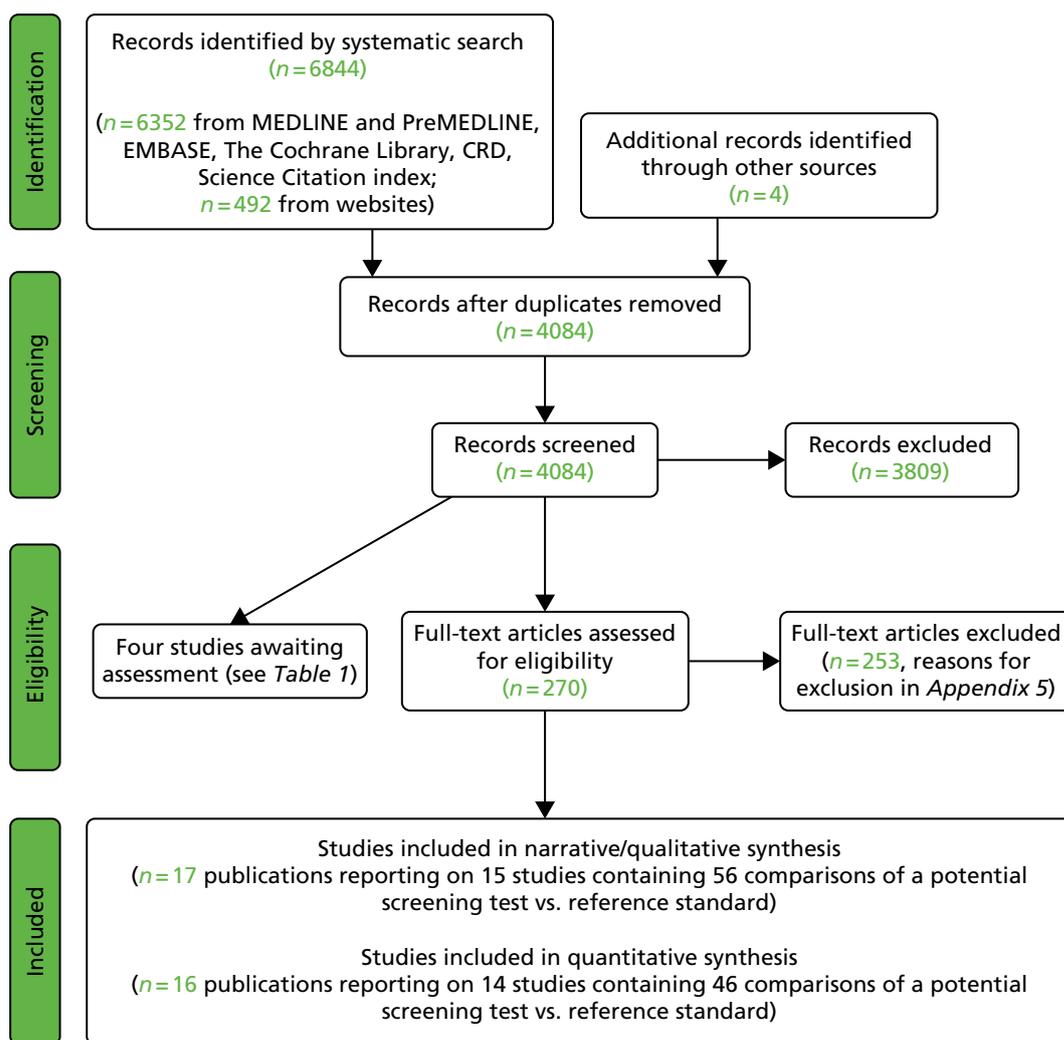
FIGURE 2 The PRISMA flow chart.

with a diagnosis of AF who had had curative treatment such as ablation or cardioversion, people with pacemakers/paced rhythms, stroke inpatients and outpatients, cardiology inpatients and outpatients, anticoagulant outpatients and patients in intensive care, and studies with < 40 participants); index test reasons (invasive, not possible in primary care or does not detect arrhythmia); reference standard (i.e. not a 12-lead ECG interpreted by a cardiologist).

Seventeen publications[19,36,60–74] (15 studies) containing 56 comparisons of an index test with the reference standard were found to meet the full set of eligibility criteria and were included in the review. One publication contained two eligible studies.[60] The protocol and results of one study were reported in three publications.[19,36,61] One study also had an erratum published.[62,63]

In addition, four potentially relevant DTA studies were identified (*Table 1*); these trials are awaiting assessment against our eligibility criteria as study reports and results were not identified.

Study characteristics

The characteristics of the included studies are summarised in *Table 2*.

Study design
There were 10 studies with a single set of inclusion criteria (cohort or one-gate studies),[19,60,63–69] three studies with two sets of inclusion criteria (case–control or two-gate studies)[70–72] and two studies of unclear design.[73,74] One of the case–control studies grouped participants into those with known sustained AF, those with other non-AF arrhythmias and those in sinus rhythm.[71]

Population and setting
Only four studies were performed in primary care.[19,67,70,72] The other studies were performed in outpatient settings (hypertension outpatients,[63] presurgery[64]), secondary care or higher settings[60,65,66,68,69] or a mixture of outpatient and secondary care or higher settings [outpatient clinics including cardiology outpatients and the preadmission clinic (ECGs carried out prior to surgery in patients aged > 40 years);[74] outpatient and inpatient settings;[60] outpatient hypertension clinic, patients admitted to University Department of Medicine wards and healthy volunteers[71]]. In one study the setting was unclear (a computerised electrocardiographic system).[73]

A secondary objective of this review was to determine the DTA of index tests in systematic (opportunistic, targeted and population) screening settings. To do this we had to assess whether the populations studied represented the populations who would attend if a population-based systematic screening programme was in place. In practice, this was difficult to achieve as it was unclear how the study populations were selected and recruited. We were able to categorise only four studies, three of which represented population

TABLE 1 Details of potentially relevant DTA studies that are either ongoing or for which no publication was identified

Clinical trial identifier	Title	Status
NCT02262351	Programme for the Identification of 'Actionable' Atrial Fibrillation in the Family Practice Setting (PIAAF-FP)	Completed
NCT02270112	Enhanced Diagnostics for Early Detection of Atrial Fibrillation (DETECT-AF)	Completed
NCT02124629 (now NCT02401451)	Six Lead Identification of Atrial Fibrillation (SL-AF)	Completed
NCT02162394	A Study to Determine the Feasibility of Wireless Echocardiography	Completed

TABLE 2 Characteristics of included studies

Study	Study design[a]	Country; setting,[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification;[c] test description	Interpreter; cut-off point (if multiple cut-off points investigated)
Antonicelli 2012[64]	Cohort/one gate	Italy; outpatients; 107; patients were enrolled from the presurgical evaluation unit in the outpatient day surgery service at the Italian National Research Center on Aging in Ancona during a 1-month period	NR	0.93 (0 to 4.3)[d] [1/107]	Mean age ± SD (range) (years): 66 ± NR (60–72); sex: 53.3% male	> 1- and < 12-lead ECG; 3-lead tele-ECG (CG-7100, Card Guard Scientific Survival Ltd)	Cardiologist (cardiologist 1)
						> 1- and < 12-lead ECG; 3-lead tele-ECG (CG-7100, Card Guard Scientific Survival Ltd)	Cardiologist (cardiologist 2)
						12-lead ECG; portable 12-lead ECG recorder (CG-7000DX-BT, Card Guard Scientific Survival Ltd)	Cardiologist (cardiologist 1)
						12-lead ECG; portable 12-lead ECG recorder (CG-7000DX-BT, Card Guard Scientific Survival Ltd)	Cardiologist (cardiologist 2)
Gregg 2008[65]	Cohort/one gate	USA; secondary, tertiary or quaternary care; 1785; ECG database collected consecutively from two teaching hospitals (one ECG per person)	Extreme artefact and paced rhythms excluded	6.11 (5 to 7.22) [109/1785]	Unclear[e]	12-lead ECG	Automatic/algorithm (the Philips 12-lead diagnostic algorithm)
						> 1- and < 12-lead ECG; reconstructed 12-lead ECG from V2 and V5	Automatic/algorithm (the Philips 12-lead diagnostic algorithm)
						> 1- and < 12-lead ECG; reconstructed 12-lead ECG from V1 and V4	Automatic/algorithm (the Philips 12-lead diagnostic algorithm)

TABLE 2 Characteristics of included studies (*continued*)

Study	Study design[a]	Country; setting;[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
Hobbs 2005,[19] Mant 2007,[36] Swancutt 2004[61]	Cohort/one gate	UK; primary care; 2556 (12-lead ECG interpreted by interpretative software), 2578 (pulse palpation). This was a prospective substudy of the SAFE study, in which 2595 people aged ≥ 65 years were screened for AF. A random sample of people aged ≥ 65 years from the 25 intervention practices was invited for screening and another random sample was allocated to opportunistic screening, during which they received an ECG if they were identified as having an irregular pulse	Patients who were terminally ill	8.41 (7.34 to 9.49) [215/2556] (12-lead ECG interpreted by interpretative software); 8.46 (7.38 to 9.53) [218/2578] (pulse palpation)	Mean age ± SD (range) (years): 73.5 ± 6.02 (65–98); sex: 46.9% male; ethnicity: white 95.5%, black 3.2%, Hispanic 1.2%	12-lead ECG Pulse palpation	Automatic/algorithm (interpretative software) Health professional (practice nurse)
Kaleschke 2009[66]	Cohort/one gate	Germany; secondary, tertiary or quaternary care; 505; patients with a clinical indication for 12-lead surface ECG recording were consecutively enrolled in the AFNET (German Competence Network on Atrial Fibrillation) centres at the University Hospitals of Hamburg, Magdeburg, Munich and Munster from July 2007 to February 2008	Aged < 18 years; presence of a pacemaker or implantable defibrillator	28.32 (24.39 to 32.25) [143/505]	Mean age ± SD (range) (years): 61.4 ± 14.5 (18–96); sex: 66% male[f]	Single-lead ECG; Omron HeartScan, single-channel ECG	Cardiologist (first author of the study)

Study	Study design[a]	Country; setting;[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
Kearley 2014[67]	Cohort/one gate	UK; primary care; 999;[g] patients aged ≥ 75 years, living at home, from six general practices in the UK were invited to take part until the sample size had been achieved (i.e. 1000 participants had been recruited)	Implanted pacemakers or defibrillators; unable to give informed consent; GP considered participation inappropriate (e.g. terminal illness)	7.91 (6.23 to 9.58) [79/999][g]	Mean age ± SD (range) (years): 79.7 ± NR (75.1–99.8); sex: 49.3% male; hypertension 53.3%, diabetes mellitus 12.2%, heart failure/congestive heart failure 3.1%, previous stroke 3.1%, previous TIA 6.5%, AF (any/not specified) 11%, receiving antiarrhythmic drugs 8.7%[h]	Modified blood pressure monitor; WatchBP, a modified oscillometric blood pressure monitor (administered by a nurse)	Automatic/algorithm
						Single-lead ECG; Omron (model HCG-801)	Automatic/algorithm
						Two-stage screening; modified blood pressure monitor (WatchBP) and then single-lead ECG with autoanalysis (Omron) if positive	Automatic/algorithm
						Single-lead ECG; Omron (model HCG-801)[i]	Cardiologist (meta-analysis of four cardiologists' interpretations)
						Single-lead ECG; Omron (model HCG-801)[i]	Cardiologist (cardiologist 1)
						Single-lead ECG; Omron (model HCG-801)[i]	Cardiologist (cardiologist 2)
						Single-lead ECG; Omron (model HCG-801)[i]	Cardiologist (cardiologist 3)
						Single-lead ECG; Omron (model HCG-801)[i]	Cardiologist (cardiologist 4)
						Single-lead ECG; Merlin ECG event recorder[i]	Cardiologist (meta-analysis of four cardiologists' interpretations)
						Single-lead ECG; Merlin ECG event recorder[i]	Cardiologist (cardiologist 1)

TABLE 2 Characteristics of included studies (*continued*)

Study	Study design[a]	Country; setting,[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
						Single-lead ECG; Merlin ECG event recorder[i]	Cardiologist (cardiologist 2)
						Single-lead ECG; Merlin ECG event recorder[i]	Cardiologist (cardiologist 3)
						Single-lead ECG; Merlin ECG event recorder[i]	Cardiologist (cardiologist 4)
Langley 2012[68]	Cohort/one gate	UK; secondary, tertiary or quaternary care; 167; ECG recordings from 167 subjects with a range of cardiac rhythms collected in the Freeman Hospital, Newcastle upon Tyne, UK. ECGs were grouped according to type of rhythm: AF (*n* = 55), sinus rhythm (*n* = 72), sinus rhythm with ectopic beats (*n* = 27) and other rhythms (*n* = 13) including paced rhythms and atrial flutter	NR	32.93 (25.81 to 40.06) [55/167]	NR	12-lead ECG	Automatic/algorithm; coefficient of variation – quantifies beat interval variability as the SD of all analysed beats. Cut-off value of 0.12, 10 seconds of ECG
						12-lead ECG	Automatic/algorithm; mean successive beat interval difference – quantifies the beat-to-beat variability. Cut-off value of 0.11, 10 seconds of ECG
						12-lead ECG	Automatic/algorithm; entropy-based algorithm – quantifies the regularity of beat interval patterns. Cut-off value of –1.19, 10 seconds of ECG

Study	Study design[a]	Country; setting;[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
Lau 2013[74]	Unsure	Australia; mixed; 204; patients in the validation set were recruited from outpatient clinics including cardiology outpatients and the preadmission clinic (ECGs carried out prior to surgery in patients aged > 40 years). The aim was to recruit between one-quarter and one-third who were in AF at the time of recording	Pacemaker/paced rhythm	23.53 (17.71 to 29.35) [48/204]	NR	Single-lead ECG; iPhone (Apple Inc., Cupertino, CA, USA) ECG (AliveCor)	Automatic/algorithm (optimised algorithm)
Lee 2007[60]	Cohort/one gate	South Korea; mixed; 157; 157 clinical cases [32 AF cases and 125 cases sinus (normal) rhythm and other rhythm] from the outpatient and inpatient departments of Chonnam University Hospital were recorded simultaneously using two different ECG machines	NR	20.38 (14.08 to 26.68) [32/157]	NR	12-lead ECG; 12-lead ECG (Bionet EKG3000)	Automatic/algorithm
						12-lead ECG; 12-lead ECG (GE Marquette system MAC5000)	Automatic/algorithm
Lee 2007[60]	Cohort/one gate	South Korea; secondary, tertiary or quaternary care; 1270; 1270 consecutive cases including 149 cases (11.7%) with AF from the Chonnam University Hospital using Bionet EKG3000	NR	11.73 (9.96 to 13.5) [149/1270]	NR	12-lead ECG; 12-lead ECG (Bionet EKG3000)	Automatic/algorithm
Lewis 2011[69]	Cohort/one gate	UK (Wales); secondary, tertiary or quaternary care; 592; unclear/not reported: patients, aged > 60 years attending hospital outpatient departments or inpatients in two hospitals	None	16.05 (13.09 to 19) [95/592]	NR	Photoplethysmography; AFS instrument that uses plethysmographic analysis of finger-tip pulsation to detect AF	Automatic/algorithm; cut-off value of 0.30
						Photoplethysmography; AFS instrument that uses plethysmographic analysis of finger-tip pulsation to detect AF	Automatic/algorithm; cut-off value of 0.25

TABLE 2 Characteristics of included studies *(continued)*

Study	Study design[a]	Country; setting;[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
						Photoplethysmography; AFS instrument that uses plethysmographic analysis of finger-tip pulsation to detect AF	Automatic/algorithm; cut-off value of 0.20
Marazzi 2014,[62] 2012[63]	Cohort/one gate	Italy; outpatients; 503; consecutive patients referred to a hypertension clinic	Aged < 18 years; the presence of a pacemaker and/or an implanted defibrillator	20.08 (16.58 to 23.58) [101/503]	Mean age ± SD (range) (years): 67 ± 10.5 (NR); sex: 54.27% male	Modified blood pressure monitor; Microlife BP A200 Plus (automated oscillometric device for the self-monitoring of blood pressure at arm level)	Automatic/algorithm
						Modified blood pressure monitor; OMRON M6 (automatic device for the self-measurement of blood pressure at arm level using the oscillometric method)	Automatic/algorithm
Slocum 1992[73]	Unsure	USA; unclear; 107; 12-lead ECGs from a computerised electrocardiographic system. The database consisted of 221 rhythms that were separated into three groups: a group of rhythms with AF, a control group and a group of sinus rhythms. The group of rhythms with AF consisted of 74 rhythms taken from consecutive patients in AF. The group of sinus rhythms consisted of 66 ECGs taken from consecutive patients in sinus rhythm. The group of rhythms with AF were divided into a training set and a test set	NR	38.32 (29.11 to 47.53) [41/107]	NR[i]	12-lead ECG	Automatic/algorithm (discrimination algorithm)

24

Study	Study design[a]	Country; setting;[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
Somerville 2000[70]	Case–control/two gate	UK; primary care; 86;[k] Participants were 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'	NR	30.23 (20.53 to 39.94) [26/86][k]	NR	Pulse palpation	Health professional (nurse A, who had a background in both community and accident and emergency nursing and who had experience of taking and interpreting ECGs)
							Health professional (nurse B, a practice nurse with no additional ECG training)
						Single-lead ECG; bipolar ECG	Health professional (nurse A, who had a background in both community and accident and emergency nursing and who had experience of taking and interpreting ECGs
							Health professional (nurse B, a practice nurse with no additional ECG training)
						Single-lead ECG; bipolar ECG	Health professional (nurse C, a practice nurse who had worked on a coronary care unit and who had been trained to interpret ECGs)
					Pulse palpation	Single-lead ECG; bipolar ECG	

continued

TABLE 2 Characteristics of included studies (*continued*)

Study	Study design[a]	Country; setting,[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
						Single-lead ECG; bipolar ECG	Clinician (GP)
						12-lead ECG	Health professional (nurse A, who had a background in both community and accident and emergency nursing and who had experience of taking and interpreting ECGs)
						12-lead ECG	Health professional (nurse B, a practice nurse with no additional ECG training)
						12-lead ECG	Clinician (GP)
						Two-stage screening; pulse palpation assessed by a nurse (A) followed by bipolar ECG interpreted by a nurse (A) for those who screened positive	Health professional (interpreted both pulse palpation and ECG) (nurse A)
						Two-stage screening; pulse palpation assessed by a nurse (A) followed by bipolar ECG interpreted by a GP for those who screened positive	Health professional and clinician (nurse A performed pulse palpation and GP interpreted ECG)
						Two-stage screening; pulse palpation assessed by a nurse (A) followed by 12-lead ECG interpreted by a nurse (A) for those who screened positive	Health professional (interpreted both pulse palpation and ECG) (nurse A)

Study	Study design[a]	Country; setting;[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
						Two-stage screening; pulse palpation assessed by a nurse (A) followed by 12-lead ECG interpreted by a GP for those who screened positive	Health professional and clinician (nurse A performed pulse palpation and GP interpreted ECG)
						Two-stage screening; pulse palpation assessed by a nurse (B) followed by bipolar ECG interpreted by a GP for those who screened positive	Health professional and clinician (nurse B performed pulse palpation and GP interpreted ECG)
						Two-stage screening; pulse palpation assessed by a nurse (B) followed by 12-lead ECG interpreted by a GP for those who screened positive	Health professional and clinician (nurse B performed pulse palpation and GP interpreted ECG)
Stergiou 2009[71]	Case–control/two gate	Greece; mixed; 72; subjects with known sustained AF or other non-AF arrhythmias and controls with sinus rhythm were recruited among those attending an outpatient hypertension clinic, patients admitted to University Department of Medicine wards and healthy volunteers	Aged < 35 years; presence of a pacemaker, and/or an implanted defibrillator; refusal to participate	37.50 (26.32 to 48.68) [27/72]	Mean age ± SD (range) (years): 70.5 ± 10.6 (NR) [cases 75.7 ± 6.3 (NR)]; sex: 65.8% male (cases 67.9% male); hypertension: 63% (cases 60.7%); diabetes mellitus: 15.1% (cases 25%)	Modified blood pressure monitor; Microlife BPA100 Plus, an automated oscillometric device for self-home blood pressure monitoring	Automatic/algorithm; cut-off value: positive on the first reading (measurements had been taken in triplicate)
						Modified blood pressure monitor; Microlife BPA100 Plus, an automated oscillometric device for self-home blood pressure monitoring	Automatic/algorithm; cut-off value: positive on either of the first two readings (measurements had been taken in triplicate)
						Modified blood pressure monitor; Microlife BPA100 Plus, an automated oscillometric device for self-home blood pressure monitoring	Automatic/algorithm; cut-off value: positive on any of the three readings (measurements had been taken in triplicate)

continued

TABLE 2 Characteristics of included studies (*continued*)

Study	Study design[a]	Country; setting,[b] number in analysis; selection and recruitment details	Exclusion criteria	AF prevalence (95% CI) (%) [proportion]	Participant characteristics	Test classification; test description[c]	Interpreter; cut-off point (if multiple cut-off points investigated)
						Modified blood pressure monitor; Microlife BPA100 Plus, an automated oscillometric device for self-home blood pressure monitoring	Automatic/algorithm; cut-off value: positive on two of the three readings (measurements had been taken in triplicate)
Vaes 2014[72]	Case–control/two gate	Belgium; primary care; 181; GPs invited patients with diagnosed paroxysmal or chronic AF and patients without a history of AF to participate in the study. 'To end up with a prevalence of atrial fibrillation of at least 50% at the moment of the study the vast majority of the invited patients were people with a known history of atrial fibrillation (161/191) and only 30 people without a history of atrial fibrillation'	Pacemaker (configured in active pacing mode)	53.04 (45.77 to 60.31) [96/181]	Mean age ± SD (range) (years): 74.6 ± 9.7 (50–99) [cases 77 ± 8 (NR), controls 71 ± 11 (NR); sex: 52.36% male [cases 55.3% male, controls 51.1% male]; hypertension: cases 91.3%, controls 70.5%; diabetes mellitus: cases 25.2%, controls 17%; coronary artery disease: cases 8.7%, controls 18.2%; previous stroke: cases 16.5%, controls 4.5%; chronic AF: 84.29%; median CHA$_2$DS$_2$ score: cases 3 (IQR 2–4), controls 3 (IQR 2–3); receiving anticoagulants: cases 100%, controls 75%[m]	Single-lead ECG; MyDiagnostick (three readings were taken)	Automatic/algorithm

AFS, atrial fibrillation screening; IQR, interquartile range; NR, not reported; SD, standard deviation.

a   Study design was divided into cohort/one gate and case–control/two gate.

b   Setting was classified into community; primary care; outpatients; secondary, tertiary, quaternary care; mixed; and unclear.

c   Tests were classified as pulse palpation; photoplethysmography; modified blood pressure monitor; single-lead ECG; > 1- and < 12-lead ECG; 12-lead ECG; ambulatory monitoring; and two-stage screening.

d   Estimated using Bayesian methods with a Jeffreys prior for a proportion, to allow for low counts.

e   Some patient demographics reported but the number of males and females reported sums to more than the number of people in the study.

f   Also reports mean ± SD body mass index (26.6 ± 4.3 kg/m²).

g   Number of participants included in each analysis varied and the prevalence also varied slightly.

h   Percentage receiving antiarrhythmic drugs actually participants with AF on antiarrhythmic medication.

i   The one-lead ECGs were interpreted by four cardiologists, with two separate pairs of cardiologists reading approximately half of the ECGs each. Results are presented for each cardiologist separately in addition to the meta-analysis of the four cardiologists' interpretations.

j   The number of people with AF on their ECG is known.

k   Some analyses were performed only on a subset of patients. The number of participants included in each analysis varied and prevalence also varied slightly.

l   Data for participants in sinus rhythm (healthy controls) and patients with other arrhythmias reported separately, but results for this paper combine the other arrhythmias and sinus rhythm groups. Also reports other patient characteristics. Whole sample (mean ± SD): body mass index 27.0 ± 4.3 kg/m²; arm circumference (cm) 28.2 ± 3.4; smokers 5.5%; number of drugs for hypertension 1.8 ± 1.5; systolic blood pressure 138 ± 19 mmHg; diastolic blood pressure 80 ± 12 mmHg; pulse rate 74 ± 17 beats per minute; ECG heart rate 78 ± 21 beats per minute; cardiovascular disease 39.7%. Cases of AF only: body mass index 26.3 ± 3.3 kg/m²; arm circumference (cm) 28.0 ± 3.3; smokers 7.1%; number of drugs for hypertension 1.8 ± 1.6; systolic blood pressure 141 ± 21 mmHg; diastolic blood pressure 79 ± 14 mmHg; pulse rate 78 ± 16 beats per minute; ECG heart rate 84 ± 24 beats per minute; cardiovascular disease 60.7%.

m   Also reports other patient characteristics(mean ± SD): body mass index: cases 27.7 ± 6.9 kg/m², controls 25.9 ± 6.5 kg/m²; systolic blood pressure: cases 130 ± 19 mmHg, controls 130 ± 18 mmHg; diastolic blood pressure: cases 75 ± 14 mmHg, controls 75 ± 10 mmHg. Note that percentage with hypertension is actually more specifically percentage with arterial hypertension; percentage with diabetes is actually percentage with diabetes mellitus type II; percentage with previous stroke is actually percentage with TIA or cerebrovascular accident. The prevalence of peripheral arterial illness was also reported: 3.9% in cases and 4.5% in controls. The percentage receiving anticoagulants was further subdivided: cases: 5.8% new OACs, 67% warfarin and 27.2% new OACs; controls: 27.3% platelet aggregation inhibitors, 34.1% warfarin and 13.6% new OACs.

systematic screening (one recruited individuals aged ≥ 75 years from primary care,[67] one case–control study recruited cases and controls aged ≥ 65 years from primary care[70] and one case–control study recruited a convenience sample of people from primary care[72]) and one that represented targeted screening (patients referred to a hypertension clinic[63]). However, it should be noted that case–control studies, by design, include individuals who are clearly cases or controls and who do not tend to fully reflect a population-based screening population. We were unable to classify the population in one study in primary care because it was part of a RCT that compared systematic population and systematic opportunistic screening, with participants from both arms feeding into the DTA study.[19]

In the included cohort studies, the prevalence of AF varied between 0.93%[64] and 32.93%.[68]

Age was an inclusion criterion in seven studies. Participants had to be ≥ 18 years in two studies,[63,66] ≥ 35 years in one study,[71] ≥ 60 years in one study,[69] ≥ 65 years in two studies[19,70] and ≥ 75 years in one study.[67]

Seven of the studies excluded patients with a pacemaker and/or an implanted defibrillator (in some cases only if they were in active pacing mode).[63,65–67,71,72,74]

Only seven studies reported any characteristics of the included cohorts[19,63,64,66,67,71,72] and, of these, only three reported on comorbidities and/or treatments received.[67,71,72]

Index test
The 15 studies contained 56 eligible comparisons of index test with the reference standard. Not all comparisons in the included studies were included because of ineligible cohorts[68,73,74] or because the results were not reported in a usable format (e.g. pooling multiple readings of the same scans[19] or data being presented per reading rather than per person[71]).

The index test used was classified into one of eight categories to facilitate the analyses:

- pulse palpation
- photoplethysmography
- modified blood pressure monitor
- single-lead ECG
- > 1- and < 12-lead ECG
- 12-lead ECG
- ambulatory monitoring
- two-stage screening.

Two studies investigated pulse palpation.[19,70] In both studies, the pulse palpation was performed by a nurse, but in one study the accuracy of pulse palpation by nurses with different backgrounds (a nurse who had a background in both community and accident and emergency nursing and a practice nurse) was compared.[70] Pulse palpation may have been investigated as an index test in one other study, but no results for this index test were presented.[67]

Photoplethysmography was investigated in one study.[69] This study used a portable device (AFS instrument), which is a finger-probe instrument, as used in pulse oximetry. The resulting pulse waveform was interpreted using fast Fourier transform analysis and the results were presented for three different cut-off values. The data on photoplethysmography for the 0.20 cut-off value was obtained directly from raw data kindly provided to us by Malcolm Lewis (Wales Postgraduate Deanery, Cardiff, 2015, personal communication) and was used to estimate the data for the 0.25 and 0.30 cut-off values (using the number of false negatives and false positives reported in the publication). However, we had confirmation from the authors that what they have termed specificity in their paper is in fact (1 – the proportion of false positives in the whole population). It was therefore possible to obtain the actual 2 × 2 data only for the 0.20 cut-off value and so only this cut-off value was included in the meta-analysis.

A modified blood pressure monitor was used as the index test in three studies.[63,67,71] These devices have been modified to detect pulse irregularity during blood pressure measurement. In all three studies, AF was detected automatically/by an algorithm. In one study, the WatchBP device was used.[67] One study compared two different blood pressure monitors: Microlife BP A200 Plus and OMRON M6.[63] One study used the Microlife BPA100 Plus and different cut-off values for the number of positive results needed to screen positive (readings were taken in triplicate and the cut-off values investigated were positive on the first reading, positive on either of the first two readings, positive on any of the three readings, positive on two of the three readings).[71] This study recruited 27 people with AF, 23 people with non-AF arrhythmias and 23 people with sinus rhythm.[71] We specified that, when possible, when studies were made up of three groups of participants (people with AF, people with another arrhythmia, people in sinus rhythm) we would extract results for people with AF and people in sinus rhythm (because the people in sinus rhythm are more similar to our target screening population). However, in this study the authors pooled the people with a non-AF arrhythmia and the people in sinus rhythm as non-AF cases. This may have impacted on the observed specificity of the test. Because of this, although all results are reported in *Figure 5*, we excluded the data derived from this study[71] in the meta-analysis.

Single-lead ECGs were used in five studies.[66,67,70,72,74] Various different devices were used. The Omron (model HCG-801), which requires placing one electrode on the individual's bare chest while the index finger of the right hand holds the device, was included in two studies,[66,67] interpreted by a cardiologist. The Merlin ECG event recorder, which resembles a watch that the individual covers with the palm of his or her right hand to take a reading, was included in one study,[67] interpreted by a cardiologist. Bipolar ECG using just a single bipolar lead from a 12-lead ECG machine was included in a single study,[70] interpreted by nurses with different backgrounds (training/experiences with ECGs). The AliveCor device, which takes a single-lead ECG recording when fingers from each hand are placed on electrodes on the back of an iPhone (Apple Inc., Cupertino, CA, USA) case, was used in one study,[74] with interpretation by an optimised algorithm. The MyDiagnostick device, in the form of a rod with a metal handle on both ends containing electrodes, was included in a single study[72] and was also analysed automatically.

Greater than 1- and < 12-lead ECGs were used in two studies.[64,65] In one study 12-lead ECGs were reconstructed from V2 and V5 leads or V1 and V4 leads and interpreted by an algorithm.[65] In one study three-lead tele-ECGs were interpreted by two different cardiologists.[64]

Twelve-lead ECGs interpreted by different diagnostic algorithms were compared with 12-lead ECGs interpreted by a cardiologist in six studies, with more than one algorithm tested in some cases.[19,60,65,68,73] One of these studies also investigated the effect of the ECG machine by comparing the diagnostic accuracy of an algorithm for ECGs recorded using two different machines (GE Marquette system MAC5000 and Bionet EKG3000).[60] Twelve-lead ECGs interpreted by nurses with different backgrounds (experience/training with ECGs) and a GP were compared with 12-lead ECGs interpreted by a cardiologist in one study.[70] We excluded one arm of a study investigating the accuracy of 12-lead tele-ECGs because of the index test being a variant of the reference standard (sensitivity and specificity were 100%).[64]

The results of one of the studies that reported the DTA of 12-lead ECGs interpreted by a diagnostic algorithm were reported in two publications.[19,36] In these publications, people could have a positive, negative or uncertain result on this test. The number of people obtaining an uncertain result was slightly different in the two publications. The numbers extracted and reported in this review are from the later publication.[36] In addition, in this publication individuals obtaining an uncertain result were considered to have obtained a negative test result to calculate the sensitivity and specificity of the test. In practice, those with an uncertain test result may be referred for further testing (effectively treated as test positives). If the uncertain results were considered to be positive results it would increase the sensitivity but decrease the specificity.

None of the included studies used any form of ambulatory monitoring.

Two studies investigated some form of two-stage screening. One study modelled a two-stage screening process in which individuals who had a positive result with the modified blood pressure monitor (WatchBP) went on to have a single-lead ECG with autoanalysis (Omron).[67] One study looked at six different combinations of pulse palpation by different nurses (nurse A, who had a background in both community and accident and emergency nursing and who had experience of taking and interpreting ECGs; and nurse B, a practice nurse with no additional ECG training) followed by a bipolar or 12-lead ECG interpreted by either a nurse or a GP for those who screened positive:[70]

1. pulse palpation assessed by a nurse (A) followed by a bipolar ECG interpreted by a nurse (A) for those who screened positive
2. pulse palpation assessed by a nurse (A) followed by a bipolar ECG interpreted by a GP for those who screened positive
3. pulse palpation assessed by a nurse (A) followed by a 12-lead ECG interpreted by a nurse (A) for those who screened positive
4. pulse palpation assessed by a nurse (A) followed by a 12-lead ECG interpreted by a GP for those who screened positive
5. pulse palpation assessed by a nurse (B) followed by a bipolar ECG interpreted by a GP for those who screened positive
6. pulse palpation assessed by a nurse (B) followed by a 12-lead ECG interpreted by a GP for those who screened positive.

Reference standard

The reference standard in all studies was a 12-lead ECG interpreted by a cardiologist, apart from one study that used an ECG classified by a clinician and validated by a researcher (a cardiac electrophysiologist),[68] which was judged to be equivalent to interpretation by a cardiologist.

Quality assessment

The methodological quality of the included studies assessed using the QUADAS-2 criteria is presented in *Table 3*. Many studies investigated multiple index tests. The index tests for each study are listed and grouped according to their quality assessment.

Risk of bias: patient selection

Four studies were judged to be at low risk of bias for patient selection. Three of the cohort studies reported that consecutive or randomly selected individuals were included[62,65,66] and, although no details were provided about the method of inclusion, it seemed possible that it was consecutive in one other study.[69]

The three studies with two sets of inclusion criteria (case–control or two-gate studies)[70–72] and the two studies of unclear study design[73,74] were judged to be at high risk of bias.

The other studies were judged to be at an unclear risk of bias because of concerns over or lack of information on the method of enrolment and/or concerns over or lack of information on exclusion criteria.[19,60,64,67,68]

Risk of bias: index test

The majority of the studies were scored as being at low risk of bias on the index test domain. However, one study was judged to have an unclear risk of bias for some of the index tests investigated because it seemed possible that the test interpreters for some of the index tests were aware of the results of another of the index tests.[67] Three studies were judged to be at high risk of bias on the index test domain, mainly because the cut-off point used to screen positives was not prespecified.[68,69,71]

Risk of bias: reference standard

No studies were judged to be at high risk of bias for this domain. Studies were assessed as being at unclear risk of bias because they did not explicitly report that the interpreters of the reference standard

TABLE 3 Quality assessment of the included studies the using QUADAS-2 criteria[a]

Study	Test classification (test description); interpreter; cut-off value (when necessary)	Risk of bias				Applicability		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Antonicelli 2012[64]	>1- and <12-lead ECG (3-lead tele-ECG); cardiologist (cardiologist 1)	Unclear	Low	Low	Low	High	High	Low
	>1- and <12-lead ECG (3-lead tele-ECG); cardiologist (cardiologist 2)						Low	Low
Gregg 2008[65]	12-lead ECG; automatic/algorithm	Low	Low	Unclear	Low	High	Low	Low
	>1- and <12-lead ECG (reconstructed 12-lead ECG from V2 and V5); automatic/algorithm							
	>1- and <12-lead ECG (reconstructed 12-lead ECG from V1 and V4); automatic/algorithm							
Hobbs 2005,[19] Mant 2007,[36] Swancutt 2004[61]	12-lead ECG; automatic/algorithm (interpretative software)	Unclear	Low	Low	Low	Low	Low	Low
	Pulse palpation; health professional (practice nurse)							
Kaleschke 2009[66]	Single-lead ECG (Omron HeartScan); cardiologist (first author of the study)	Low	Low	Low	Low	High	High	Low
Kearley 2014[67]	Modified blood pressure monitor (WatchBP); automatic/algorithm	Unclear	Low	Low	Low	Low	Low	Low
	Single-lead ECG (Omron); automatic/algorithm							
	Two-stage screening [modified blood pressure monitor (WatchBP) and then single-lead ECG with autoanalysis (Omron) if positive]; automatic/algorithm		Unclear					
	Single-lead ECG (Omron); cardiologist (meta-analysis of four cardiologists' interpretations)			Low			High	
	Single-lead ECG (Omron); cardiologist (cardiologist 1)							
	Single-lead ECG (Omron); cardiologist (cardiologist 2)							
	Single-lead ECG (Omron); cardiologist (cardiologist 3)							
	Single-lead ECG (Omron); cardiologist (cardiologist 4)							
	Single-lead ECG (Merlin); cardiologist (meta-analysis of four cardiologists' interpretations)	Unclear	Low	Low	Low	Low	High	Low

continued

TABLE 3 Quality assessment of the included studies the using QUADAS-2 criteria[a] (*continued*)

Study	Test classification (test description); interpreter; cut-off value (when necessary)	Risk of bias				Applicability		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Langley 2012[68]	Single-lead ECG (Merlin); cardiologist (cardiologist 1)							
	Single-lead ECG (Merlin); cardiologist (cardiologist 2)							
	Single-lead ECG (Merlin); cardiologist (cardiologist 3)							
	Single-lead ECG (Merlin); cardiologist (cardiologist 4)							
	12-lead ECG; automatic/algorithm; coefficient of variation, cut-off value of 0.12	Unclear	High	Unclear	Low	High	Low	Unclear
	12-lead ECG; automatic/algorithm; mean successive beat interval difference, cut-off value of 0.11							
	12-lead ECG; automatic/algorithm; entropy-based algorithm, cut-off value of −1.19							
Lau 2013[74]	Single-lead ECG [iPhone ECG (AliveCor)]; automatic/algorithm (optimised algorithm)	High	Low	Low	Low	High	Low	Low
Lee 2007 (study 1)[60]	12-lead ECG (EKG3000); automatic/algorithm	Unclear	Low	Unclear	Low	High	Low	Low
	12-lead ECG (MAC5000); automatic/algorithm							
Lee 2007 (study 2)[60]	12-lead ECG (EKG3000); automatic/algorithm	Unclear	Low	Unclear	Low	High	Low	Low
Lewis 2011[69]	Photoplethysmography (AFS instrument); automatic/algorithm; cut-off value of 0.30	Low	High	Low	Low	Unclear	Low	Low
	Photoplethysmography (AFS instrument); automatic/algorithm; cut-off value of 0.25							
	Photoplethysmography (AFS instrument); automatic/algorithm; cut-off value of 0.20							
Marazzi 2014,[62] 2012[63]	Modified blood pressure monitor (Microlife BP A200 Plus); automatic/algorithm	Low	Low	Low	Low	Unclear	Low	Low
	Modified blood pressure monitor (OMRON M6); automatic/algorithm							
Slocum 1992[73]	12-lead ECG; automatic/algorithm (discrimination algorithm)	High	Low	Low	Low	Unclear	Low	Low

Study	Test classification (test description); interpreter; cut-off value (when necessary)	Risk of bias				Applicability		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Somerville 2000[70]	Pulse palpation; health professional (nurse A)	High	Low	Low	Low	Low	Low	Low
	Pulse palpation; health professional (nurse B)							
	Single-lead ECG (Bipolar ECG); health professional (nurse A)							
	Single-lead ECG (Bipolar ECG); health professional (nurse B)							
	Single-lead ECG (Bipolar ECG); health professional (nurse C)							
	Single-lead ECG (Bipolar ECG); clinician (GP)							
	12-lead ECG; health professional (nurse A)							
	12-lead ECG; health professional (nurse B)							
	12-lead ECG; clinician (GP)							
	Two-stage screening [pulse palpation assessed by a nurse (A) followed by bipolar ECG interpreted by a nurse (A) for those who screened positive]; health professional (interpreted both pulse palpation and ECG) (nurse A)							
	Two-stage screening [pulse palpation assessed by a nurse (A) followed by bipolar ECG interpreted by a GP for those who screened positive]; health professional and clinician (nurse A performed pulse palpation and GP interpreted ECG)							
	Two-stage screening [pulse palpation assessed by a nurse (A) followed by 12-lead ECG interpreted by a nurse (A) for those who screened positive]; health professional (interpreted both pulse palpation and ECG) (nurse A)							
	Two-stage screening [pulse palpation assessed by a nurse [A] followed by 12-lead ECG interpreted by a							

TABLE 3 Quality assessment of the included studies the using QUADAS-2 criteria[a] (*continued*)

Study	Test classification (test description); interpreter; cut-off value (when necessary)	Risk of bias				Applicability		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
	GP for those who screened positive]; health professional and clinician (nurse A performed pulse palpation and GP interpreted ECG)	High						
	Two-stage screening [pulse palpation assessed by a nurse (B) followed by bipolar ECG interpreted by a GP for those who screened positive]; health professional and clinician (nurse B performed pulse palpation and GP interpreted ECG)							
	Two-stage screening [pulse palpation assessed by a nurse (B) followed by 12-lead ECG interpreted by a GP for those who screened positive]; health professional and clinician (nurse B performed pulse palpation and GP interpreted ECG)							
Stergiou 2009[71]	Modified blood pressure monitor (Microlife BPA100 Plus); automatic/algorithm; cut-off value: positive on the first reading	High	High	Unclear	Low	High	Low	Low
	Modified blood pressure monitor (Microlife BPA100 Plus); automatic/algorithm; cut-off value: positive on either of the first two readings							
	Modified blood pressure monitor (Microlife BPA100 Plus); automatic/algorithm; cut-off value: positive on any of the three readings							
	Modified blood pressure monitor (Microlife BPA100 Plus); automatic/algorithm; cut-off value: positive on two of the three readings							
Vaes 2014[72]	Single-lead ECG (MyDiagnostick); automatic/algorithm	High	Low	Low	Low	Low	Low	Low

a The tailored signalling questions and guidance for this review are detailed in *Appendix 2*.

were unaware of the results of the index test[60,65,71] or because the index test was interpreted by a clinician and validated by a researcher (a cardiac electrophysiologist).[68]

Risk of bias: flow and timing
All studies were judged to be at low risk of bias for the domain of flow and timing.

Applicability: patient selection
Only those studies performed in primary care were judged to have a low level of concern with regard to applicability.[19,67,70,72]

Applicability: index test
Studies in which the index test was interpreted by a cardiologist were judged to be less applicable (high level of concern regarding applicability) than studies interpreted in primary care/by an algorithm.[64,66,67]

Applicability: reference standard
All of the studies were judged to have a low level of concern regarding applicability apart from one, because the index test was interpreted by a clinician and validated by a researcher (a cardiac electrophysiologist).[68]

Study results

The results of the studies are presented in *Figures 3–9*. The raw results are presented regardless of whether or not they were included in the statistical analyses.

Statistical modelling results

The data included in the model are described in *Appendix 6*. There were too few data to estimate asymmetrical HSROC curves. We therefore present results from the symmetrical HSROC model. Evidence of heterogeneity across and within studies was found, justifying our choice of random effects at both of these levels (see *Appendix 6*). In terms of model fit, there was little difference between the model with and the model without distinguishing between interpreter type, suggesting that there is no evidence that DTA depends on interpreter type (see *Appendix 6*). This finding is, however, partly due to a lack of evidence, with some index tests investigated using only one particular interpreter type. Furthermore, because the performance of the tests with different interpreter types is of practical interest in a screening programme, we present the results with and without the inclusion of interpreter type in the definitions of the index tests.

Figures 10 and *11* show the estimated HSROC curves for the different index tests, with interpreter and other test specifics indicated by the different symbols. *Figures 12* and *13* show the SROC curves for each index test and interpreter combination, for those index tests with evidence for more than one interpreter. *Tables 4* and *5* show the point estimates of sensitivity and specificity at the mean of the HSROC model, together with diagnostic odds ratios (DORs) for different index tests and index test/interpreter combinations respectively.

Single-lead ECG generally has high sensitivity and specificity, but there is a lot of heterogeneity, as seen in the results interpreted by a cardiologist (where the majority of the evidence lies). The sensitivity of 12-lead ECG is good for all interpreters, but specificity varies by interpreter. Twelve-lead ECG has very good sensitivity and specificity when interpreted by a GP, suggesting that this could be considered a gold standard test. ECGs with between 1 and 12 leads have good specificity, but sensitivity is variable depending on the interpreter (but good when interpreted by a cardiologist). A modified blood pressure monitor with automatic interpretation has both good sensitivity and good specificity. Photoplethysmography has very good sensitivity, but slightly less good specificity. Pulse palpation by a nurse has the poorest diagnostic performance, with the lowest

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Hobbs 2005,[19] Mant 2007,[36] Swancutt 2004[61] Pulse palpation_Practice nurse	190	441	28	1919	0.87 (0.82 to 0.91)	0.81 (0.80 to 0.83)
Somerville 2000[70] Pulse palpation_Nurse A	26	14	0	46	1.00 (0.87 to 1.00)	0.77 (0.64 to 0.87)
Somerville 2000[70] Pulse palpation_Nurse B	12	4	1	21	0.92 (0.64 to 1.00)	0.84 (0.64 to 0.95)
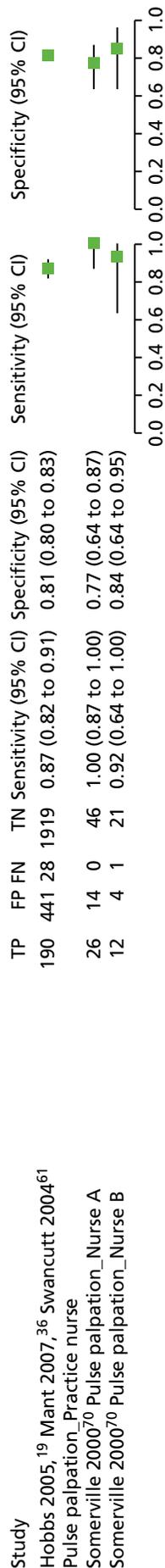
FIGURE 3 Forest plot of pulse palpation. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lewis 2011[69] AFS instrument_algorithm: a cut-off point of 0.20	95	48	0	449	1.00 (0.96 to 1.00)	0.90 (0.87 to 0.93)
Lewis 2011[69] AFS instrument_algorithm: a cut-off point of 0.25	91	35	4	462	0.96 (0.90 to 0.99)	0.93 (0.90 to 0.95)
Lewis 2011[69] AFS instrument_algorithm: a cut-off point of 0.30	84	23	11	474	0.88 (0.80 to 0.94)	0.95 (0.93 to 0.97)
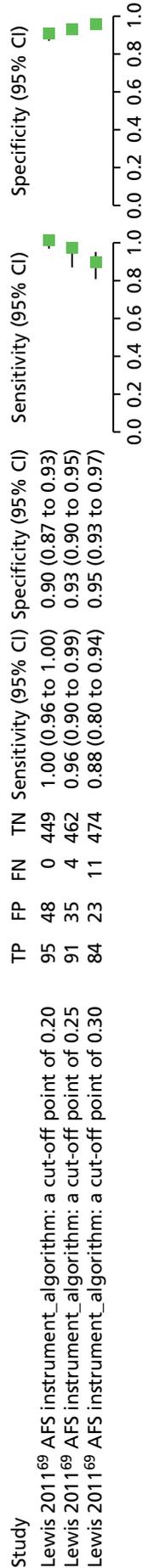
FIGURE 4 Forest plot of photoplethysmography. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kearley 2014[67] Watch BP_Algorithm	75	95	4	825	0.95 (0.88 to 0.99)	0.90 (0.88 to 0.92)
Marazzi 2014[62] 2012[57] Microlife BP A200 Plus_Algorithm	93	19	8	383	0.92 (0.85 to 0.97)	0.95 (0.93 to 0.97)
Marazzi 2014[62] 2012[57] OMRON M6_Algorithm	101	23	0	379	1.00 (0.96 to 1.00)	0.94 (0.92 to 0.96)
Stergiou 2009[71] Microlife_Algorithm: positive on first reading	25	5	2	40	0.93 (0.76 to 0.99)	0.89 (0.76 to 0.96)
Stergiou 2009[71] Microlife_Algorithm: positive on first/second reading	27	11	0	34	1.00 (0.87 to 1.00)	0.76 (0.60 to 0.87)
Stergiou 2009[71] Microlife_Algorithm: positive on second/third reading	27	5	0	40	1.00 (0.87 to 1.00)	0.89 (0.76 to 0.96)
Stergiou 2009[71] Microlife_Algorithm: positive on any reading	27	14	0	31	1.00 (0.87 to 1.00)	0.69 (0.53 to 0.82)

FIGURE 5 Forest plot of modified blood pressure monitors. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kaleschke 2009[66] Omron Heartscan_Cardiologist	142	14	1	348	0.99 (0.96 to 1.00)	0.96 (0.94 to 0.98)
Kearley 2014[67] Merlin_Cardiologist (meta-analysis of 4)	72	89	5	814	0.94 (0.85 to 0.98)	0.90 (0.88 to 0.92)
Kearley 2014[67] Merlin_Cardiologist 1	34	105	1	326	0.97 (0.85 to 1.00)	0.76 (0.71 to 0.80)
Kearley 2014[67] Merlin_Cardiologist 2	32	21	3	410	0.91 (0.77 to 0.98)	0.95 (0.93 to 0.97)
Kearley 2014[67] Merlin_Cardiologist 3	41	40	1	432	0.98 (0.87 to 1.00)	0.92 (0.89 to 0.94)
Kearley 2014[67] Merlin_Cardiologist 4	37	41	5	431	0.88 (0.74 to 0.96)	0.91 (0.88 to 0.94)
Kearley 2014[67] Omron_Algorithm	78	219	1	700	0.99 (0.93 to 1.00)	0.76 (0.73 to 0.79)
Kearley 2014[67] Omron_Cardiologist (meta-analysis of 4)	75	50	4	867	0.95 (0.88 to 0.99)	0.95 (0.93 to 0.96)
Kearley 2014[67] Omron_Cardiologist 1	35	110	2	332	0.95 (0.82 to 0.99)	0.75 (0.71 to 0.79)
Kearley 2014[67] Omron_Cardiologist 2	35	10	2	432	0.95 (0.82 to 0.99)	0.98 (0.96 to 0.99)
Kearley 2014[67] Omron_Cardiologist 3	38	12	4	463	0.90 (0.77 to 0.97)	0.97 (0.96 to 0.99)
Kearley 2014[67] Omron_Cardiologist 4	41	25	1	450	0.98 (0.87 to 1.00)	0.95 (0.92 to 0.97)
Lau 2013[74] iPhone ECG (AliveCor)_Algorithm	47	5	1	151	0.98 (0.89 to 1.00)	0.97 (0.93 to 0.99)
Somerville 2000[70] Bipolar ECG_GP	25	1	1	59	0.96 (0.80 to 1.00)	0.98 (0.91 to 1.00)
Somerville 2000[70] Bipolar ECG_Nurse A	24	7	2	53	0.92 (0.75 to 0.99)	0.88 (0.77 to 0.95)
Somerville 2000[70] Bipolar ECG_Nurse B	12	2	1	23	0.92 (0.64 to 1.00)	0.92 (0.74 to 0.99)
Somerville 2000[70] Bipolar ECG_Nurse C	13	0	0	35	1.00 (0.75 to 1.00)	1.00 (0.90 to 1.00)
Vaes 2014[72] MyDiagnostick_Algorithm	90	6	6	79	0.94 (0.87 to 0.98)	0.93 (0.85 to 0.97)

FIGURE 6 Forest plot of single-lead ECGs. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antonicelli 2011[64] 3-lead tele-ECG_Cardiologist 1	1	0	0	106	1.00 (0.03 to 1.00)	1.00 (0.97 to 1.00)
Antonicelli 2011[64] 3-lead tele-ECG_Cardiologist 2	1	0	0	106	1.00 (0.03 to 1.00)	1.00 (0.97 to 1.00)
Gregg 2008[65] 12-lead from V1 and V4_Algorithm	96	17	13	1659	0.88 (0.80 to 0.93)	0.99 (0.98 to 0.99)
Gregg 2008[65] 12-lead from V2 and V5_Algorithm	92	17	17	1659	0.84 (0.76 to 0.91)	0.99 (0.98 to 0.99)

FIGURE 7 Forest plot of > 1- and < 12-lead ECGs. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity	Specificity
Gregg 2008[65] 12-lead ECG_Algorithm	97	17	12	1659	0.89 (0.82 to 0.94)	0.99 (0.98 to 0.99)		
Hobbs 2005,[19] Mant 2007,[36] Swancutt 2004[61] 12-lead ECG_Algorithm	179	21	36	2320	0.83 (0.78 to 0.88)	0.99 (0.99 to 0.99)		
Langley 2012[68] 12-lead ECG_Algorithm: entropy	52	8	3	104	0.95 (0.85 to 0.99)	0.93 (0.86 to 0.97)		
Langley 2012[68] 12-lead ECG_Algorithm: coefficient of variation	52	8	3	104	0.95 (0.85 to 0.99)	0.93 (0.86 to 0.97)		
Langley 2012[68] 12-lead ECG_Algorithm: mean successive	52	10	3	102	0.95 (0.85 to 0.99)	0.91 (0.84 to 0.96)		
Lee 2007[60] (study 1) 12-lead ECG (EKG3000)_Algorithm	32	2	0	123	1.00 (0.89 to 1.00)	0.98 (0.94 to 1.00)		
Lee 2007[60] (study 1) 12-lead ECG (MAC5000)_Algorithm	30	2	2	123	0.94 (0.79 to 0.99)	0.98 (0.94 to 1.00)		
Lee 2007[60] (study 2) 12-lead ECG (EKG3000)_Algorithm	137	25	12	1096	0.92 (0.86 to 0.96)	0.98 (0.97 to 0.99)		
Slocum 1992[73] 12-lead ECG_Algorithm	28	0	13	66	0.68 (0.52 to 0.82)	1.00 (0.95 to 1.00)		
Somerville 2000[70] 12-lead ECG_GP	26	1	0	59	1.00 (0.87 to 1.00)	0.98 (0.91 to 1.00)		
Somerville 2000[70] 12-lead ECG_Nurse A	25	4	1	56	0.96 (0.80 to 1.00)	0.93 (0.84 to 0.98)		
Somerville 2000[70] 12-lead ECG_Nurse B	13	6	0	19	1.00 (0.75 to 1.00)	0.76 (0.55 to 0.91)		

FIGURE 8 Forest plot of 12-lead ECGs. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity	Specificity
Kearley 2014[67] Watch BP then Omron if positive	75	61	4	859	0.95 (0.88 to 0.99)	0.93 (0.92 to 0.95)		
Somerville 2000[70] Pulse palpation, 12-lead ECG (nurse A)	25	1	1	59	0.96 (0.80 to 1.00)	0.98 (0.91 to 1.00)		
Somerville 2000[70] Pulse palpation, bipolar ECG (nurse A)	24	3	2	57	0.92 (0.75 to 0.99)	0.95 (0.86 to 0.99)		
Somerville 2000[70] Pulse palpation (nurse A) 12-lead ECG (GP)	26	1	0	59	1.00 (0.87 to 1.00)	0.98 (0.91 to 1.00)		
Somerville 2000[70] Pulse palpation (nurse A) bipolar ECG (GP)	25	1	1	59	0.96 (0.80 to 1.00)	0.98 (0.91 to 1.00)		
Somerville 2000[70] Pulse palpation (nurse B) 12-lead ECG (GP)	12	0	1	25	0.92 (0.64 to 1.00)	1.00 (0.86 to 1.00)		
Somerville 2000[70] Pulse palpation (nurse B) bipolar ECG (GP)	11	0	2	25	0.85 (0.55 to 0.98)	1.00 (0.86 to 1.00)		

FIGURE 9 Forest plot of two-stage screening. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

(a)

(b)

(c)
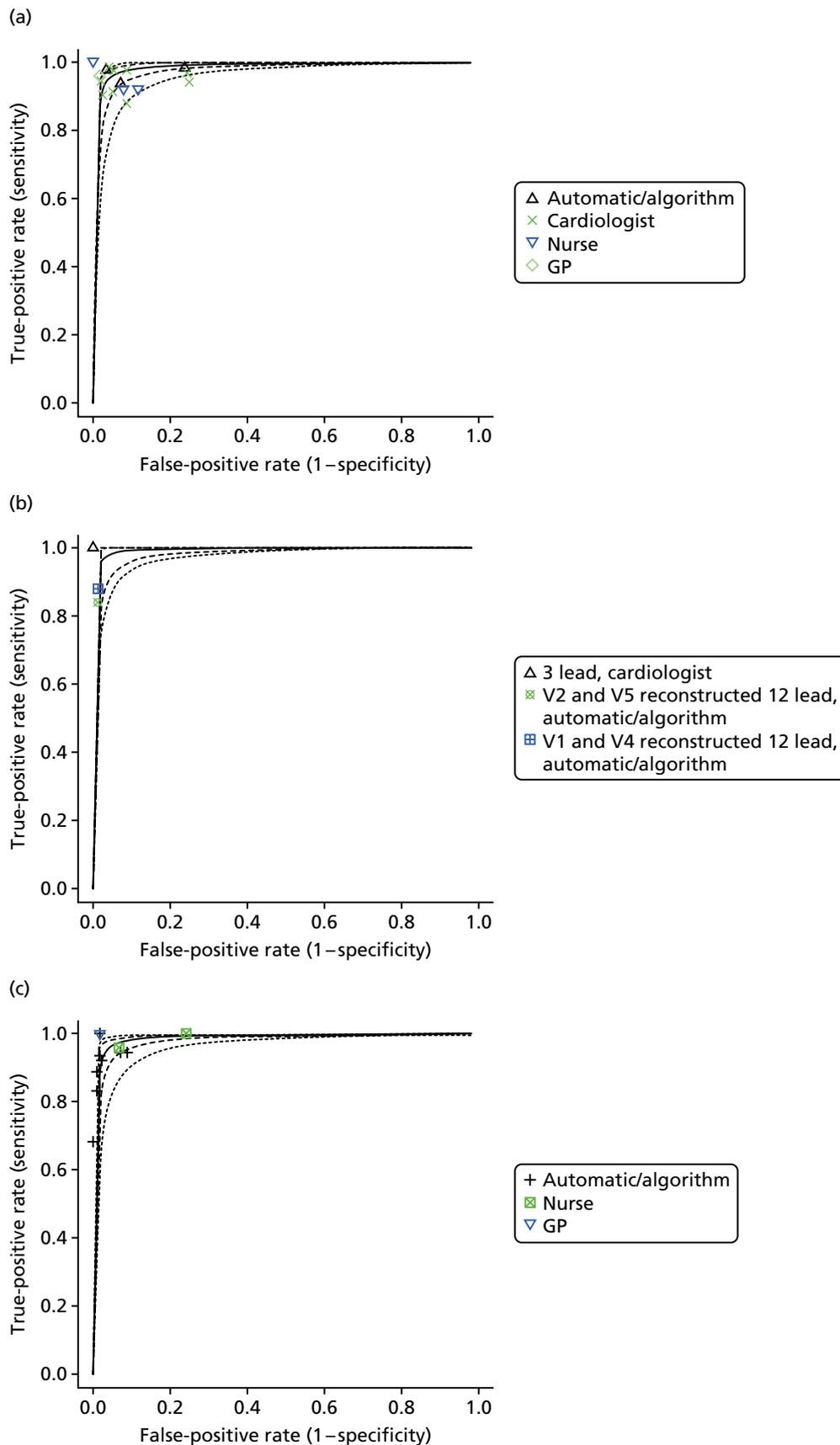
FIGURE 10 Hierarchical summary receiver operating characteristic curves for each ECG index test compared with 12-lead ECG interpreted by a cardiologist: (a) single-lead ECG; (b) > 1- and < 12-lead ECG; and (c) 12-lead ECG. Dashed lines represent 95% credible intervals and dotted lines represent 95% prediction intervals. V1, V2, V4 and V5 refer to leads used by Gregg *et al.*[65]

(a)

(b)

(c)
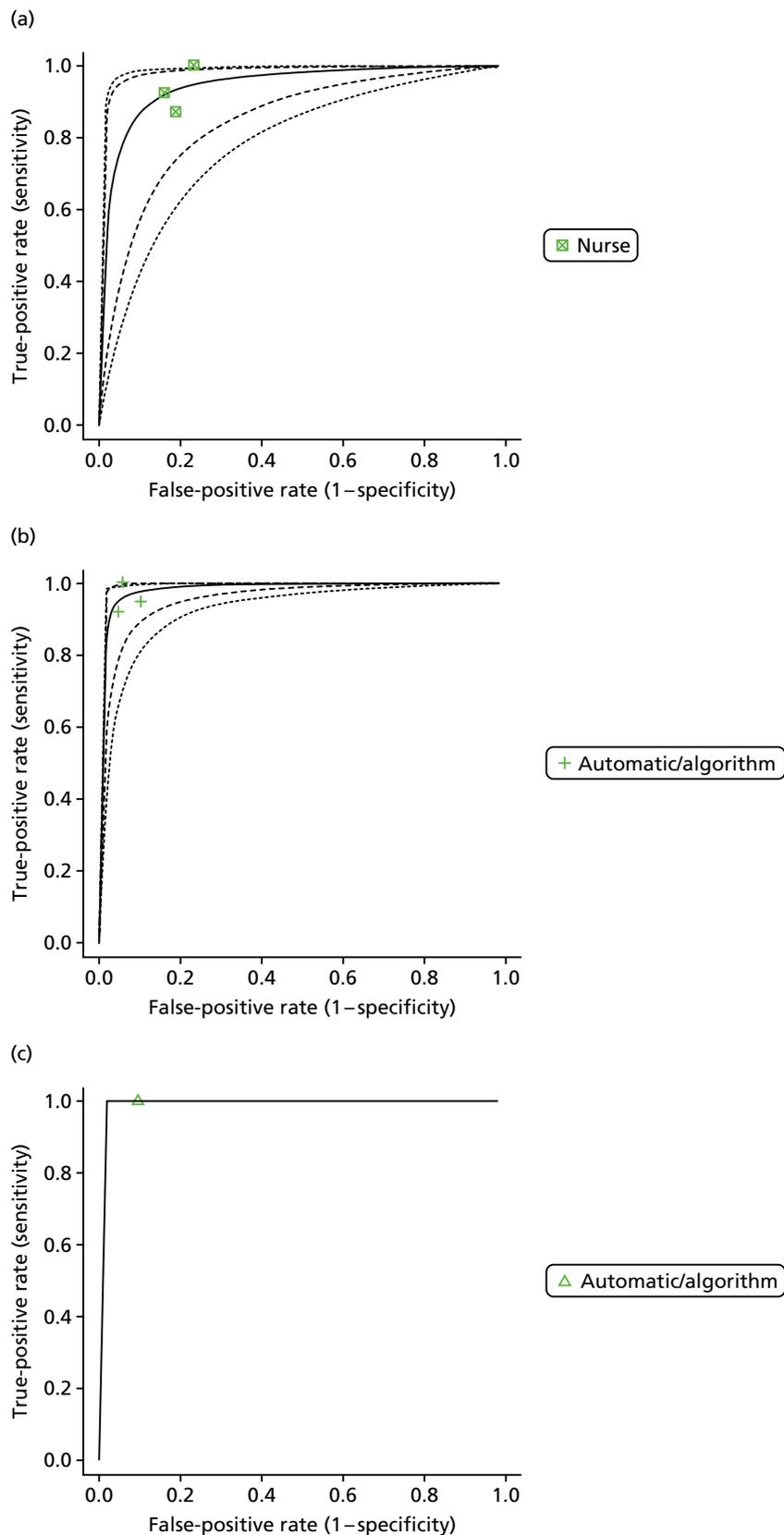
FIGURE 11 Hierarchical summary receiver operating characteristic curves for each ECG index test compared with 12-lead ECG interpreted by a cardiologist: (a) pulse palpation; (b) modified blood pressure monitor; (c) photoplethysmography; and (d) two-stage screening strategy. Dashed lines represent 95% credible intervals and dotted lines represent 95% prediction intervals. (*continued*)
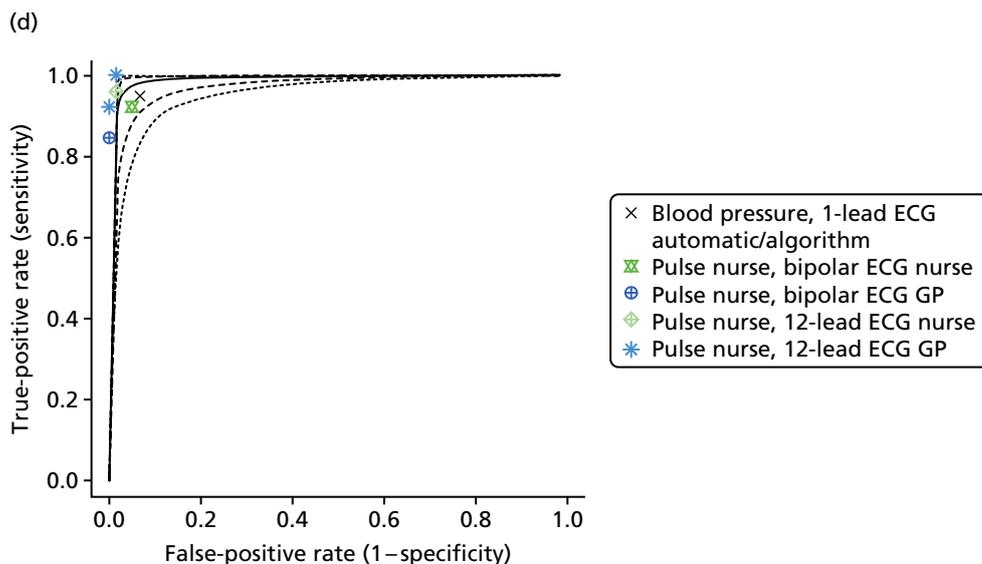
FIGURE 11 Hierarchical summary receiver operating characteristic curves for each ECG index test compared with 12-lead ECG interpreted by a cardiologist: (a) pulse palpation; (b) modified blood pressure monitor; (c) photoplethysmography; and (d) two-stage screening strategy. Dashed lines represent 95% credible intervals and dotted lines represent 95% prediction intervals.

FIGURE 12 Summary receiver operating characteristic curves for each ECG (< 12 leads) index test and interpreter compared with 12-lead ECG interpreted by a cardiologist: (a) single-lead ECG – automatic/algorithm; (b) single-lead ECG – nurse; (c) single-lead ECG – GP; (d) single-lead ECG cardiologist; (e) > 1- and < 12-lead ECG – automatic/algorithm; and (f) > 1- and < 12-lead ECG – cardiologist. Dashed lines represent 95% credible intervals and dotted lines represent 95% prediction intervals. V1, V2, V4 and V5 refer to leads used by Gregg et al.[65] (continued)

(b)

(c)

(d)

FIGURE 12 Summary receiver operating characteristic curves for each ECG (< 12 leads) index test and interpreter compared with 12-lead ECG interpreted by a cardiologist: (a) single-lead ECG – automatic/algorithm; (b) single-lead ECG – nurse; (c) single-lead ECG – GP; (d) single-lead ECG cardiologist; (e) > 1- and < 12-lead ECG – automatic/algorithm; and (f) > 1- and < 12-lead ECG – cardiologist. Dashed lines represent 95% credible intervals and dotted lines represent 95% prediction intervals. V1, V2, V4 and V5 refer to leads used by Gregg *et al.*[65] (*continued*)

(e)

(f)

FIGURE 12 Summary receiver operating characteristic curves for each ECG (< 12 leads) index test and interpreter compared with 12-lead ECG interpreted by a cardiologist: (a) single-lead ECG – automatic/algorithm; (b) single-lead ECG – nurse; (c) single-lead ECG – GP; (d) single-lead ECG cardiologist; (e) > 1- and < 12-lead ECG – automatic/algorithm; and (f) > 1- and < 12-lead ECG – cardiologist. Dashed lines represent 95% credible intervals and dotted lines represent 95% prediction intervals. V1, V2, V4 and V5 refer to leads used by Gregg *et al.*[65]

(a)

(b)

(c)

FIGURE 13 Summary receiver operating characteristic curves for each 12-lead ECG index test and interpreter compared with 12-lead ECG interpreted by a cardiologist: (a) 12-lead ECG – automatic/algorithm; (b) 12-lead ECG – nurse; and (c) 12-lead ECG – GP. Dashed lines represent 95% credible intervals and dotted lines represent 95% prediction intervals.

TABLE 4 Diagnostic test accuracy summaries for the different index tests

Index test	Number of observations	Number of studies	Sensitivity at mean of HSROC model	Specificity at mean of HSROC model	DOR (95% CrI)
Modified blood pressure monitor	3	2	0.955 (0.864 to 0.992)	0.919 (0.777 to 0.982)	2.51 (2.17 to 2.67)
Single-lead ECG	16	5	0.961 (0.917 to 0.986)	0.94 (0.882 to 0.976)	2.56 (2.42 to 2.65)
Two-stage screening strategy	7	2	0.943 (0.838 to 0.988)	0.966 (0.9 to 0.992)	2.63 (2.46 to 2.7)
Photoplethysmography[a]	1	1	1 (1 to 1)	0.867 (0.534 to 0.987)	2.39 (1.71 to 2.68)
12-lead ECG	12	7	0.927 (0.859 to 0.968)	0.974 (0.95 to 0.989)	2.65 (2.59 to 2.69)
> 1- and < 12-lead ECG	4	2	0.839 (0.553 to 0.973)	0.993 (0.978 to 0.999)	2.7 (2.66 to 2.72)
Pulse palpation	3	2	0.916 (0.75 to 0.986)	0.788 (0.51 to 0.945)	2.21 (1.67 to 2.57)

CrI, credible interval.
a Sensitivity of photoplethysmography did not converge.

TABLE 5 Diagnostic test accuracy summaries for the different index test and interpreter combinations

Index test/interpreter	Number of observations	Number of studies	Sensitivity at mean of HSROC model	Specificity at mean of HSROC model	DOR (95% CrI)
Modified blood pressure monitor	3	2	0.953 (0.851 to 0.993)	0.916 (0.759 to 0.983)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	3	3	0.967 (0.9 to 0.995)	0.9 (0.742 to 0.975)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	1	0.929 (0.711 to 0.995)	0.92 (0.7 to 0.992)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	1	0.94 (0.671 to 0.999)	0.973 (0.838 to 1)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	9	2	0.959 (0.878 to 0.992)	0.927 (0.802 to 0.984)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	2	0.941 (0.826 to 0.989)	0.964 (0.889 to 0.993)	2.62 (2.43 to 2.7)
Photoplethysmography	1	1	1 (1 to 1)	0.866 (0.516 to 0.989)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	9	6	0.903 (0.803 to 0.961)	0.98 (0.958 to 0.993)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	1	0.967 (0.824, 1)	0.84 (0.484 to 0.982)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	1	1 (1 to 1)	0.973 (0.843 to 1)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	2	1	0.83 (0.474 to 0.978)	0.985 (0.937 to 0.999)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	2	1	0.981 (0.756 to 1)	1 (0.999 to 1)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	3	2	0.919 (0.749 to 0.988)	0.787 (0.497 to 0.949)	2.21 (1.64 to 2.58)

CrI, credible interval.

mean specificity and sensitivity. All of the two-stage tests have good specificity, but sensitivity depends on the second-stage test, with good sensitivity with a 12-lead ECG as the second-stage test.

The DOR is an overall measure of test performance, with higher values indicating improved test accuracy. The DOR is lowest for pulse palpation by a nurse, followed by photoplethysmography with automatic interpretation. All other index tests have very similar DORs, with the highest DORs for 12-lead ECG (regardless of interpreter), between 1- and 12-lead ECG (automatic or cardiologist interpretation), two-stage tests and single-lead ECG interpreted by a GP. It should be noted that the DOR gives equal weight to false-positive and false-negative results. In the context of a screening programme in which a follow-up diagnostic test is likely to be conducted to confirm the diagnosis, it could be argued that false-negative results are of more importance than false-positive results. The DOR should therefore be interpreted alongside the mean sensitivity of the test.

Subgroup and sensitivity analyses

We conducted various subgroup and sensitivity analyses using the analysis in which the test definition includes the interpreter. *Table 6* summarises the proportion of observations for each test by subgroup, risk of bias and concern of applicability category, indicating where there are enough data to conduct subgroup/sensitivity analyses. The full results of the subgroup and sensitivity analyses can be found in *Appendix 7* and the results are summarised in the following sections.

Study year
There were too few data to perform a subgroup analysis for recent studies. Only one study reported before 2000 (in 1992[73]). The study by Slocum *et al.*[73] was also of unclear design and setting, and it was unclear which screening approach it reflected. Therefore, this study was excluded from the various subgroup analyses reported in the following sections.

Study design
Study design was classified into cohort, case–control and unclear. Ten out of 15 studies (34 observations) were cohort studies, and three out of 15 studies (20 observations) were case–control studies. We conducted analyses including cohort studies only (our preferred study design), but found that the results were robust to restricting to cohort studies only. It should be noted that single-lead ECG – nurse, single-lead ECG – GP, 12-lead ECG – nurse and 12-lead ECG – GP were studied only in non-cohort studies. Tests studied exclusively in cohort studies included photoplethysmography and modified blood pressure monitors.

Screening approach
Three out of 15 studies (29 observations) most closely reflected population-based systematic screening, one out of 15 studies (two observations) most closely reflected targeted screening and the remainder were unclear. We conducted analyses restricting to studies most closely reflected by population-based systematic or targeted screening, which are those most relevant to our decision question. The results were robust to restricting to these populations; however, some tests, including photoplethysmography, were not studied in systematic or targeted screening studies and so the results for these tests could be less applicable to our decision question.

Setting
Four out of 15 studies (31 observations) were carried out in a primary care setting, 10 out of 15 studies (24 observations) were carried out in a non-primary care setting (outpatients or secondary care and above) and in one study (one observation) the setting was unclear. We present the results of subgroup analysis for the primary care setting only as this is most relevant to our decision question. No differences were found from the base-case results but, again, it should be noted that photoplethysmography was studied only in the non-primary care setting and so the results for this test may be less applicable to screening in a primary care setting.

TABLE 6 Percentage of observations by subgroup, risk of bias and concern of applicability category

Index test/interpreter	Cohort study (%)	Systematic or targeted screening (%)	Primary care (%)	Low RoB: patient selection (%)	Low RoB: index test (%)	Low RoB: reference test (%)	Low concern of applicability: patient selection (%)	Low concern of applicability: index test (%)	Low concern of applicability: reference test (%)
Modified blood pressure monitor	100	100	33	67	100	100	33	100	100
Single-lead ECG – automatic/algorithm	33	67	67	0	100	100	67	100	100
Single-lead ECG – nurse	0	100	100	0	100	100	100	100	100
Single-lead ECG – GP	0	100	100	0	100	100	100	100	100
Single-lead ECG – cardiologist	100	89	89	11	56	100	89	0	100
Two-stage screening strategy	14	100	100	0	100	100	100	100	100
Photoplethysmography	100	0	0	100	0	100	0	100	100
12-lead ECG – automatic/algorithm	89	0	11	11	67	22	11	100	67
12-lead ECG – nurse	0	100	100	0	100	100	100	100	100
12-lead ECG – GP	0	100	100	0	100	100	100	100	100
>1- and <12-lead ECG – automatic/algorithm	100	0	0	100	100	0	0	100	100
>1- and <12-lead ECG – cardiologist	100	0	0	0	100	100	0	0	100
Pulse palpation – nurse	33	67	100	0	100	100	100	100	100

Mean age/sex

The mean age of individuals was reported in only seven out of 15 studies (27 observations). Mean age ranged from 61.4 to 79.7 years, with the majority of studies in the range 65–75 years. Because of the high level of missing data on age and the similarity in mean age across studies it was not possible to perform subgroup analyses based on age. The proportion of male participants was reported in the same subset of studies that reported mean age, varying from 47% to 66%. It was therefore also not possible to report the results according to sex.

Prevalence of atrial fibrillation and atrial fibrillation subtypes

It was possible to estimate the prevalence of AF in the included studies and this ranged from 0.9% to 53%, with no distinct groupings. We included prevalence as a continuous covariate in a meta-regression, adding a regression coefficient for the effect of AF prevalence on false positive rate and true positive rate respectively. We excluded case–control studies from the regression as they do not provide meaningful estimates of prevalence. No data were reported on type of AF in any of the included studies and we were therefore unable to report the results according to AF type.

Regression on atrial fibrillation prevalence

We found no evidence of a relationship between diagnostic accuracy and prevalence. Credible intervals (CrIs) for the regression coefficients were wide and included no effect, also supported by the mode fit being similar with and without inclusion of prevalence as a covariate.

Prevalence of comorbidities/stroke risk score, length of monitoring for ambulatory tests and cut-off values for tests with quantitative reading

Only two studies reported information on comorbidities[67,72] and only one reported information on CHA_2DS_2-VASc score.[72] It was therefore not possible to report the results according to comorbidities/stroke risk score. No data were reported on length of monitoring for ambulatory tests and so no subgroup analysis was possible. The only study that reported results for different cut-off values was that by Lewis *et al.*[69] but, as previously noted, the results were reliable only for a single cut-off value in this study (the one with the best sensitivity). It was therefore not possible to perform a subgroup analysis on the cut-off values used.

QUADAS-2 risk of bias domains

We present the results restricting the analysis to low-risk studies only for each domain separately:

- risk of bias according to patient selection
- risk of bias according to index test
- risk of bias according to reference test
- applicability according to patient selection
- applicability according to index test
- applicability according to reference test.

All studies were rated as being at low risk of bias on the flow/timing of tests domain and so no sensitivity analysis was necessary for this domain.

We found that the results were robust to all of these sensitivity analyses; however, it should be noted that there was a lack of data to fully explore these. In particular, for most tests there was at least one domain in which no observations were graded as being at low risk of bias/applicability. The exception was modified blood pressure monitors, for which most of the observations were at low risk of bias/applicability. Photoplethysmography was the only test to have no observations at low risk of bias for the index test, suggesting that we should be cautious when interpreting the results for this test.

Discussion

Summary of results

This review of 15 studies of screening tests for detecting AF found that most tests had a sensitivity (probability of detecting AF in patients with AF) in excess of 0.9. In support of the view that screening could be carried out in primary care, 12-lead ECG interpreted by a GP had a sensitivity of 1 (95% CrI 1 to 1) and also a high specificity of 0.97 (95% CrI 0.84 to 1). Photoplethysmography also had a sensitivity of 1 (95% CrI 1 to 1), but with a lower specificity of 0.87 (95% CrI 0.52 to 0.99). Specificity was in general lower than sensitivity for all of the tests and was lowest for pulse palpation by a nurse (specificity 0.79), 12-lead ECG interpreted by a nurse (specificity 0.84) and photoplethysmography (specificity 0.87). Tests with the highest DOR were the 12-lead ECG (regardless of interpreter), between 1- and 12-lead ECG (automatic or cardiologist interpretation), two-stage tests and single-lead ECG interpreted by a GP, with all of these tests having similar DORs. However, the DOR gives equal weight to true positives and true negatives, whereas for a screening programme test sensitivity takes precedence over specificity (if the test is not too burdensome), so that AF cases are not missed. False positives should be picked up subsequently by a confirmatory diagnostic test. We are therefore interested in finding a test that primarily has good sensitivity; specificity is a secondary consideration to avoid too many individuals being referred for a confirmatory diagnostic test.

In general, for a given interpreter, the results for single-lead ECGs were less accurate and more variable than those for ECGs with more than one lead. Nurse interpretation of single-lead ECGs performed similarly to single-lead ECGs with other interpretation methods, but nurse interpretation of 12-lead ECGs did not perform as well as other interpretation methods. Automatic interpretation did not have a consistent impact on test accuracy, with automatic interpretation of single-lead ECGs having a high sensitivity but variable specificity. In contrast, automatic interpretation of ECGs with more leads had good specificity but variable sensitivity.

The different two-stage screening strategies all had very high specificity, but sensitivity was high only when a 12-lead ECG was used as the second-stage test. This supports the need for a 12-lead ECG as a diagnostic test. Interestingly, a second-stage 12-lead ECG test interpreted by a (trained) nurse performed similarly to that interpreted by a GP.

Strengths and limitations

Our review followed a published protocol, registered on PROSPERO.[56] Inclusion was restricted to studies that best represented a population-based screening programme conducted in primary care. A comprehensive search strategy was used and the risk of bias and applicability of studies was assessed. The effect of interpreter as well as the screening test utilised were both investigated. Preplanned subgroup and sensitivity analyses were conducted when data allowed.

Analyses were limited by the small number of studies that met the inclusion criteria and hence the lack of statistical power to detect meaningful differences. For example, not all combinations of screening test and interpreter were included in the studies and, when present, few observations were reported. For this reason tests were grouped in the statistical modelling. In particular, > 1- and < 12-lead ECGs included 3-lead ECGs as well as reconstructed ECGs resulting from different sets of leads. Similarly, two-stage screening strategies encompassed varying combinations of tests in the two stages.

There was substantial heterogeneity in the test/interpreter combinations when there were repeated observations, and it is likely that the same heterogeneity would be seen for the other tests if multiple observations were available. Furthermore, we would expect more heterogeneity in test accuracy in practice than is seen in research studies.

None of the included studies was assessed to be at low risk of bias across all domains and to have a low level of concern regarding applicability. In particular, only four studies were conducted in primary care. Although we had preplanned a wide range of subgroup analyses and sensitivity analyses, there were enough data available to perform only some of these. The way that the data were reported was in some cases challenging. Some studies were designed so that the same ECGs were read by different individuals, whereas in other studies different individuals read different sets of ECGs. Some studies presented the results for patients with AF, individuals in sinus rhythm and patients with non-AF arrhythmias, but some combined the sinus rhythm and non-AF groups when reporting results, which will underestimate test accuracy as applied to the general population in which the proportion of other arrhythmias is much lower. Although photoplethysmography performed well, only a single study[69] used this method, and this study was able to provide data only for a single cut-off point, which was chosen retrospectively for having the best sensitivity. The index test in this study was rated as being at high risk of bias and the applicability of patient selection was unclear. We are therefore cautious in interpreting the results from this screening test and would ideally want the results to be replicated in a new cohort of individuals.

Findings in the context of previous research

European Society of Cardiology guidance recommends pulse palpation as the first step for AF screening.[32,75] We found pulse palpation to have the lowest DOR, driven by its low specificity.

Guidance from NICE (medical technologies guidance MTG13)[76] recommends the use of modified blood pressure monitors for the detection of AF in individuals undergoing diagnostic testing or monitoring for hypertension. We found that modified blood pressure monitors had high sensitivity, although specificity was not as high as with other tests.

During the course of this review, two systematic reviews were published. One reported the accuracy of methods for detecting an irregular pulse and suspected AF;[77] the other reported the accuracy of methods for diagnosing AF using 12-lead ECG.[78] In addition, a systematic review of the DTA of automated blood pressure monitors for opportunistic AF detection,[79] and a review and meta-analysis of the DTA of the algorithm for AF detection during automated blood pressure measurement[80] have been published. These reviews are summarised in *Table 7*.

TABLE 7 Systematic reviews and meta-analyses of potential screening tests for AF

Study	Population	Index test	Reference standard	Findings	Differences from our review
Taggar 2016[77]	Adults aged ≥ 18 years	Screening tests for AF	Any ECG interpreted by a competent professional	Modified blood pressure monitors and non-12-lead ECG devices had the greatest diagnostic accuracy. Different categories of tests had similar sensitivities but pulse palpation had markedly lower specificity than the other tests. Subgroup analyses for studies performed in primary care only (pulse palpation and non-12-lead ECGs) had similar findings to those of the main analyses	Included studies conducted in cardiology patients, patients recruited from anticoagulation clinics and patients with AF attending elective cardioversion. Reference standard was broader (any ECG interpreted by a competent professional, rather than 12-lead ECG interpreted by a cardiologist). The accuracy of 12-lead ECGs and the results of two-stage screening strategies were not extracted. When studies reported findings using multiple thresholds for the same intervention, only data in which thresholds maximised the sensitivity of the index test were extracted. Categorisation of the index tests was slightly different (the review divided the tests into blood pressure monitors, non-12-lead ECGs, pulse palpation and smartphone applications)
Taggar 2015[78]	Adults aged ≥ 18 years	Any method of interpreting a 12-lead ECG to show AF [the interpreters of the index test were divided into automated software and health-care professionals; subgroups of health-care professionals were secondary care physicians and primary care physicians (GPs and practice nurses)]	12-lead ECG interpreted by a trained specialist	The sensitivity of all interpreters was similar, but automated software analysis had slightly greater specificity. The authors suggest that the accuracy of diagnosing AF in primary care in their results is being driven by GPs' diagnosis of AF and that specificity is lower when ECGs are interpreted by nurses	Population inclusion criteria; only one type of test
Kane 2016[79]	Any	Automated blood pressure monitors used for opportunistic AF detection	Any	All seven studies included reported specificity of > 85% and six reported sensitivity of > 90%	Population inclusion criteria; only one type of test; reference standard; no meta-analysis performed
Verberk 2016[80]	Any	The AF detection algorithm of the Microlife automated blood pressure monitor	12-lead ECG interpreted by a cardiologist	Overall sensitivity 0.98 (95% CI 0.95 to 1.00), specificity 0.92 (95% CI 0.88 to 0.96)	Population inclusion criteria; only one type of test

Chapter 4　Results: systematic review of randomised controlled trials comparing screening strategies

Study selection

The systematic review is summarised in a PRISMA diagram (*Figure 14*). Seventeen records were identified from the original Cochrane review[49] and a further three records were kindly provided to us from the ongoing Cochrane update review. An additional 2934 records were identified from our update searches from April to May 2012 and from July to December 2015. In total, 2954 records were identified from all sources. After removing duplicates, 2204 records remained. Sifting of these records revealed six eligible articles[26,40,81–83] representing four RCTs, all involving two study arms. These four RCTs, together with the
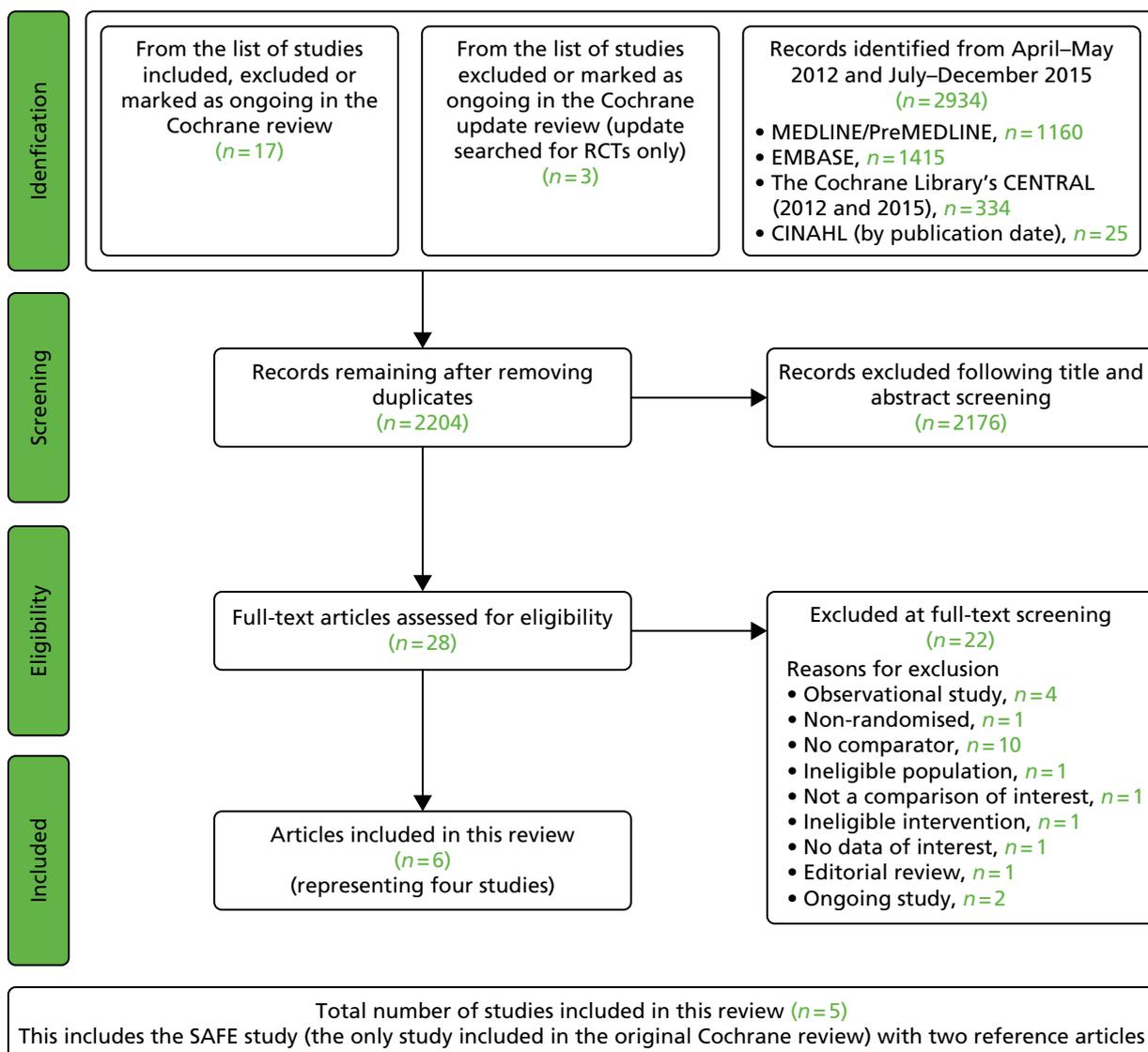
FIGURE 14 The PRISMA flow chart: review of RCTs of screening strategies. CINAHL, Cumulative Index to Nursing and Allied Health Literature.

previously identified RCT from the original Cochrane review involving three study arms (with two reference articles),[19,84] were included in this review.

We also found one ongoing study and one study that has yet to publish its results (described in *Table 8*).[85,86]

Study characteristics

Table 9 provides a summary of the characteristics of the five included RCTs. All of the RCTs were conducted in Europe, with two each conducted in the UK[19,82] and Spain[81,83] and one conducted in Sweden.[40] Four studies[19,26,81,82] were multicentre RCTs and one was a single-centre RCT.[83] Four RCTs involved individuals aged ≥ 65 years[19,81–83] and one RCT involved individuals aged 75 and 76 years.[40] All five RCTs involved both sexes. Two RCTs involved only individuals without a previous diagnosis of AF.[81,83] Three RCTs each evaluated systematic population screening[19,40,82] and systematic opportunistic screening,[19,79,82] one RCT evaluated systematic targeted screening directly,[83] and one RCT evaluated systematic targeted screening indirectly by using results from those individuals identified at 'high risk' among those randomised to systematic population screening.[19] Screening tests employed across the RCTs were mainly pulse palpation and ECG with AF diagnosis predominantly by ECG. However, it was clear in only one RCT that ECG interpretation was by a cardiologist.[19] The primary outcome across a majority of the included RCTs was the number of new cases of AF in the study arms. The duration of study follow-up varied across the RCTs, ranging from 1 year to 5 years.

TABLE 8 Ongoing RCTs and RCTs yet to publish results that compared screening strategies

Study and study design	Population and setting	Intervention	Control	Primary outcome	Estimated completion date
IDEAL-MD;[85] two-arm cluster-randomised, multicentre trial	People aged ≥ 65 years, general practice	Opportunistic screening with single-lead ECG (MyDiagnostick). When the device indicates a positive result, the single-lead ECG will be assessed to confirm/reject the diagnosis	No screening/ usual care	Difference in yield of newly detected cases of AF in 1 year between intervention and control practices	March 2017 (no results available to date)
D₂AF;[86] two-arm cluster-randomised, multicentre trial	People aged ≥ 65 years without previously documented AF, general practice	Opportunistic screening with single-lead ECG (MyDiagnostick), pulse palpation and a modified blood pressure monitor with an algorithm for irregular beat detection (WatchBP Home A). Patients with a positive result on at least one of the tests and a random sample of patients with negative tests will have their AF status confirmed by 12-lead ECG and if this is negative they will additionally be tested with a 2-week Holter monitor. Patients tested with the 2-week Holter monitor will concurrently use the MyDiagnostick three times per day	No screening/ usual care	Difference in yield of newly detected cases of AF in 1 year between intervention and control practices	January 2018

D₂AF, Detecting and Diagnosing Atrial Fibrillation; IDEAL-MD, Improving DEtection of Atrial fibriLlation in Primary Care With the MyDiagnostick.

TABLE 9 Characteristics of the five included studies

Study (country)	Design	Setting (centres)	Participant age (sex)	Participant health status	Number randomised (% male)	Interventions and comparator	Screening test	Diagnostic test	Primary outcome assessed	Duration of follow-up (months)
SAFE[19,84] (UK)	Multicentre RCT with cluster randomisation to screening or no screening and individual randomisation within screening clusters to systematic population or systematic opportunistic screening	General practice (n = 50)	≥ 65 years (both)	All individuals aged ≥ 65 years, excluding patients who were terminally ill	n = 14,802 (42.6)	Systematic population and systematic opportunistic screening and no screening	Systematic population screening arm: 12-lead ECG interpreted by a cardiologist; systematic opportunistic screening arm: pulse palpation by a GP or nurse followed by 12-lead ECG interpreted by a cardiologist if pulse irregular	12-lead ECG interpreted by a cardiologist	New AF (noted in medical record)	12
STROKESTOP[26,40] (Sweden)	Two-arm RCT	Screening centre (multicentre – n = 7 screening centres)	75 and 76 years (both)	All 75- and 76-year-olds living in Stockholm County or in the Halland region	n = 28768 (NR)	Systematic population screening and no screening	ECG at index visit. Participants in sinus rhythm were then given a single-lead ECG recorder and instructed to use it twice daily and whenever they noticed palpitations for a period of 2 weeks. The ECGs were interpreted by a nurse. All abnormal ECGs were referred to a cardiologist	Single-lead ECG interpreted by a cardiologist	Stroke (new AF not the focus of the study)	60

TABLE 9 Characteristics of the five included studies *(continued)*

Study (country)	Design	Setting (centres)	Participant age (sex)	Participant health status	Number randomised (% male)	Interventions and comparator	Screening test	Diagnostic test	Primary outcome assessed	Duration of follow-up (months)
Morgan[82] (UK)	Two-arm RCT	General practice (n = 4)	65–100 years (both)	No exclusion of individuals based on health status	n = 3001 (41.25)	Systematic opportunistic and population screening	Pulse palpation by a GP or a trained nurse	Lead II rhythm strip ECG but interpreter not reported. However, for opportunistic screening, diagnosis could be made clinically	AF	6
DOFA-AP[81] (Spain)	Two-arm cluster RCT	Primary care [multicentre (NR)]	≥ 65 years (both)	Individuals without AF	n = 12,870 according to the study protocol	Systematic opportunistic screening and no screening	Pulse palpation by a GP	ECG (type and interpreter NR)	New AF	12
EARLY[83] (Spain)	Two-arm RCT	Primary care (n = 1)	≥ 65 years (both)	Patients without AF but with one or more risk factors for AF	n = 4000 (49% of those screened)	Targeted screening and no screening	ECG (number of leads and interpreter NR), physical examination and medical history every 6 months over 2 years. Participants randomised to targeted screening were also instructed about warning signs and how to take their own pulse	NR	New AF	24

DOFA-AP, Detección Oportunista de Fibrilación Auricular en Atención Primaria; EARLY, Early diagnosis of Atrial fibrillation: a Randomized triaL in primarY care; NR, not reported.

Overview of the five included studies

The SAFE study[19,84]

This study compared systematic population screening, systematic opportunistic screening and no screening. The study employed a two-stage randomisation process. The first step was a balanced cluster randomisation (stratified based on Townsend score quartiles and practice size) of 50 practices in the UK to either the intervention (screening) arm or the control (no screening) arm. This was then followed by a random assignment of individuals aged ≥ 65 years in the intervention centres to either systematic opportunistic or systematic population screening. Individuals aged ≥ 65 years were also identified in the control arm.

Individuals randomised to systematic opportunistic screening had a flag placed in their notes to encourage practice staff to take a pulse recording during routine consultation. If their pulse was found to be irregular individuals were invited to attend a screening clinic. Individuals in the systematic population screening arm were all invited by post to attend a screening clinic. The screening clinic was the same for both screening arms (i.e. they differ only in how individuals were selected for invitation to the screening clinic). At the screening clinic consent was obtained, baseline characteristics were recorded and a 12-lead ECG was performed, which was interpreted by two blinded cardiologists, with interpretation by a third blinded cardiologist in the case of disagreement. An acceptability questionnaire was also delivered to patients.

This study also identified 'high-risk patients' with cardiac failure, hypertension, rheumatic heart disease, previous MI, angina, diabetes mellitus, hyperthyroidism, previous stroke or previous TIA (using practice disease registers) among the individuals randomised to systematic population screening. If these patients were selectively invited to screening, this would represent a targeted screening programme. The results of targeted screening (using the whole population screening as the denominator) were compared with the results of no screening.

Patients with previously diagnosed AF were included in the study, but the number of such patients was reported so that the number of newly diagnosed cases could be ascertained. To identify previously diagnosed cases of AF, patient notes were searched for 'G573 AF/flutter', '327 ECG supraventricular arrhythmia', '181 palpitations', 'digoxin', 'amiodarone', 'verapamil', 'sotalol', 'metoprolol', 'warfarin' and 'aspirin'. Case notes of patients identified were reviewed and a diagnosis of AF was accepted if there were hospital letters referring to AF or confirmatory ECGs within the previous 5 years.

The primary study outcome was the number of newly diagnosed cases of AF during the 12 months of follow-up. Newly diagnosed cases were identified using the same computer searches as performed at baseline.

The STROKESTOP study[26,40]

The STROKESTOP study aims to test the hypothesis that screening 75- and 76-year-olds for AF will reduce stroke incidence cost-effectively. All 75- and 76-year-olds in two regions of Sweden were randomised to either the systematic population screening arm or the no screening arm.

Participants who were randomised to the screening arm were sent an invitation by post. At the screening centre, participants were asked about their medical history and an index ECG (using a hand-held single-lead ECG) was taken. Participants without a history of AF who were in sinus rhythm were given a hand-held single-lead ECG and instructed to use it twice daily and whenever palpitations were felt over a period of 2 weeks. The ECG transmitted the results to a database where they were interpreted by a nurse. All abnormal ECGs were referred to the investigating cardiologist. When results were unclear, referral for interpretation by a consensus group was used.

An interim publication has reported the AF detection rates in individuals in the screening arm who presented at the screening clinic.[40] Both the number of people with known AF and the number of people with newly detected AF were reported. Results for the control arm will be obtained from a National Patient Registry (Emma Svennberg, Danderyds Sjukhus AB, Stockholm, Sweden, 2015, personal communication) but are not available currently.

The primary outcome of this study is stroke prevention resulting from early identification of AF through screening, and subsequent treatment, with 5 years' follow-up. However, the study is still ongoing and not due to report until 2019. Because the results are currently available only for the screening arm, this study did not contribute data to our statistical analysis.

The Morgan and Mant[82] study

This study compared systematic opportunistic screening with systematic population screening. Four general practices in the UK were chosen to represent each quartile of the small area standardised mortality ratio of the geographical area served and approximately 750 individuals aged ≥ 65 years from each practice were randomly selected and then randomised to either the systematic opportunistic or the systematic screening arm.

In the systematic opportunistic screening arm a flag was inserted into the patients' notes so that if they consulted with their GP or nurse for any reason during the duration of the study their pulse was taken. If their pulse was found to be irregular the nurse or GP could choose to request a confirmatory ECG. In the reported results from this study it was assumed that there was a clinical diagnosis of AF if the pulse was irregular but a confirmatory ECG was not requested. In the systematic screening arm, all individuals were invited by post to a screening appointment with a nurse who took a radial pulse reading for at least 20 seconds to record evidence of irregularity. A two-lead ECG reading interpreted by a GP was used for validation. The primary study outcome was the number of AF cases after 6 months' follow-up, but the study also reports the number of cases with no previous evidence of AF in their medical records. However, the number of people with and without AF in their medical records at baseline was not reported.

The two arms differed in the validation of AF cases: in the systematic screening arm, a two-lead ECG interpreted by a GP was used for the validation of new cases of AF, whereas in the systematic opportunistic screening arm ECG validation took place at the discretion of the nurse or GP taking the pulse, with some cases identified by clinical diagnosis only and others based on an ECG (number of leads and interpreter unspecified). This difference in validation could have influenced the reported results.

The DOFA-AP study[81]

This study compares systematic opportunistic screening with a control group (termed 'usual care'). Details of the study are currently only available in the protocol[81] and in an unpublished abstract shared with us by the ongoing Cochrane update review authors (P Moran, Health Information and Quality Authority, Ireland, 4 November 2015, personal communication). Eligibility criteria for the study were participant aged ≥ 65 years and attending their health centre in Spain for any reason without a previous diagnosis of AF.

In the systematic opportunistic screening arm, it is reported that individuals underwent 'active screening for AF, regardless of the reason of their visit' (p. 4).[81] In the control arm an individual's pulse was measured if he or she presented with symptoms of sequelae associated with AF. AF was confirmed using an ECG, but the ECG type and the information regarding the interpreter were not provided.

However, we have concerns about the interpretation of the results of the study. It is possible that to be eligible for inclusion in the control group individuals had to be symptomatic, as the protocol reports that individuals presenting with AF symptoms (dyspnoea, chest pain palpitations or dizziness) or AF-induced sequelae (stroke, TIAs, peripheral embolism, heart failure or other associated health disorders) were eligible. Another statement in the protocol that supports this suggestion is 'since the probability of finding patients meeting the eligibility criteria is much lower for the [control group (usual care arm)] than for the [intervention arm], stratified randomization will be used to assign at least 200 primary care health professionals to the two groups, in the ratio of a 2 : 3 ratio' (p. 4).[81] This suggests that the study may in fact be better described as a comparison of targeted testing of symptomatic individuals (control arm) with systematic opportunistic screening (intervention arm).

From the reported results in the abstract, the denominators are also not clear from the information available.

The study result is also unintuitive (the authors reported more AF cases in the control arm than in the screening arm), although this could be explained by differential inclusion criteria for each arm. Because of the lack of information about the denominator for the control arm (both the actual number and what it represents, i.e. the number randomised or the number randomised who presented with symptoms or sequelae), the results from this study were excluded from our analysis.

The EARLY study[83]

This study compared targeted screening for AF with usual care (no systematic screening). Individuals in a primary health-care centre in Spain were selected from electronic records on the basis of not having a diagnosis of AF but having at least one of the main risk factors for AF. Individuals were then randomised to either an early detection programme or the control arm. Individuals in the early detection programme arm were invited by telephone to participate in the study. The early detection programme involved visits every 6 months over a 2-year period, at which an ECG, a physical examination and a medical history were taken/performed. On the first visit, advice on pulse taking was also given. The primary end point of the study was the proportion of individuals with newly diagnosed AF at 6 months. A secondary outcome was the number of individuals diagnosed with AF during the 2-year follow-up. Cases of AF were determined from the medical records system.

Many individuals were excluded from each arm for different reasons. Of the 2000 people randomised to screening, 616 were not found, 425 declined to participate, 78 were no longer assigned to the health centres, three were deceased, 153 fulfilled at least one exclusion criterion (were disabled, had a pacemaker, did not really have a risk factor, had AF at baseline) and 262 were not contacted. Of the 2000 people randomised to the control group, 42 were no longer assigned to the health centres, six were deceased, 38 fulfilled at least one exclusion criterion and 1449 were not contacted.

Although the study reported that an intention-to-treat analysis was performed, it considered only the individuals included in the study. In the intervention group, individuals could decline to participate, but in the control group they could not. The study actually compared the incidence of AF in all people eligible for screening but who received usual care with the incidence of AF in a self-selecting subgroup of people who both were eligible and attended screening.

On the basis of these methodological limitations, we felt that it was inappropriate to include the results from this study in our statistical analysis.

Risk of bias within the included studies

Table 10 and Figure 15 show the risk-of-bias assessments for the included studies. A complete assessment of the STROKESTOP study[26,40] could not be made as the results for only one arm of the study have been published. None of the studies was assessed to be at low risk of bias across all of the domains. None of the studies could blind participants and personnel to which arm they had been randomised to and therefore they were all judged to be at high risk of bias for this domain. The SAFE study,[19,84] on which the results of most of this review are based, was judged to be at low or unclear risk of bias for the other domains.

Study results and meta-analysis

Newly diagnosed cases of atrial fibrillation among individuals without atrial fibrillation at baseline

Of the five RCTs included in this review, only one (the SAFE study[19,84]) provided usable data on the relative effects of screening on the proportion of newly diagnosed cases of AF among individuals without AF at baseline. The results of this study are presented in Table 11.

TABLE 10 Risk-of-bias assessments for the five included RCTs

Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias(es)
SAFE[19,84]	Low Randomisation of general practices: 'After stratification for practice size and deprivation (based on Townsend score), we used MINITAB to select randomly two equal size groups from those practices within a particular stratum' (p. 2)[84] Randomisation of individuals within intervention practices to systematic population or systematic opportunistic screening: 'After stratification for known atrial fibrillation (see below) we randomly allocated patients from the 25 intervention practices to systematic or opportunistic screening. We used SPSS to allocate patients randomly from this list to either systematic or opportunistic screening to create two equal size groups of patients' (p. 2)[84]	Low Although the study reports that there was no allocation concealment ('There was no deliberate concealment of allocation to the trial arms . . . the trial statistician determined allocation, which was implemented by the trial coordinator'), as clusters were recruited prior to randomisation and individuals were identified and randomised without knowing which group would receive systematic screening and which would receive opportunistic screening, it seems unlikely that randomisation could have been subverted	High It would not have been possible to blind participants. GPs/nurses who saw an individual randomised to the opportunistic screening arm would have known his or her allocation because of the presence of a flag on his or her notes. It is possible that the nurses running the screening clinics may have been unaware of whether people had been sent an invitation and were part of the systematic screening arm or whether they had been invited to attend after being found to have an irregular pulse in the opportunistic arm. However, it is not explicitly reported that they were blinded	Unclear The primary outcome measure was the number of new cases of AF detected during the 12-month period and the new cases were ascertained from computer searches of patient notes (which were then reviewed manually). A diagnosis was accepted if there were hospital letters referring to AF or confirmatory ECGs. Twelve-lead ECGs taken from individuals attending screening clinics (after being randomised to systematic population screening or after being found to have an irregular pulse during opportunistic screening) were interpreted by cardiologists blind to the method of recruitment	Low AF status was available for > 99% of participants	Low All outcomes specified in the trial protocol were reported	*Baseline outcome measures* Unclear The baseline prevalence of AF was 7.9% (95% CI 7.2% to 8.7%) in the control arm and 6.9% (95% CI 6.2% to 7.6%) in the opportunistic and systematic screening arms *Baseline characteristics* Low Age and sex balance were similar across arms *Contamination* Unclear Randomisation to intervention (screening) or control was carried out by practice – low risk of contamination. However, randomisation to opportunistic or systematic screening was carried out by individual. It is possible that there was contamination between these arms and

Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias(es)
							that people in the opportunistic arm were more thoroughly investigated for AF because of the systematic screening programme that was also occurring. This is borne out by the fact that in the opportunistic arm there were 75 incident cases, of which 31 were identified by the screening programme and 44 were identified outside the screening programme. In the systematic arm there were 74 cases identified, 52 through the screening programme and 22 outside the programme
Morgan and Mant[82]	Low 'Randomisation of the practice lists was done at individual patient level, using the random sample command of STATA statistical software' (p. 377)[82]	Low Individuals were randomly selected from each practice list and then randomised prior to consent being sought and therefore selection bias from anticipation of allocation could not have occurred. As participants were selected randomly and statistical software was used it	High It would not have been possible to blind participants. GPs/ nurses who saw an individual randomised to the opportunistic screening arm would have known his or her allocation because of the presence of a flag on his or her notes. Nurses performing the nurse screening	High In the systematic screening arm an AF diagnosis was confirmed using a lead II rhythm strip. However, in the opportunistic arm, an AF diagnosis could be made without a confirmatory ECG. The interpreter of the lead II rhythm strips is likely to have known the arm that	High 'Analysis was performed on an intention-to-screen basis, including all individuals initially included in the screened and opportunistic case finding [screening] arms' (p. 377).[82] However, although an intention-to-treat analysis was performed, as 73.3%	Unclear No protocol identified	High Cases of AF in the opportunistic screening arm could be diagnosed on the basis of clinical judgement alone

TABLE 10 Risk-of-bias assessments for the five included RCTs *(continued)*

Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias(es)
		seems unlikely that personnel could have anticipated allocation/ subverted the randomisation	would have known that attendees had been randomised to systematic screening	participants were randomised to and the GPs making a clinical diagnosis of AF in those in the opportunistic arm would also have known the group that they were randomised to	of the systematic arm and 29.2% of the opportunistic arm had their pulse assessed there are missing data on AF status for 26.7% of the systematic arm and 70.8% of the opportunistic arm. No patient record review was performed to determine whether or not AF had been detected in people outside the screening programme. Within the four practices, the proportion of people who had their pulse assessed after being randomised to opportunistic screening varied from 8.0% to 52.1%. In addition, 4.3% of the notes of those randomised to opportunistic screening were not flagged		

Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias(es)
DOFA-AP/ FAMDAP[81]	'Randomisation and allocation to trial groups will be carried out by a central computer system' (p. 1)[81]	Unclear Patient selection was performed by consecutive sampling. However, as individuals could choose not to consent, and it seems likely that the health professionals who recruited them knew to which arm they were allocated to, selection bias could have been introduced	High It will not have been possible to blind participants or personnel	Unclear The protocol states that AF will be confirmed by 'ECG according to the clinical protocol' (p. 5)[81] but it is unclear who interpreted this ECG (and whether or not they were blinded); in addition, prior testing before referral for ECG would have been unblinded	Unclear The ClinicalTrials.gov record states that 6990 individuals were recruited and the abstract includes data for 6971 people (99.7% of participants). However, the numbers are very different from the 12,870 individuals that the protocol suggests will be recruited	Unclear The results of this study have not yet been published in full	High It is possible that to be eligible for inclusion in the control group individuals had to be symptomatic. If this is the case the entry criteria for the two arms would have been different *Baseline outcome measures* Low A previous diagnosis of AF was an exclusion criterion *Baseline characteristics* Unclear Not reported *Contamination* Unclear

continued

TABLE 10 Risk-of-bias assessments for the five included RCTs (*continued*)

Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias(es)
							'Researchers will be assigned to one of the study groups by stratified randomisation based on the centre and the type of health professional' (p. 4).[81] Health professionals in the same centre could have been assigned to different groups so contamination is a possibility
STROKESTOP[26,40]	Low	Low	High	High	NA	Unclear	
	'A computerised 1 : 1 randomisation was performed in the 75- to 76-year-old population with stratification for sex, year of birth, and region' (p. 2177)[40]	'A computerised 1 : 1 randomisation was performed in the 75- to 76-year-old population with stratification for sex, year of birth and region' (p. 2177).[40] As randomisation was central and performed on all 75- to 76-year-olds it could not have been subverted	It will not have been possible to blind participants or personnel	Only participants in the intervention (screening group) received the hand-held ECG so cases identified by this means will have obviously been from the intervention group	No outcome data have been published for the control group and a comparison between the intervention group and the control group has not been made. In the interim publication AF status is reported for people who attended screening only (7173, whereas 13,331 were invited and 14,387 were randomised)	Only interim results have been published so far. AF incidence/ prevalence was not an outcome mentioned in the protocol	

Study	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias(es)
EARLY[83]	Unclear How randomisation was performed was not reported: 'patients were randomised, with a 1 : 1 allocation ratio' (p. 2)[83]	High Although individuals were randomly preselected for randomisation, not all participants were contacted, with 262 of the 2000 intervention group individuals not contacted and 1449 of the 2000 control group individuals not contacted. Individuals in the intervention arm were contacted post randomisation to invite them to participate (they could decline) but this did not occur in the control group	High It will not have been possible to blind participants or personnel	Unclear It is unclear exactly how cases were ascertained, but it seems possible that it was carried out by reviewing the electronic medical records system	High No intention-to-treat analysis was performed or could be calculated from the data presented. Reasons for exclusion were unbalanced between the two groups. People could decline participation in the intervention group (and were therefore excluded from the analyses) but not the control group	Unclear No protocol identified	

DOFA-AP, Detección Oportunista de Fibrilación Auricular en Atención Primaria; EARLY, Early diagnosis of Atrial fibrillation: a Randomized triaL in primarY care; FAMDAP, Effectiveness of Early Detection of Atrial Fibrillation; N/A, not applicable.

Study	Interventions	Comparator	Sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other biases	Baseline outcome measures	Baseline characteristics	Contamination
SAFE[19,84]	Population and opportunistic screening	No screening	+	+	–	?	+	+		?	+	?
STROKESTOP[26,40]	Population screening	No screening	+	+	–	–		?				
Morgan 2002[82]	Population screening	Opportunistic screening	+	+	–	–	–	+	–			
DOFA-AP[81]	Opportunistic screening	No screening	+	?	–	?	?	?	–	+	?	?
EARLY[83]	Targeted screening	No screening	?	–	–	?	–	?				

(Cluster RCTs: Baseline outcome measures, Baseline characteristics, Contamination)

FIGURE 15 Risk-of-bias assessment of all included studies. Blank cells indicate fields that were not applicable.

TABLE 11 Number of newly diagnosed AF cases among individuals without AF at baseline

Study	Screening strategy	Cases	No. randomised	Proportion (95% CI) (%)	OR (95% CI; p-value)	NNS
SAFE[19,84]	Control	47	4547	1.03 (0.71 to 1.29)		
	Systematic population screening	74	4594	1.61 (1.24 to 1.96)	1.57 (1.09 to 2.26; $p = 0.017$)	173
	Systematic opportunistic screening	75	4593	1.63 (1.24 to 1.96)	1.59 (1.10 to 2.29; $p = 0.013$)	167
	Targeted screening	53	4594	1.15 (0.8 to 1.4)	1.12 (0.75 to 1.66; $p = 0.58$)	833

This SAFE study examined systematic population screening and systematic opportunistic screening compared with no systematic screening (usual practice). Both systematic population screening and systematic opportunistic screening increased the odds of detecting new cases of AF among individuals without known AF at baseline compared with usual practice (systematic population screening OR 1.57, 95% CI 1.09 to 2.26; $p = 0.017$; systematic opportunistic screening OR 1.59, 95% CI 1.10 to 2.29; $p = 0.013$). The NNS for one screened individual with newly diagnosed AF to be identified by 12 months was 173 individuals for systematic population screening and 167 individuals for systematic opportunistic screening.

The study also modelled a targeted screening strategy, considering what would happen if only individuals at high risk in the systematic population screening arm were invited for screening. Targeted screening of high-risk individuals had a similar odds of detecting AF as usual practice (OR 1.12, 95% CI 0.75 to 1.66; $p = 0.58$), and the NNS was 833 individuals. These numbers were calculated from the number of people identified through screening who were at high risk plus the number of people in the systematic screening arm who were identified as having AF outside the screening programme. As this was a modelled outcome, it is possible that other cases of AF in low-risk individuals who were detected by the screening programme

would have been identified outside the screening programme if a systematic population screening programme had not been in place and therefore the odds of detecting AF in a targeted screening programme could be slightly higher.

Results from Morgan and Mant,[82] who compared systematic opportunistic screening with systematic population screening, could not be used because the number of people with AF at baseline was not reported. However, if we assume that the number of people with a previous diagnosis of AF was balanced between the groups we can assess the relative merits of systematic opportunistic screening compared with systematic population screening. Of the 1502 people randomised to systematic opportunistic screening there were seven newly identified cases of AF over the 6 months of follow-up, whereas of the 1499 people randomised to systematic population screening there were 12 new cases of AF (OR systematic population vs. systematic opportunistic 1.72, 95% CI 0.68 to 4.39; $p = 0.25$).

Although the SAFE study provides strong evidence that screening identifies more cases of AF than usual care, neither the SAFE study[19,84] nor the study by Morgan and Mant[82] provides strong evidence that one screening approach is superior to another.

Prevalence of atrial fibrillation at baseline
The prevalence of AF in the absence of screening could be calculated from only one RCT (the SAFE study[19,84]). Averaging across the study arms, the prevalence of AF before screening was 7.22% (95% CI 6.80% to 7.63%).

Prevalence of diagnosed atrial fibrillation at the end of screening and the change in prevalence of diagnosed atrial fibrillation
The change in prevalence of diagnosed AF after screening could be calculated only from the one RCT from which the baseline prevalence of AF could be calculated (the SAFE study).[19] The prevalence of diagnosed AF increased after the trial in all arms, including the control arm (*Table 12*), reflecting incident AF over the trial period. Compared with the control arm, the prevalence of diagnosed AF changed by approximately an additional 50% in the systematic opportunistic and systematic population screening arms, but targeted screening saw a similar change in diagnosed prevalence to that in the control arm.

Acceptability of systematic population screening programmes
In screening programmes that sent invitations to a defined population, one measure of the acceptability of screening is the uptake of screening among those invited. *Table 13* shows the uptake in studies that investigated systematic population and systematic targeted screening (number who attended screening divided by the number sent invitations). Studies with methodological or reporting limitations that precluded inclusion in other parts of this review have been included here. The fact that patients with a previous diagnosis of AF were not excluded from the SAFE,[19,84] STROKESTOP[26,40] and Morgan and Mant[82]

TABLE 12 Prevalence of diagnosed AF after systematic screening and change in prevalence of diagnosed AF

Study	Screening strategy	Baseline prevalence (*n* with AF/*n* randomised) (%)	Prevalence after screening (*n* with AF/*n* randomised) (%)	Change (%)
SAFE[19,84]	Control	389/4936 (7.88)	436/4936 (8.83)	0.95
	Systematic opportunistic screening	340/4933 (6.89)	415/4933 (8.41)	1.52
	Systematic population screening	339/4933 (6.87)	413/4933 (8.37)	1.50
	Targeted screening	339/4933 (6.87)	392/4933 (7.95)	1.07

TABLE 13 Uptake of systematic screening

Study	Systematic population, % (number who attended screening/number sent invitations)	Systematic targeted, % (number who attended screening/number sent invitations)
SAFE[19,84]	53.2 (2357/4433)	–
STROKESTOP[26,40]	53.8 (7173/13,331)	–
Morgan[82]	73.3 (1099/1499)	–
EARLY[83]	–	52.1 (463/888)

studies should be considered when interpreting these results. With the exception of the study by Morgan and Mant,[82] uptake of screening was fairly consistent across the studies, with approximately 50% of those invited attending for screening. In *Chapter 5* (see *Uptake of systematic population screening*) we report uptake rates for single-arm studies as well as RCTs and combine the results in a meta-analysis.

It is more difficult to judge the acceptability of systematic opportunistic screening, as systematic opportunistic screening relies on the person visiting the GP/screening centre, the nurse/GP offering the screening test and the individual consenting to it. In the SAFE study, 69.2% (3278/4738) of people randomised to systematic opportunistic screening who had their notes flagged had their pulse palpated.[19,84] Of the 361 people found to have an irregular pulse, 238 (65.9%) attended the screening clinic, where a 12-lead ECG was performed, although it should be noted that, of the 123 individuals who had an irregular pulse who did not attend for an ECG, 56 (45.5%) had already been diagnosed with AF.[19,84] The SAFE study also reports that there was variation at the practice level in both the proportion of individuals who had a pulse taken (from 33.5% to 93.0%) and the uptake of ECGs when the pulse was found to be irregular (from 14.3% to 87.5%).

In the study by Morgan and Mant,[82] 30.5% (439/1437) of people randomised to systematic opportunistic screening who had their notes flagged had their pulse palpated. In this study, the number of individuals who had their pulse palpated also varied by practice, from 8.0% in one practice to 52.1% in another. In both studies it is unclear how many eligible individuals visited their GP.

Only the SAFE study[19,84] reported the reasons why individuals refused to participate in systematic screening. Of the 904 individuals who replied to decline screening, 38.9% gave no reason, 9.4% replied 'can't get to surgery', 9.2% replied that they were 'already in NHS', 7.9% had 'had a recent ECG', 6.9% did not attend because of illness, 6.4% did not attend because of 'old age', 4.2% were 'not interested', 3.0% had 'mental health problems' and the remainder gave other reasons. All individuals undergoing an ECG in the SAFE study were asked to complete a screening acceptability questionnaire. Of those who completed the questionnaire, 94.4% felt that the letter/information sheet explained the tests properly, 3.7% considered screening to be inconvenient, 95.4% thought that health screening was important, 4.8% would have liked someone to discuss it more with first, 3.2% would have liked to talk about the tests with a GP first and 4% would have liked to come to a clinic appointment for more information.

The SAFE study recorded private individual costs associated with attending for screening, reporting a mean cost of £3.13 (95% CI £2.97 to £3.29) (2005 price year). Converting to 2016 prices (inflation factor 1.37, source ONS)[87] gives a mean cost of £4.29 (95% CI £4.07 to £4.51).

Characteristics of all screened individuals and screen-identified atrial fibrillation patients

Of the five included RCTs, four[19,40,82,83] reported the characteristics of the individuals who participated in the screening processes and the individuals who were identified as having AF through screening. The reported characteristics varied across studies, for both the individuals invited to screening and those identified as

having AF through screening. *Tables 14* and *15* show the characteristics of all screened individuals and screen-detected AF patients in the included studies respectively.

The mean age and sex distributions of the individuals invited to screening were comparable in the three studies that reported these individual characteristics. None of the reported characteristics of the patients identified to have AF through screening was comparable across the studies.

Adverse events with systematic screening

Few data were reported on adverse events associated with systematic screening, the rate and severity of complications or adverse events associated with the ECG or other forms of AF testing, psychological distress, change in quality of life and impact on well-being measured using a validated scale. The SAFE study[19,84] was the only study that reported these outcomes.

A random sample of individuals randomised to the screening arms of the trial was sent the postal version of the EuroQol-5 Dimensions (EQ-5D) (to measure quality of life)[88] and the shortened Spielberger anxiety questionnaire[89] on study entry and again at the end of the screening period (approximately 17 months later), when it was also sent to all participants who had screened positive. In addition, all participants were asked to complete the Spielberger anxiety questionnaire immediately after screening.

The EQ-5D scores were similar across the systematic population screening and systematic opportunistic screening arms at baseline and also at 12 months' follow-up. Similar results were found for anxiety scores.

TABLE 14 Characteristics of individuals invited to screening

Characteristic	SAFE[19,84]	STROKESTOP[26,40]	Morgan[82]	DOFA-AP[81]	EARLY[83]
Mean age (years)	73.5	–	75.5	–	69
Mean weight (kg)	–	79.2	–	–	–
Mean BMI (kg/m²)	–	26.7	–	–	–
Mean CHA_2DS_2-VASc score	–	3.6	–	–	–
% female	53.1	53.7	58.8	–	51
% hypertensive	–	49.7	–	–	72
% diabetes	–	11.1	–	–	18
% CHF	–	3.4	–	–	15
% stroke/TIA	–	9	–	–	–
% vascular disease	–	9.4	–	–	–
% IHD	–	–	–	–	11
% VHD	–	–	–	–	5
% on aspirin	–	22.7	–	–	–
% on anticoagulants	–	8.7	–	–	–
% other characteristics	PRA 11, previous AF 7.2	Previous AF 9.3	–	–	One RF 38, two RFs 44, three RFs 15, four RFs 1.7, five RFs 0.2

BMI, body mass index; CHF, congestive heart failure; IHD, ischaemic heart disease; PRA, pulse reading abnormality; RF, risk factor; VHD, valvular heart disease.

TABLE 15 Characteristics of patients identified as having AF through screening

Characteristics	[a]SAFE[19,84]	[b]STROKESTOP[26,40]	[c]Morgan[82]	DOFA-AP[81]	[b]EARLY[83]
Mean age (years)	–	–	–	–	–
Mean weight (kg)	–	81.4	–	–	–
Mean BMI (kg/m^2)	–	27.3	–	–	–
Mean CHA$_2$DS$_2$-VASc score	–	3.5	–	–	–
Mean CHADS$_2$ score[d]	1.4	–	–	–	–
% female	–	45.4	–	–	18.2
% hypertensive	37.6	52.1	57	–	63.6
% diabetes	9.4	13.3	12	–	36.4
% CHF	11.4	2.8	–	–	9.1
% stroke/TIA	7.4	9.6	21	–	–
% vascular disease	–	14.3	–	–	–
% IHD	–	–	–	–	36.4
% VHD	–	–	–	–	9.1
% on aspirin	–	24.8	39	–	–
% on anticoagulants	–	2.3 (oral)	25	–	–
% other characteristics	–	–	One RF 70, previous PRA 49, previous AF 82	–	One RF 9.1, two RFs 36.4, three RFs 54.5

BMI, body mass index; CHF, congestive heart failure; IHD, ischaemic heart disease; PRA, pulse reading abnormality; RF, risk factor; VHD, valvular heart disease.
a   The SAFE study characteristics are based on 149 cases.
b   Denotes characteristics of only new AF cases.
c   Characteristics for the study by Morgan and Mant[82] are for the systematic population screening arm only.
d   CHADS$_2$ is constructed as a sum of the following risk factors: cardiac failure, hypertension, age, diabetes, stroke (× 2).

However, based on the end-of-study questionnaire, individuals who were screen positive had a higher mean anxiety score (38.1, 95% CI 35.9 to 40.4) than individuals who were screen negative (34.6, 95% CI 32.4 to 36.8). EQ-5D scores were also significantly different, with screen-positive individuals reporting a lower quality of life score.

Cost-effectiveness of screening
Two RCTs[19,82] reported the cost-effectiveness of systematic screening. The SAFE study[19] reported the incremental cost per additional newly diagnosed AF case for systematic opportunistic, targeted and systematic population screening relative to no screening. The Morgan and Mant study[82] reported the incremental cost per additional newly detected AF case for systematic population screening relative to systematic opportunistic screening. The results are shown in *Table 16* in 2015 prices (using inflation factor 1.67 to convert the costs in the SAFE study[19] from 2003 prices and inflation factor 1.77 to convert the costs in the Morgan and Mant[82] study from 2002 prices (source ONS Consumer Price Index: Medical Services).[87] It should be noted that the SAFE study[19] conducted a detailed cost assessment, whereas the Morgan and Mant[82] study calculated the cost for the time of a short appointment with a nurse as £6 (2002 prices), which is £10.62 in 2015 prices. The SAFE study also included costs incurred by individuals attending for screening.

Subgroup and sensitivity analyses
The SAFE study[19,84] provides results by sex and age (65–74 years and ≥ 75 years). *Table 17* shows the proportion of newly diagnosed AF cases among those undiagnosed at baseline by age and sex. In the control arm, the rate of newly diagnosed AF was similar for men and women. For men, all screening

TABLE 16 Cost-effectiveness of systematic screening strategies: study incremental cost (£, 2015 prices) per additional case detected

Study	Systematic opportunistic screening vs. no screening	Systematic population screening vs. no screening	Targeted screening vs. no screening	Systematic population screening vs. systematic opportunistic screening
SAFE[19,84]	562	2528	5879	Dominated (costs an additional £52,527, but 22 fewer cases detected)
Morgan[82]	NA	NA	NA	329 (95% CI 245 to 531)

NA, not applicable.

TABLE 17 Proportion of newly diagnosed AF cases among those undiagnosed at baseline and ORs compared with the control arm for male/female and age 60–74 and ≥ 75 years subgroups[a]

Subgroup	Control	Systematic opportunistic screening	Systematic population screening	Targeted screening
Sex *Men*				
Proportion, *n/N* (%)	16/1896 (0.8)	38/1947 (2.0)	44/1968 (2.2)	32/1968 (1.6)
OR versus control (95% CI)		2.31 (1.28 to 4.16)	2.64 (1.48 to 4.71)	1.92 (1.05 to 3.52)
Women				
Proportion, *n/N* (%)	31/2651 (1.2)	37/2646 (1.4)	30/2626 (1.1)	21/2626 (0.8)
OR versus control (95% CI)		1.19 (0.73 to 1.93)	0.97 (0.58 to 1.61)	0.68 (0.39 to 1.19)
Women versus men				
Ratio of ORs[b]		0.50 (0.24 to 1.05)	0.36 (0.16 to 0.75)	0.34 (0.15 to 0.78)
Age (years) *65–74*				
Proportion, *n/N* (%)	18/2479 (0.7)	31/2637 (1.2)	30/2573 (1.2)	28/2573 (1.1)
OR versus control (95% CI)		1.61 (0.9 to 2.9)	1.6 (0.89 to 2.88)	1.49 (0.82 to 2.71)
≥ 75				
Proportion, *n/N* (%)	29/2068 (1.4)	44/1956 (2.2)	44/2021 (2.2)	35/2021 (1.7)
OR versus control (95% CI)		1.6 (0.99 to 2.57)	1.55 (0.96 to 2.49)	1.23 (0.75 to 2.02)
≥ 75 versus 65–74				
Ratio of ORs[b]		0.97 (0.45 to 2.00)	0.99 (0.44 to 2.05)	0.84 (0.37 to 1.84)

a Ratios of ORs comparing subgroups are given with 95% CIs. All results are from the SAFE study.
b Ratio of ORs: posterior medians and 95% CrIs from a Bayesian analysis.

strategies had a higher odds of newly diagnosed AF than usual practice and this was higher for systematic opportunistic and systematic population screening than for targeted screening. For women, there is no clear evidence that that there are any differences in new AF diagnoses under any of the screening strategies compared with usual practice. The ratio of ORs for women compared with men measures the 'interaction' effect of sex on the relative efficacy of the different screening arms. There is clear evidence of an interaction effect, with the ORs for the screening strategies compared with usual practice much closer to 1 for women than for men. This effect is stronger in the systematic population screening and targeted screening arms and less strong in the systematic opportunistic screening arm.

For age, there is clearly an increase in the proportion of newly diagnosed AF cases in the ≥ 75 years subgroup compared with those aged 65–74 years. However, the relative effectiveness (ORs) for the different screening interventions compared with the control was very similar for the different age ranges, with ratios of ORs estimated as close to 1, indicating no interaction between screening efficacy and age.

These analyses were also performed in the Cochrane review, which found strong evidence of a difference between sexes for systematic population screening ($\chi^2 = 6.64$, $p = 0.001$) but not for systematic opportunistic screening ($\chi^2 = 2.95$, $p = 0.09$). There was no strong evidence of a difference in AF detection between the two age groups. These findings are in line with our results.

There were too few data to perform any other subgroup or sensitivity analyses.

Discussion

Summary of findings

Both systematic population and systematic opportunistic screening strategies were found to be effective compared with no screening. Both systematic population and systematic opportunistic screening had similar efficacy with an estimated 170 individuals needed to screen to detect one additional AF case compared with usual practice. Systematic screening targeted to high-risk individuals found a similar number of new cases as usual practice. One reason for this may be that individuals with comorbidities are more likely to undergo an increased number of visits to primary care and, as part of investigating and managing these conditions, will be screened in earnest or by default for AF.

Uptake of systematic population screening was typically around 50% although it was as high as 70% in one study. There was variability in uptake between practices, with uptake being between 22% and 70% in the SAFE study.[19] Reasons for not attending for screening were varied, although older age and decreased mobility were common themes. The proportion of individuals having their pulses checked under systematic opportunistic screening varied across studies (30% and 66%) and between practices within studies (from 8% to 93%). The proportion of individuals consulting with their GP was not reported and so it is unclear how much these uptake rates are driven by consultation rates, GPs offering pulse palpation and uptake of pulse palpation by individuals. Of those with an irregular pulse who did not have a previous diagnosis of AF, approximately 18% did not attend for an ECG test, although again this was variable across practices.

Subgroup analyses based on the SAFE study[19,84] indicated that both systematic population and systematic opportunistic screening were more effective for men than for women, and this was especially the case for systematic screening. The Cochrane review[49] discusses possible explanations for this, concluding that 'differences in effect observed between the subgroups of men and women could be due to a combination of higher prevalence and greater rates of participation among men' (p. 14).[49] The efficacy of screening was not found to vary with age, despite AF prevalence being strongly associated with age group.

The findings from the subgroup analyses have implications for a screening programme, in particular whether such a programme would be effective in women and whether or not there are interventions that could encourage women to participate in screening. Reasons given for not attending for screening included poor mobility, thus provision of transport to and from practices may increase participation. Systematic opportunistic screening (when a journey to the GP has already been arranged) appeared to be more effective for women, and this may be a consideration in determining the format of a population-based screening programme. Most evaluations of population-based screening have included individuals aged ≥ 65 years. The results of the subgroup analysis for age suggest that the relative efficacy of screening is likely to be unchanged if the age threshold is increased. However, the absolute rate of detection of new cases (per person screened) and cost-effectiveness are likely to increase with age if the proportion of screened individuals with undiagnosed AF increases with age.

Strengths and limitations

Although five studies were identified, because of studies having reported only interim results or concerns over methodology, only two had data extracted, with only one reporting the number of people with AF at baseline and therefore used to study the primary outcome. The SAFE study[19,84] therefore remains the primary source of evidence for a comparison of different screening strategies. This situation should change, as three ongoing studies[26,40,85,86] are due to report within the next 3 years. In particular, the STROKESTOP study[26,40] aims to report the impact of screening on long-term outcomes (stroke and mortality), for which there is currently no robust evidence.

The SAFE study, on which most of our results are based, was a well-conducted study at low risk of bias on all domains except blinding of participants/personnel (which is not possible) and blinding of outcome assessment (unclear risk of bias). One issue with the SAFE study was that the baseline prevalence of AF was slightly higher in the control group than in the screening group. This could potentially introduce bias because, if more AF cases were previously diagnosed in the control practices, there may be fewer cases that could be diagnosed subsequently through screening or routine care. This was explored in the SAFE study using an individual participant analysis controlling for differences in baseline prevalence, which showed that the conclusions of the SAFE study were robust to adjustments to baseline prevalence.[19,84]

The SAFE study[19,84] found the incidence and 12-month prevalence of diagnosed AF to be similar between systematic opportunistic and systematic population screening. However, it should be noted that a large proportion of AF cases in the systematic opportunistic arm were detected outside the screening programme (i.e. in primary care but not as a result of activating the flag). Of the 74 new cases of AF identified in the systematic population screening arm, 22 (30%) were diagnosed outside the screening programme over the 12 months of the study. This proportion was greater for the systematic opportunistic screening arm, in which of the 75 newly identified cases, 44 (59%) were diagnosed outside the screening programme. One explanation for this is that GPs in the systematic opportunistic screening arm changed their usual practice to check for AF more frequently (the 'Hawthorne effect', where behaviour changes when under observation). If this is the case, then the full benefits seen in the systematic opportunistic arm may not be realised outside the context of a RCT. In contrast, Morgan and Mant[82] concluded that systematic population screening is superior to systematic opportunistic screening, although there is no strong evidence to support this based on comparisons of new cases of AF. The Morgan and Mant study[82] is complicated by the fact that an unreported number of patients with a previous diagnosis of AF were included. Also, only approximately 30% of people randomised to systematic opportunistic screening had their pulse assessed and, of those with an irregular pulse not referred for an ECG, it was assumed that a clinical diagnosis was made (which could have introduced detection bias).

In a screening programme it is likely that screening is offered repeatedly at given intervals (e.g. every 5 years). However, none of the RCTs identified in our review considered repeated screening, only a one-off screen. We therefore do not have any evidence on the effectiveness of such a screening strategy.

Findings in the context of previous research

A systematic review[90] of cohort studies (including 26 prospective cohorts, two retrospective cohorts and two RCTs) that used an ECG or pulse palpation to identify AF in the general ambulant population (excluding studies performed in people with specific comorbidities such as stroke, hypertension or cardiac disease) found that the overall incidence of previously unknown AF was 1.0% (95% CI 0.89% to 1.04%), increasing to 1.4% (95% CI 1.2% to 1.6%) in those aged ≥ 65 years. The authors observed that incidence was similar in GP/outpatient clinics and community-based populations. The AF prevalence was 2.3% (95% CI 2.2% to 4.4%), increasing to 4.4% (95% CI 4.1% to 4.6%) in those aged ≥ 65 years. The denominator for these estimates was the number screened and not the number randomised (even for the RCTs), as used in our review. Despite this difference, the SAFE study[19,84] found the 12-month incidence of unknown AF, with either opportunistic or systematic screening, to be a comparable 1.6% for individuals aged ≥ 65 years. However, the prevalence of AF after screening in the SAFE study[19,84] was 8.8% in the control group and 8.4% in the screening groups, much higher than in the systematic review by Lowres et al.[90]

Chapter 5 Methods for the economic evaluation of atrial fibrillation screening

Introduction

In this chapter we set out the methods used in the economic evaluation of systematic screening strategies for AF. We begin by defining the decision question that our economic model addressed, the modelling approach taken and the outputs reported. We then present the results of a review of previous economic models evaluating population-based screening strategies for AF. This is followed by a description of the structure of our model and the inputs required to populate it. Next, we describe a supplementary literature review of observational studies that we used to identify evidence on the natural history of AF, to provide inputs to the model, and present the results from this review. We then summarise all of the evidence used for inputs to the model and describe the sensitivity analyses that we conducted.

Decision question

The economic evaluation aimed to assess the cost-effectiveness of different population-based screening strategies for AF (including no screening) in a primary care setting in England and Wales. We took the perspective of the NHS, excluding costs incurred by individuals, employers and other agencies.

Screening strategies

We assumed the presence of an established screening programme, with cohorts of individuals invited to attend for screening when they reach a certain age.

We compared the following population-based screening strategies in the base case:

- no systematic screening (current practice)
- single systematic population screen, inviting all individuals meeting the screening criteria
- single systematic opportunistic screen, inviting individuals meeting the screening criteria when they consult with their GP.

We did not include a targeted strategy in which high-risk individuals are screened, based on the findings from the SAFE study,[19] which found that this strategy did not perform any better than no systematic screening.

In the base case, we report the results for one-off screening at ages 55, 60, 65, 70, 75 and 80 years, exploring various different screening tests (see *Screening tests*).

In sensitivity analyses, we explored repeated screening every 5 years for various initial screening ages (55, 60, 65, 70, 75 and 80 years) and final screening ages (55, 60, 65, 70, 75 and 80 years), which gave a total of 21 repeat screening strategies. We assumed that the optimal screening test and recruitment method (systematic opportunistic or systematic population) would be used for the repeat screening.

Population

The target population for systematic screening was the general population in England and Wales of a given age without a diagnosis of AF.

Screening tests

We included all screening tests identified in the systematic review (see *Tables 2* and *3*) for which we had sensitivity and specificity estimates compared with a gold standard (12-lead ECG interpreted by at least one cardiologist). The exception was the two-stage screening tests because we assumed that a diagnostic test would be given to those testing positive on a screening test, and so in effect all of the strategies represent two-stage screening. We acknowledge that this gold standard is appropriate only for persistent/permanent AF; however, we felt that persistent/permanent AF is likely to be the primary target of a population-based screening programme. Furthermore, estimates of the costs and benefits of anticoagulation are based on trials that do not generally include paroxysmal AF patients. However, we acknowledge that some paroxysmal AF cases may be detected and treated and we included this in our model.

We assumed that all individuals with a positive screening test result would receive a 12-lead ECG diagnostic test to confirm the diagnosis. The 12-lead ECG is likely to be performed in primary care, with a diagnosis being made in primary care when the result is clear and referral to a cardiologist made in a proportion of cases when the result is unclear.

Screening tests were defined according to the technology used and the interpreter (including algorithms). The following diagnostic tests were compared in the model (see *Tables 2* and *3*):

- 12-lead ECG:

  - interpreted by GP
  - interpreted by nurse
  - interpreted by algorithm

- single-lead ECG (including iPhone and other hand-held devices):

  - interpreted by cardiologist
  - interpreted by GP
  - interpreted by nurse
  - interpreted by algorithm

- > 1- but < 12-lead ECG:

  - interpreted by cardiologist
  - interpreted by algorithm

- pulse palpation (various methods):

  - interpreted by nurse

- modified blood pressure monitor (various devices):

  - interpreted by algorithm

- photoplethysmography:

  - interpreted by algorithm.

Although we have included tests interpreted by a cardiologist (as they were included in the DTA review), we note that cardiologists are unlikely to interpret screening tests in a primary care setting.

Outcomes of the economic evaluation

The economic model was evaluated using a probabilistic analysis in which the model inputs were simulated from distributions that reflected parameter uncertainty in the model.[91] Results from the model are averages (means) over these simulations. We present the mean total incremental costs, mean total incremental quality-adjusted life-years (QALYs) and mean incremental net benefit (INB) at a willingness to pay of £20,000 per QALY[92] for each strategy relative to no screening (current practice). The INB is calculated as the incremental QALYs multiplied by £20,000 minus the incremental costs. Positive values of the mean INB indicate strategies that are cost-effective at the £20,000 per QALY threshold. We aimed to maximise the mean INB. We also present the incremental cost-effectiveness ratio (ICER), which is the mean incremental costs divided by the mean incremental QALYs. The INB and ICER are related, so that a positive mean INB would have an ICER of < £20,000 and would be considered cost-effective at a willingness to pay of £20,000 per QALY. Conversely, a negative mean INB would have an ICER of > £20,000 and would not be considered cost-effective at a willingness to pay of £20,000 per QALY. Uncertainty in the optimal strategy is represented using cost-effectiveness acceptability curves (CEACs) and by displaying the simulations in the incremental cost-effectiveness plane (incremental QALYs plotted against incremental costs). The CEAC plots the probability that each strategy is most cost-effective against willingness to pay per QALY. For clarity we omit strategies from the CEAC plots that have a < 10% probability of being most cost-effective at any willingness-to-pay threshold. To compare different screening strategies by age at first screen, and number of repeated screens, we multiply the per-person INB by the population size that will benefit (for a given age cohort). Costs and outcomes over a lifetime time horizon were included because treatment for AF aims to reduce future events. Future costs and benefits are discounted at an annual rate of 3.5% in line with recommendations of the UK treasury.[93]

Review of previous economic models of screening strategies for atrial fibrillation

We searched the NHS Economic Evaluation (NHS EED) and HTA databases from inception until close (end of December 2014 for NHS EED and end of December 2015 for HTA databases) using the search terms Atrial Fibrillation AND Screening to identify economic evaluations of screening strategies. This identified 15 references, of which seven studies[19,21,25,94–97] were considered relevant after title and abstract screening (reviewed by NJW).

Because NHS EED stops at the end of December 2014, we ran Cochrane AF screening review searches[49] in EMBASE and MEDLINE with CRD economic filters[98] from January 2015 to the end of December 2015. This identified 261 records, of which eight papers[99–106] not previously identified were considered relevant after title and abstract screening and a further four papers[107–110] were highlighted as providing information on costs.

We also searched the Cost-Effectiveness Analysis (CEA) Registry on 19 January 2016, using the search term Atrial Fibrillation. This resulted in 76 hits but did not identify any other studies on the cost-effectiveness of AF screening. However, it did identify two studies on the cost-effectiveness of diagnostic tests for AF.[111,112]

The search for studies of the natural history of AF and AF screening (in EMBASE and MEDLINE, search dates from 1 January 2000 to 20 January 2016) (see *Model inputs*) identified three further cost-effectiveness studies.[113–115]

Finally, hand-searching reference lists of papers and reports that we were aware of identified three further studies.[116–118]

In total, after title and abstract screening we identified 27 studies for which full-text articles were obtained and screened. After screening the full texts, nine studies[19,21,25,96,99,100,116,117] of economic evaluations of population-based screening strategies were identified.

Figure 16 shows a diagram illustrating the literature searches.
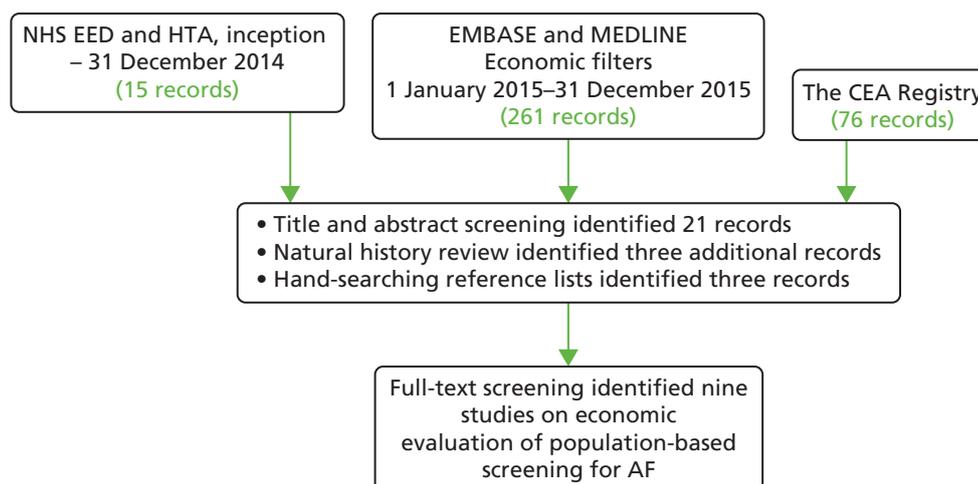
FIGURE 16 Flow chart for the literature search for studies of economic evaluations of population-based screening.

In the following section, we summarise the nine economic evaluations of population-based screening for AF, which are also described in *Table 18*. The quality of these economic evaluations was assessed using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist.[119] These assessments are provided in *Appendix 8*.

Summary of previous economic evaluations of screening for atrial fibrillation

Hobbs *et al.*[19]

Hobbs *et al.*[19] (the SAFE study) present results from (1) a within-trial cost-effectiveness analysis that reports cost per additional true-positive AF case detected and (2) a longer-term economic model that captures costs and benefits beyond the follow-up period of the trial and reports cost per QALY. Both analyses assume screening with a 12-lead ECG interpreted by a cardiologist, although this is relaxed in sensitivity analyses.

Results from the within-trial analysis comparing systematic opportunistic screening, systematic population screening, screening targeted at high-risk individuals according to stroke risk and no screening have previously been described (see *Chapter 4*, *Cost-effectiveness of screening*). The systematic population and targeted screening strategies were dominated (cost more for the NHS and identified fewer new cases) by systematic opportunistic screening. The incremental cost per case detected for opportunistic compared with no screening was £337 (2003 prices).

The model-based analysis compared different screening intervals, screening tests and interpreters and screening methods (systematic opportunistic and systematic population screening) for 65-year-olds. A discrete event simulation model was used to describe patient pathways. If a patient is not identified through screening he or

TABLE 18 Summary of economic evaluations of population-based screening for AF identified in our review

Study	Setting	Screening methods	Screening tests	Model structure	Outcomes
Hobbs 2005[19]	UK general population, based on 50 general practices, West Midlands, UK	Systematic opportunistic screening when individuals consult with their GP. Systematic population screening of the general population. Aged ≥ 65 years	Compared different index tests (12-lead ECG, limb-lead rhythm strip ECG, single-lead ECG), interpreters (consultant, GP, nurse, automatic) and screening intervals	Discrete event simulation model	Cost per QALY with results presented for both a NHS and a societal perspective

TABLE 18 Summary of economic evaluations of population-based screening for AF identified in our review (*continued*)

Study	Setting	Screening methods	Screening tests	Model structure	Outcomes
Moran 2015[25]	Ireland, primary care	Systematic opportunistic screening when individuals present to their GP. Various age thresholds (base-case ≥ 65 years)	Pulse palpation with individuals with an irregular pulse referred for an ECG	Decision tree for screening outcomes followed by a Markov model for long-term costs and benefits	Cost per QALY with results presented for both payer and societal perspectives
Aronsson 2015[100]	Sweden, general population (two regions)	Systematic population screening of 75- and 76-year-olds	Intermittent hand-held ECG twice daily (or when palpitations) for 2 weeks	Decision tree for screening outcomes followed by a Markov model for long-term costs and benefits of screen-detected AF	Cost per QALY; societal perspective reported but only health service costs included
Lowres 2014[21]	Australia, 10 pharmacies (Sydney)	Systematic opportunistic screening of people visiting the pharmacies aged ≥ 65 years	Pulse palpation and hand-held iPhone-based single-lead ECG, interpreted by a nurse and cardiologist	Not reported explicitly	Cost per QALY and cost per stroke prevented from a health service perspective
Maeda 2004[96]	Japan general population	Systematic population screening of 65-year-olds annually until age 85 years	Compared a 12-lead ECG with pulse palpation followed by a 12-lead ECG	Decision tree for screening outcomes followed by a Markov model for long-term costs and benefits of screen-detected AF	Cost per QALY from a health service perspective
Lord 2013[116]	England and Wales	Not applicable	Suspected AF in primary care referred to a specialist for an ECG. Ambulatory monitoring may be used if paroxysmal AF is suspected after a negative ECG	Discrete event simulation model	Cost per QALY from a NHS perspective. No results reported comparing screening methods
Gordon 2012[118]	35 general practices in north-east Essex	Systematic opportunistic screening in individuals aged ≥ 65 years without known AF and attending for influenza vaccination	Pulse palpation followed by a 12-lead ECG in individuals with an irregular pulse	Decision tree	Cost per screen-detected case and cost of preventing one stroke in the following year from a NHS perspective
Rhys 2013[117]	Single general practice in Leek, UK	Systematic opportunistic screening in individuals aged ≥ 65 years attending for influenza vaccination	Pulse palpation followed by a 12-lead ECG in individuals with an irregular pulse	Decision tree	Cost per screen-detected case and cost of preventing one stroke in the following year from a NHS perspective
Aronsson 2015[99]	Not reported	1,073,741,824 different screening designs compared	Not reported	Not reported	Cost per QALY; perspective not reported

she may be diagnosed at a later date through routine care (on average 3 years) or after an event (e.g. ischaemic stroke). Incidence data were used to model the development of AF. The sensitivity and specificity of the tests were taken from the SAFE trial, along with the uptake of screening and costs of screening. It was assumed that all AF patients would be prescribed warfarin, although a sensitivity analysis was reported assuming that all patients received aspirin instead. Rate and rhythm control were not included in the model. The results showed that QALYs gained were very similar across all strategies (including no screening), with differences seen only in costs. However, differences in costs were small, with slightly lower costs for systematic opportunistic screening than for no screening.

Moran et al.[25]

Moran et al.[25] reported a HTA of a national screening programme for AF in primary care in Ireland. The model compared systematic opportunistic screening with no screening for men and women aged ≥ 65 years. The screening test was pulse palpation followed by an ECG read by a GP and an algorithm if an irregular pulse was detected. The uptake and effectiveness of screening were taken from the SAFE study,[19] prevalence and incidence were taken from the Irish Longitudinal Study on Ageing (TILDA)[120] and a Markov model was used to capture the costs and benefits of treatment over the long term (until age 90 years). Patients were assumed to take DOACs, warfarin, antiplatelets or no therapy in proportions estimated from the TILDA study and other routine data sources in Ireland and the UK. The ICER for screening compared with no screening was €20,271 per QALY. The ICER was sensitive to the start age of screening, varying from €50,578 if screening was started at age 50 years to €14,594 if screening was started at age 70 years. The ICER was also sensitive to the screening interval, decreasing as the interval between screens increased.

Aronsson et al.[100]

The study by Aronsson et al.[100] was a cost-effectiveness analysis of systematic population screening for AF in 75- and 76-year-olds in Sweden using intermittent ECG recording based on data from the ongoing STROKESTOP study.[40] Future consequences were tracked over a patient's lifetime using a Markov model, depending on CHA_2DS_2-VASc score and treatment. For patients eligible for anticoagulation it was assumed that a proportion were contraindicated (based on data from the STROKESTOP study) and the remainder used a DOAC (assumed to be apixaban). Use of warfarin rather than apixaban was considered in a sensitivity analysis. The model predicted a cost of €4313 per QALY gained and €6583 per stroke avoided compared with no screening.

Lowres et al.[21]

The study by Lowres et al.[21] was a prospective study evaluating the effectiveness and cost-effectiveness of community screening for AF in those aged ≥ 65 years visiting one of 10 pharmacies in Sydney, Australia. Screening was by pulse palpation and hand-held iPhone-based single-lead ECG interpreted by a nurse and cardiologist. Individuals with suspected AF were referred to their GP. The economic model followed a cohort of 65- to 84-year-olds for 10 years, using stroke risk information from studies on asymptomatic incidentally detected AF. Patients were assumed to be treated with warfarin. The model structure is unclear, although it is likely to be a decision tree for screening outcomes followed by a Markov model to reflect annual stroke and mortality risk and costs. Sensitivity and specificity of the iECG were taken from the authors' study. The ICER was €3142 per QALY and €15,993 per stroke prevented compared with no screening. Results were most sensitive to treatment adherence.

Maeda et al.[96]

This study compared three community-based screening strategies for 65-year-olds in Japan: annual screening with an ECG, annual screening with pulse palpation with referral for an ECG in patients with arrhythmias and no screening. A Markov model was used to estimate costs and outcomes for AF patients up to age 85 years. The two screening strategies gave similar results, with ICERs compared with no screening of approximately US$8000 and US$10,000 per QALY in men and women respectively. The results were sensitive to the proportion prescribed anticoagulants and the incidence of ischaemic stroke, with screening being more cost-effective as the proportion of patients prescribed anticoagulants increased and

also as the incidence of ischaemic stroke increased. Increasing the interval between repeat screenings to every 5 years decreased the costs but had very little impact on the QALYs gained, suggesting that 5-yearly repeated screening tests would be adequate.

Lord et al.[116]

This study developed a model to reflect the course of a cohort of AF patients diagnosed and treated following the NICE clinical guidelines for AF that were available at that time (CG36[121]). The model tracks care pathways for AF but does not compare the cost-effectiveness of screening strategies. Of interest to our project, the model includes a diagnostic pathway with outcomes modelled separately for true positives, true negatives and false negatives (figures 15 and 17 in Lord et al.[116]).

Other studies

Rhys et al.[117] and Gordon et al.[118] estimated the cost per stroke prevented over the coming 12 months based on primary data collected while screening individuals aged ≥ 65 years attending for influenza vaccination in UK general practice. The report by Aronsson et al.[99] was a conference abstract detailing the results of an economic evaluation of > 1000 million screening designs, concluding that the most cost-effective programme is an initial screen at age 75 years with a repeat screen at 80 years.

Summary

The study by Hobbs et al.[19] is the most relevant study to our study in that it used an economic model to obtain lifetime costs per QALY for different screening tests and strategies in a UK population. The study was conducted before DOACs were in routine use. The model therefore needs updating to incorporate this new class of therapy and also to reflect up-to-date evidence on the natural history of AF in a contemporary UK population. The HIQA model[25] is also very relevant, using recent evidence in an Irish population. However, the HIQA report considered only systematic opportunistic screening using pulse palpation. Most models used a decision tree for the screening part of the model followed by a Markov model for subsequent treatment, although the study by Hobbs et al.[19] used a discrete event simulation model. Our model provides a comparison of various screening tests and methods using contemporary evidence relevant to a UK population and including anticoagulation therapy options currently available in the NHS.

Model structure

We developed a model to compare the cost-effectiveness of population-based screening strategies, drawing on the models previously reported in the literature (described in the previous section) and also based on guidance from our expert advisory group. Our model has many features in common with previous models. An initial decision tree describes screening attendance, screening test findings and diagnostic test results. Then a Markov model tracks AF-related events and mortality conditional on anticoagulant therapy. Our previously published Markov model (DOAC model)[37] was based on a network meta-analysis of all trials comparing DOACs with warfarin (INR range 2–3) to capture the lifetime costs and benefits of anticoagulation therapy in AF.

Figure 17 illustrates the decision tree used to compare population (S1) and opportunistic (S2) screening with current practice (S0, no screening). The model structure, costs of care post diagnosis and utilities remain the same for each strategy; however, the sensitivity and specificity parameters depend on the screening test.

For systematic population screening, a cost is incurred for the screening invitation, which depends on the population size invited and the proportion of individuals who do not initially respond and who are sent a reminder letter. A certain proportion will not attend for screening and do not incur any further costs or receive any further health benefits. We assumed that individuals with a known AF diagnosis will not be invited for screening and that the cost of identifying and excluding these individuals using computer records is small and can be ignored. For opportunistic screening, we assumed that, during a 12-month period, a proportion of individuals who are flagged for screening will have an unrelated GP appointment and will be opportunistically tested during the appointment.
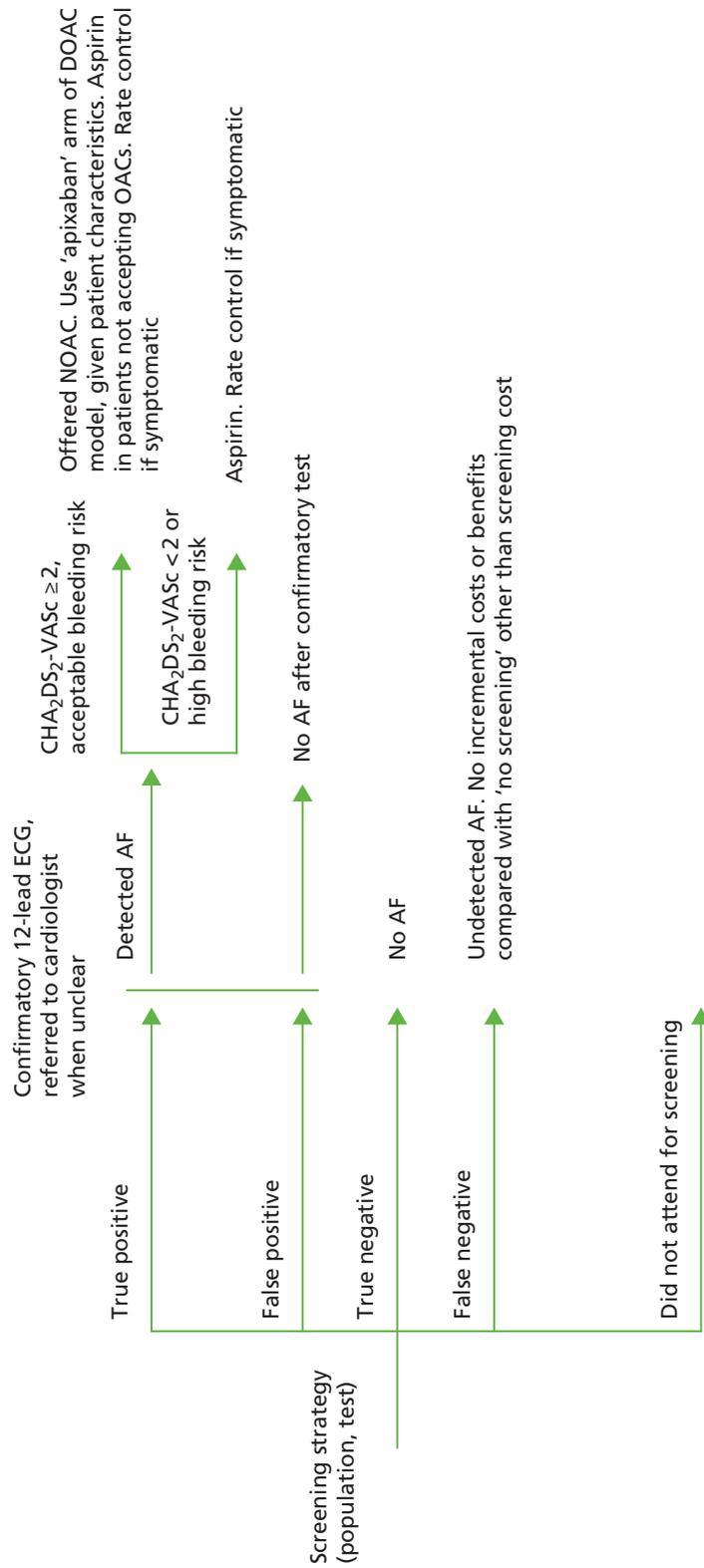
FIGURE 17 Schematic representation of the decision tree for a given screening strategy. The benefits and costs of anticoagulant therapy[37] are incurred for those with detected AF and a CHA_2DS_2-VASc score of ≥ 2.

Of those who are screened, all will incur the cost of the screening test. We assumed that a proportion of individuals with AF would have detectable AF at the time of the screening test. This proportion consists of individuals with persistent and permanent AF and a smaller proportion of paroxysmal AF individuals who happen to be in AF at the time of screening. We assumed that individuals who test positive on the screening test will receive a gold standard diagnostic test (12-lead ECG interpreted in primary care with cardiologist interpretation in a proportion of cases that are unclear), which is assumed to have 100% accuracy. We assumed that all of those offered a 12-lead ECG diagnostic test agree to have the test (but varied this assumption in sensitivity analysis) and that all unclear ECGs are resolved by cardiologist interpretation. Further details are provided in *Test costs*. We assumed no quality of life benefit or decrement as a result of administration of the screening test itself or the confirmatory ECG.

True positive on the initial screening test

Of those who are tested, a proportion will be true positives (screen-positive results that are confirmed using a 12-lead ECG gold standard diagnostic test), and these individuals may benefit from early detection of AF. All of these individuals will accrue the cost of the initial screening test plus a gold standard 12-lead ECG test interpreted by a GP, except a proportion of unclear cases that are additionally interpreted by a cardiologist.

For those individuals with a stroke risk CHA_2DS_2-VASc score of ≥ 2, anticoagulation therapy is recommended unless contraindicated (bleeding risk using the HAS-BLED score, dementia, epistaxis, frailty or patient preference). However, in a proportion of cases, individuals not contraindicated for anticoagulation are not prescribed anticoagulants by their GP. In the base case we assumed that all of the individuals not contraindicated and prescribed anticoagulation therapy use DOACs. We assumed that the most cost-effective treatment (apixaban) in the DOAC model[37] is used for all patients who receive DOACs, although it should be noted that similar results would be obtained using other DOACS.[37] In the sensitivity analyses we assumed that 50% use DOACs and the remainder use warfarin (INR range 2–3). We assumed that individuals with a CHA_2DS_2-VASc score of ≤ 1, in whom OACs are not recommended,[11] and patients contraindicated to, not prescribed or choosing not to take OACs, would receive aspirin. This was based on clinical advice, confirmed by Quality and Outcomes Framework (QOF) statistics,[3] that these patients will likely receive aspirin in practice, despite aspirin not being recommended in the latest NICE clinical guidelines.[11]

Regardless of anticoagulation therapy, we assumed that all symptomatic patients would receive a standard beta-blocker for rate control, which incurs a cost and improves quality of life benefit. The costs and benefits of rate control were assumed to be independent of anticoagulation therapy. Rhythm control is not currently recommended except for patients younger than those who would be invited to a screening programme, and so we did not include it in our model.

True negative on the initial screening test

Of those tested, a proportion will be true negatives on the initial screening test. These individuals incur the cost of the initial screening test but do not receive any further incremental benefits or costs from the screening strategy compared with current practice (S0, no screening).

False positive on the initial screening test

We assumed that individuals with a false-positive result on the initial screening test will be given a diagnostic 12-lead ECG test, which correctly identifies that they do not have AF (i.e. the confirmatory test is 100% specific). All of these individuals incur the cost of the initial screening test plus the cost of the diagnostic 12-lead ECG test interpreted by a GP and, in a proportion of unclear cases that are additionally interpreted by a cardiologist, an additional expense will be incurred. As with the true negatives, these individuals do not receive any further incremental benefits or costs from the screening strategy compared with current practice (S0, no screening). It should be noted that some false positives may be diagnosed with another condition (e.g. atrial flutter) and may receive some benefit from an AF screening programme. However, we do not capture this benefit in our model.

False negative on the initial screening test

Of those tested, a proportion will be false negatives on the initial screening test, that is, they have AF but are not referred for a diagnostic 12-lead ECG test and gain no benefits from treatment. This includes individuals with paroxysmal AF who are not in AF at the time of the screening test. These patients incur the cost of the initial screening test but no future incremental costs and benefits compared with current practice (S0, no screening).

Model inputs

In this section we describe the inputs used for the parameters required to populate the economic model. The sensitivity and specificity of the screening tests came from our systematic review of DTA studies (see *Chapter 3*, *Statistical modelling results*). The information on screening came from the studies identified in our systematic review of comparative studies of screening strategies (see *Chapter 4*), together with supplementary searches (see below) to identify observational studies of screen-detected AF. The cost-effectiveness of anticoagulation therapy is based on our recently published DOAC model.[37] The cost parameters came from our review of economic evaluations, NHS reference costs[122] and Personal Social Services Research Unit (PSSRU) unit costs[123] and primary studies reporting costs and resource use identified in our review. Supplementary searches were conducted to identify evidence sources for the remaining parameters in the model, primarily information on the natural history of AF and its management in the NHS.

We searched for longitudinal, observational and cross-sectional studies to identify, ideally in a population-based screening population:

1. the prevalence of AF in England and Wales by age and sex
2. population size eligible for screening by age and sex
3. uptake of systematic population screening and the proportion requiring a reminder letter
4. the proportion of eligible patients who have their pulses taken
5. the proportion of 12-lead ECGs interpreted by a GP that are unclear and require further interpretation by a cardiologist
6. the rate of diagnosis in the absence of screening
7. the proportion AF that is undiagnosed
8. the proportion of AF that is paroxysmal, persistent or permanent
9. the proportion of AF that is asymptomatic (by AF type)
10. disease progression through AF types
11. hazard ratios (HRs) for stroke and mortality risk by AF type and whether asymptomatic or not
12. the proportion of screen-detected AF with a CHA_2DS_2-VASc score of ≥ 2, previous stroke (or TIA), previous MI and moderate (HAS-BLED score of 2) or high (HAS-BLED score of ≥ 3) risk of bleeding
13. the proportion of eligible patients prescribed DOACs, warfarin and aspirin/no anticoagulation therapy and the proportion of patients contraindicated to anticoagulation therapy
14. incremental quality of life benefits and costs from rate control compared with no rate control.

We searched MEDLINE and EMBASE between 1 January 2000 and 22 January 2016 (search strategy is provided in *Appendix 9*). In total, 2126 references were identified after deduplication. After abstract and title screening, 134 references were identified as potentially relevant. For the prevalence of AF we restricted inclusion to UK or Irish populations, unless the study was conducted in a community or general practice screening setting. For all other model inputs we did not restrict the population. This identified 56 papers. In addition, we identified a further 22 studies as follows: (1) seven studies found in the reference lists of the Lowres *et al.*[90] and Xiong *et al.*[124] systematic reviews, (2) eight studies found in the economic evaluations review and citations in the identified studies, (3) three studies found in the NICE clinical guideline CG180[11] report and the associated consensus report (www.nice.org.uk/guidance/cg180/resources/nic-consesus-statement-on-the-use-of-noacs-243733501), (4) three studies found in the DOACs review[37] and associated review and (5) one recently published study brought to our attention by the HIQA team (P Moran, Health

Information and Quality Authority, Ireland, 2016, personal communication). After screening the full texts, we included 48 studies that reported data relevant to at least one of the model inputs listed above. A flow chart summarising the literature search is provided in *Figure 18*. In the following sections, we describe the evidence sources for each model input in turn and provide a rationale for the values and distributions assumed in our model.

Prevalence of atrial fibrillation in the UK by age and sex

The prevalence of AF strongly depends on both age and sex.[31] We searched for studies reporting prevalence in a general population (both diagnosed and undiagnosed) broken down by age and sex in Western European countries, with data collected from 2000 onwards. *Table 19* summarises the studies identified and *Figure 19* shows plots of prevalence against age for men and women separately. All studies show similar patterns, although there are some differences between studies, with, in particular, the study by Baena-Díez *et al.*[129] finding a lower prevalence of AF in older people than the other studies. We chose to use the prevalence estimates from Norberg *et al.*[2] as this was the most recent and largest study providing the most detailed results. The prevalence estimates from Norberg *et al.*[2] are shown in *Table 20*.

Population size eligible for screening at each age

The size of population eligible for screening at each age was estimated as the total number of men and women in the associated 5-year age range as reported by the Office for National Statistics (ONS)[130] for England and Wales in 2014 divided by 5. These estimates are reported in *Table 21*.

Uptake of systematic population screening

Seven studies reported the uptake of systematic screening in which individuals were invited to screening by letter. These studies are summarised in *Table 22*. Uptake of screening varied between 53% and 76% across studies/populations, with both the minimum and the maximum uptake seen in UK general practice populations. We pooled the data from all seven studies using a random-effects meta-analysis, reported in *Table 22*. However, uptake of screening is likely to depend on age. The only study that reported uptake by age is the SAFE study[19] (*Table 23*). The results for the < 75 years age group are in line with the results from the studies reported in *Table 22*, whereas uptake in the older age group is lower. We used the figures in *Table 23* to inform beta distributions for use in our model.
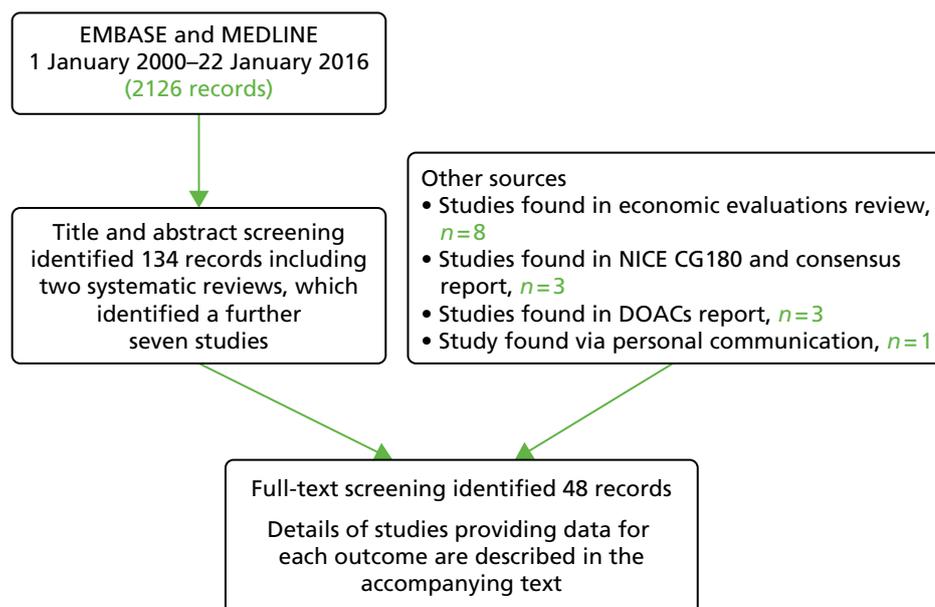
FIGURE 18 Flow diagram showing the literature search to find evidence on the natural history of AF.

TABLE 19 Summary of studies providing information on the prevalence of AF by age and sex in Western European populations

Study	Population	Comments
Fitzmaurice 2007[84]	Prevalence at 12 months following randomisation to screening (systematic opportunistic or systematic population screening); 50 UK general practice centres randomised in 2001 (n = 4915 and 4906 respectively)	
Schnabel 2012[125]	Population-based Gutenberg Health Study random sample of people from the Mainz region, from 2007 onwards (n = 5000)	Used in Irish HIQA HTA[25] of AF screening
Norberg 2013[2]	General population in Skelleftia, Sweden (2010), with extensive searching of multiple databases with ECG validation (n = 38,142)	Used in Public Health England report[1]
Frewen 2013[126]	Irish community sample from the TILDA study (n = 8175)	
Krijthe 2013[127]	Two general population cohorts in Rotterdam, the Netherlands (2002) (n = 6934)	
Gómez-Doblas 2014[128]	Spain, general population aged ≥ 40 years in 47 hospitals across 46 provinces (2010) (n = 8343)	
Baena-Díez 2014[129]	Pooled data from six population-based studies in Spain and nearby islands (1999–2011) (n = 8149)	
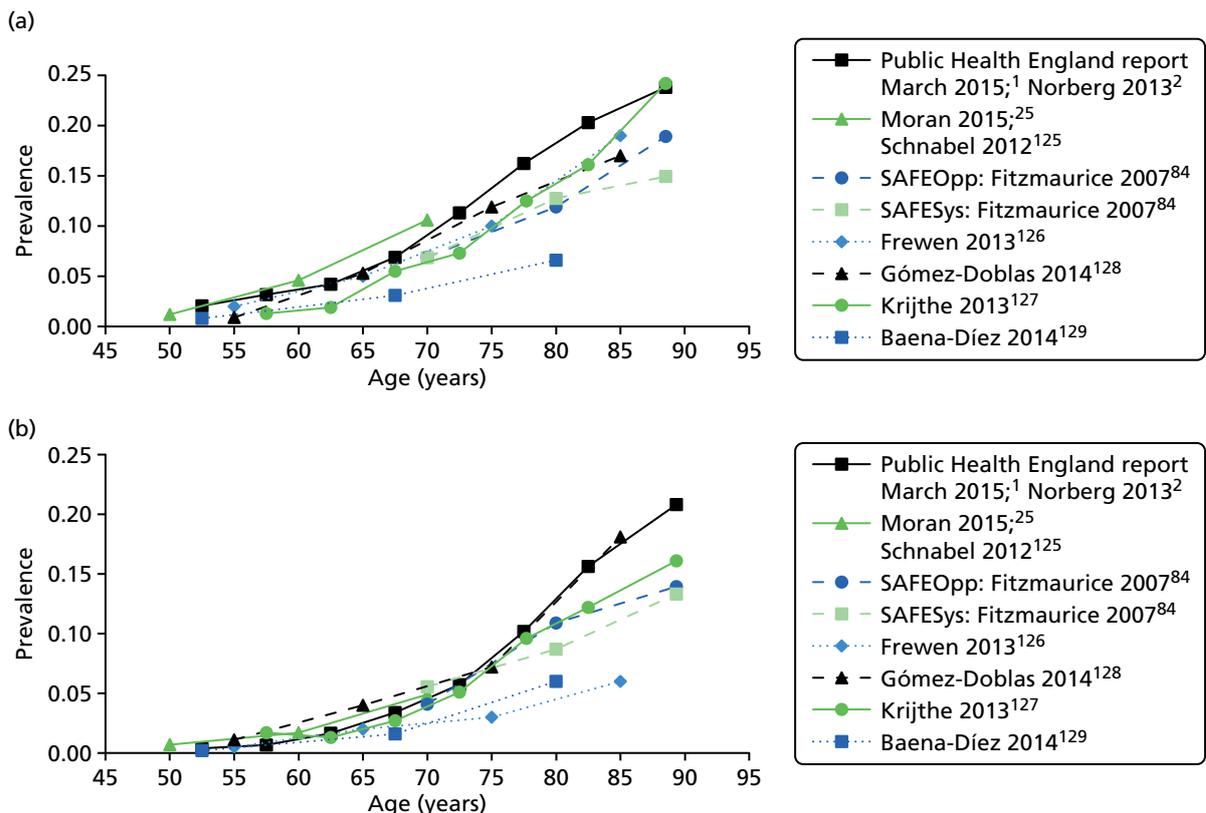
FIGURE 19 Estimated prevalence of AF by age and sex plotted for the studies identified in our literature review relevant to a contemporary Western European population. (a) Men; and (b) women.

TABLE 20 Prevalence of AF (diagnosed and undiagnosed combined) by age and sex

Age (years)	Men			Women		
	Population, *n*	Cases of AF, *n*	Prevalence (95% CI)	Population, *n*	Cases of AF, *n*	Prevalence (95% CI)
50–54	2575	53	0.021 (0.015 to 0.026)	2575	10	0.004 (0.001 to 0.006)
55–59	2710	86	0.032 (0.025 to 0.038)	2549	17	0.007 (0.004 to 0.010)
60–64	2736	115	0.042 (0.035 to 0.050)	2596	43	0.017 (0.012 to 0.021)
65–69	2383	164	0.069 (0.059 to 0.079)	2450	83	0.034 (0.027 to 0.041)
70–74	1874	212	0.113 (0.099 to 0.127)	1957	112	0.057 (0.047 to 0.068)
75–79	1405	228	0.162 (0.143 to 0.182)	1797	183	0.102 (0.088 to 0.116)
80–84	1015	206	0.203 (0.178 to 0.228)	1478	231	0.156 (0.138 to 0.175)
≥ 85	731	174	0.238 (0.207 to 0.269)	1350	281	0.208 (0.186 to 0.230)

Source: adapted from data published in Norberg *et al.*[2] © 2013 Norberg *et al.* published by Dove Medical Press Limited, and licensed under Creative Commons Attribution – Non Commercial (unreported, v3.0) License.

TABLE 21 Population size eligible for screening derived from ONS 2014 estimates for England and Wales[130]

Screening age (years)	ONS range (years)	Male population in range, *n*	Female population in range, *n*	Screening age population, *n*
55	55–59	1,670,449	1,708,826	675,855.0
60	60–64	1,517,771	1,583,693	620,292.8
65	65–69	1,540,906	1,627,388	633,658.8
70	70–74	1,110,603	1,222,424	466,605.4
75	75–79	873,540	1,026,027	379,913.4
80	80–84	595,901	801,284	279,437.0

TABLE 22 Summary of studies reporting information on the uptake of systematic population screening

Study	Population	Method	Uptake, *n*	Invited, *n*	Proportion uptake
Morgan 2002[82]	Four UK general practices, aged ≥ 65 years, either systematic population or systematic opportunistic screening	Pulse palpation + optional ECG	1099	1499	0.733
Hobbs 2005[19]	50 UK general practice centres, aged ≥ 65 years, systematic population, systematic opportunistic or no screening	ECG	2357	4433	0.532
Davis 2012[131]	English general practices 1995–9, aged ≥ 45 years invited for screening	12-lead ECG	3960	6286	0.630
Engdahl 2013[22]	75- and 76-year-olds in a region of Sweden invited to screening	12-lead ECG + 2 weeks hand-held ECG, systematic	848	1330	0.638
Gómez-Doblas 2014[128]	Spain, general population aged ≥ 40 years, 47 hospitals across 46 provinces	ECG	8400	11,055	0.760

TABLE 22 Summary of studies reporting information on the uptake of systematic population screening (*continued*)

Study	Population	Method	Uptake, *n*	Invited, *n*	Proportion uptake
Bury 2015[132]	25 Irish general practice centres, aged ≥ 70 years invited for screening	3-lead ECG, systematic	639	1003	0.637
Svennberg 2015[40]	Individuals aged ≥ 65 years invited to screening in two locations in Sweden	Hand-held ECG over 2 weeks (24 hours of continuous monitoring if unclear result), systematic	7173	13,331	0.538
Pooled random-effects posterior mean (95% CrI) 0.641 (0.543 to 0.730)[a]					

a Pooled using random-effects analysis on a log-odds scale in WinBUGS version 1.4.3 (convergence checked using Brooks–Rubin–Gelman diagnostic, *n* = 10,000 burn-in, *n* = 20,000 sample) (MRC Biostatistics Unit, Cambridge, UK). Results converted to probabilities for presentation.

TABLE 23 Summary of studies reporting information on the uptake of systematic population screening by age

Study	Population	Age group (years)	Uptake, *n*	Invited, *n*	Proportion uptake (95% CI)
Hobbs 2005[19]	50 UK general practice centres, aged ≥ 65 years, systematic population, systematic opportunistic or no screening	< 75	1471	2418	0.608 (0.589 to 0.628)
		≥ 75	741	1726	0.429 (0.406 to 0.453)

Proportion of initial invitees requiring a reminder letter

Two studies reported the proportion of reminder invitations sent during systematic screening (*Table 24*). This proportion varied from 0.4 to 0.6. We used the results from the SAFE study as being the most representative of a UK population and a normal distribution was used to reflect uncertainty in this estimate.

Pulse taking in flagged individuals in opportunistic screening

Two studies reported the number of flagged individuals who had pulse palpation under a systematic opportunistic screening strategy (*Table 25*). There is a large discrepancy between these results. We used the results from Hobbs *et al.*[19] to inform beta distributions in our base-case model as this was the more recent, larger study and the results were broken down by age group (< 75 and ≥ 75 years). Both of these studies used paper flags, whereas in current practice computerised flags are likely to be more feasible and sustainable for recurrent opportunistic screening; however, GPs may be less likely to respond to computerised flags because of 'alert fatigue'.[133] We used the results from Morgan and Mant[82] in a sensitivity analysis (assuming the same rate in both age groups as the rates were not reported separately), to represent a lower uptake of systematic opportunistic screening.

The proportion of 12-lead electrocardiograms interpreted by a general practitioner that require further interpretation by a cardiologist

We did not identify any information on the proportion of 12-lead ECGs read by a GP that would require further interpretation by a cardiologist and so took advice from our clinical experts. On the basis of this, we assumed that 10% of ECGs read by a GP would require a cardiologist interpretation to confirm the diagnosis. We assumed a 95% CI from 1% to 20%.

TABLE 24 Summary of studies reporting the proportion of patients invited to screening who required a reminder letter

Study	Population	Arm/group	Invited, *n*	Proportion reminder (95% CI)
Hobbs 2005[19]	50 UK general practice centres, aged ≥ 65 years, systematic population, systematic opportunistic or no screening	Systematic	4433	0.45 (0.435 to 0.465)
Svennberg 2015[40]	Individuals aged ≥ 65 years invited to screening in two locations in Sweden	Stockholm	10,908	0.6
		Halland	2345	0.4

TABLE 25 Summary of studies reporting the proportion of patients with flagged notes who have their pulses checked

Study	Population	Age group (years)	Pulses taken on consultation, *n*	Notes flagged, *n*	Proportion pulses taken (95% CI)
Hobbs 2005[19]	50 UK general practice centres, aged ≥ 65 years, systematic population, systematic opportunistic or no screening	< 75	1806	2585	0.699 (0.681 to 0.716)
		≥ 75	1224	1828	0.670 (0.648 to 0.691)
Morgan 2002[82]	Four UK general practices, aged ≥ 65 years, either systematic population or systematic opportunistic screening		439	1437	0.305 (0.282 to 0.329)

Rate of diagnosis of atrial fibrillation in the absence of screening

The SAFE study compared systematic and opportunistic screening against a no screening control.[19] *Table 26* shows the diagnosis rate in the SAFE study no screening arm. The rate of detection in the absence of screening is low and it is likely that in many cases diagnosis of AF occurs only after a stroke or another acute event. Our DOAC model[37] already tracks the incidence of adverse events in individuals with AF not on anticoagulation therapy and assumes that anticoagulation therapy is initiated following a stroke in those who are not contraindicated. Therefore, to avoid double counting a diagnosis of AF following acute events, we assumed that there would be no 'serendipitous' diagnosis of AF in the no screening arm.

The proportion of atrial fibrillation that is undiagnosed

It is possible to estimate the proportion of AF that is undiagnosed from screening studies (*Table 27*). However, these estimates are not directly comparable with each other because different screening tests are used. Furthermore, we might expect these to be underestimates because a one-off ECG is unlikely to detect paroxysmal AF. In the two studies with prolonged monitoring,[22,40] higher undiagnosed proportions were seen than in many studies with one-off monitoring. A recent report by Public Health England[1] used prevalence estimates from a Swedish study[2] (that comprehensively searched ECG records in a variety of sources together with patient records) and applied these estimates to the demographics of GPs in England to obtain predicted AF cases in England. Comparing this prediction with the figures reported in the QOF[3] for England GPs gives an estimate of the proportion of AF that is undiagnosed of 0.348. Given that this estimate is in line with the most intensive screening study,[22] it has been adjusted to match demographics in England and prevalence estimates are available broken down by age and sex, and we used this estimate of the proportion of undiagnosed AF in our model.

TABLE 26 Summary of the study reporting the diagnosis rate of AF in the absence of screening

Study	Population	Incident AF, n	Population, n	1-year incident AF diagnosis rate (95% CI)
Hobbs 2005[19]	25 UK general practice centres, aged ≥ 65 years, control (no screening)	47	4936	0.0095 (0.0068 to 0.0122)

TABLE 27 Summary of screening studies for AF

Study	Population	Method	Number screened with known AF	Number with screen-detected AF	Number screened with AF	Proportion AF is undiagnosed
Public Health England[1]	England – diagnosed AF from QOF. Total AF estimated using prevalence by age and sex from Swedish database data[2] applied to demographics from the QOF	Estimated using patient demographics in England and Swedish database	889,383	473,938	1,363,321	0.348
Engdahl 2013[22]	75- and 76-year-olds in a region of Sweden invited to screening	12-lead ECG + 2-week hand-held ECG	81	40	121	0.331
Svennberg 2015[40]	Individuals aged ≥ 65 years invited to screening in two locations in Sweden	Hand-held ECG over 2 weeks (24 hours of continuous monitoring if unclear result)	666	218	884	0.247
Hobbs 2005[19]	50 UK general practice centres, aged ≥ 65 years, systematic population, systematic opportunistic or no screening	Systematic: ECG	339	52	391	0.133
Hobbs 2005[19]	50 UK general practice centres, aged ≥ 65 years, systematic population, systematic opportunistic or no screening	Opportunistic: pulse palpation + ECG	340	31	371	0.084
Bury 2015[132]	25 Irish general practice centres, individuals aged ≥ 70 years invited for screening	3-lead ECG	23	12	35	0.343
Claes 2012[134]	Aged ≥ 40 years, Belgium, nationwide volunteers	Single-lead ECG	771	228	999	0.228
Kaasenbrood 2016[135]	10 general practices in the Netherlands, individuals aged ≥ 60 years invited to screening during flu vaccination	Single-lead ECG	84	37	121	0.306
Lowres 2014[21]	Australia, individuals aged ≥ 65 years attending 10 community pharmacies invited to screening	Pulse palpation + iECG (checked by cardiologist)	104	15	119	0.126

TABLE 27 Summary of screening studies for AF (*continued*)

Study	Population	Method	Number screened with known AF	Number with screen-detected AF	Number screened with AF	Proportion AF is undiagnosed
Deif 2013[136]	Australian hospital, individuals aged ≥ 65 years attending for minor surgery invited for screening	ECG in hospital	98	10	108	0.093
Rhys 2013[117]	Individuals aged ≥ 65 years attending for flu vaccination in a single centre in the UK	Pulse + ECG	21	2	23	0.087
Gómez-Doblas 2014[128]	Spain, general population aged ≥ 40 years, 47 hospitals across 46 provinces	ECG	369	41	410	0.1
Panisello-Tafalla 2015[137]	Spanish AF registry of those aged > 60 years	Unclear				0.309

The proportion of atrial fibrillation that is paroxysmal, persistent and permanent

Although several studies provided information on the proportion of patients with paroxysmal, persistent and permanent AF, these were primarily studies of diagnosed patients, which will underestimate the proportion of paroxysmal patients quite dramatically. Instead, we were interested in the proportion of paroxysmal AF in a systematic screening population (*Table 28*). The proportion of AF of each type varied widely across screening studies. Engdahl *et al.*[22] and Arronnson *et al.*[100]/Svenberg *et al.*[40] used a continuous monitoring screening test in population-based screening that was able to detect paroxysmal AF. Both of these studies detected a higher proportion of paroxysmal AF than persistent or permanent AF, suggesting that paroxysmal AF is more prevalent in people with undiagnosed AF, but that most screening studies are not able to detect it. For example, among the new AF cases detected by Sanmartín *et al.*,[138] only 6% were paroxysmal AF. In our model we assumed that 6% of screen-detected AF is paroxysmal AF and the remaining 94% is permanent or persistent AF. In sensitivity analysis, we varied the proportion of paroxysmal AF that can be detected.

For the proportion of all undiagnosed AF that is paroxysmal we pooled the data from Engdahl *et al.*[22] and Arronsson *et al.*[100]/Svenberg *et al.*[40] using a fixed-effects meta-analysis (see *Table 28*).

The proportion of atrial fibrillation that is asymptomatic (by atrial fibrillation type)

We were interested in the proportion of patients, by AF type, who are asymptomatic in a population-based screening population (*Table 29*). Two of the studies[136,139] gave results for screen-detected AF based on electrocardiography (and hence mostly persistent and permanent AF). These studies found a high proportion of AF that is asymptomatic (0.8 and 0.49 respectively), as would be expected in a screening context. Another study[14] reporting 1 year of follow-up of newly diagnosed AF patients found a similarly high proportion of asymptomatic patients (0.87). In patients with diagnosed chronic AF (persistent or permanent), there was a consistent finding of approximately 30% of patients being asymptomatic.[14,140,141] The proportion of paroxysmal AF that was asymptomatic was higher (estimates of 0.4[142] and 0.7[14]).

We pooled the data from Deif *et al.*,[136] Smyth *et al.*[139] and Nieuwlaat *et al.*[14] in our model using a fixed-effects meta-analysis (see *Table 29*).

TABLE 28 Summary of screening studies providing information on the proportion of patients with paroxysmal, persistent or permanent AF

Study	Population	Proportion paroxysmal AF	Proportion persistent or permanent AF
Engdahl 2013[22]	75- and 76-year-olds in a region of Sweden invited for screening, 12-lead ECG+ 2-week hand-held ECG	30/40 = 0.75	10/40 = 0.25
Arronsson 2015[100]/ Svennberg 2015[40]	Individuals aged ≥ 65 years invited for screening in two locations in Sweden; 2-week hand-held ECG	182/218 = 0.835	36/218 = 0.165
Sanmartín 2013[138]	Individuals aged ≥ 65 years invited for screening, three Spanish health centres, pulse palpation + 12-lead ECG	1/17 = 0.059	16/17 = 0.941

Pooled fixed-effects estimate posterior mean (95% CrI) 0.822 (0.773 to 0.865)[a]

a Data from Engdahl et al.[22] and Arronnson et al.[100]/Svenberg et al.[40] were pooled using a fixed-effects meta-analysis (insufficient number of studies to estimate a random-effects model) on a log odds scale in WinBUGS version 1.4.3 (convergence checked using the Brooks–Rubin–Gelman diagnostic, $n = 10,000$ burn-in, $n = 20,000$ sample). Results converted to probabilities for presentation.

TABLE 29 Summary of studies providing information on the proportion of screen-detected AF that is asymptomatic (by AF type)

Study	Population	Proportion asymptomatic in paroxysmal AF	Proportion asymptomatic in permanent AF	Proportion asymptomatic in persistent and permanent AF	Proportion asymptomatic in newly diagnosed AF
Deif 2013[136]	Australian hospital, individuals aged ≥ 65 years attending for minor surgery invited for screening				8/10 = 0.8
Smyth 2016[139]	37 GPs in rural Ireland, individuals aged ≥ 65 years flagged for opportunistic screening				27/55 = 0.491
Nieuwlaat 2008[14]	Euro Heart Survey, 1-year follow-up of AF patients	821/1170 = 0.702		686/2012 = 0.341	615/708 = 0.869
Senoo 2014[142]	Shinken database, paroxysmal AF from a single hospital in Japan	468/1176 = 0.398			
Rienstra 2014[140]	Permanent AF patients in the RACE trial, 33 centres, the Netherlands		157/522 = 0.301		
Frykman 2001[141]	Swedish hospital outpatients with permanent or persistent AF			91/282 = 0.323	

Pooled fixed-effects estimate posterior mean (95% CrI) 0.841 (0.814 to 0.866)[a]

RACE, RAte Control versus Electrical cardioversion for persistent atrial fibrillation.
a Data from Deif et al., Smyth et al. and Nieuwlaat et al. were pooled using a fixed-effects meta-analysis (insufficient number of studies to estimate a random-effects model) on a log odds scale in WinBUGS version 1.4.3 (convergence checked using the Brooks–Rubin–Gelman diagnostic, $n = 10,000$ burn-in, $n = 20,000$ sample). Results converted to probabilities for presentation.

Disease progression from paroxysmal atrial fibrillation to persistent/permanent atrial fibrillation

Table 30 summarises studies reporting information on the rate at which paroxysmal AF patients progress to persistent and permanent AF. The majority of the studies report the rate at which paroxysmal AF progresses to permanent AF. The results are fairly consistent across studies, with the only UK study[143] giving a very similar estimate to the largest study conducted in the US (Holmqvist[144]). Holmqvist[144] also provides the figures to estimate the rate that individuals progress from paroxysmal to chronic (persistent or permanent) AF, which is what is required in our model. Given the agreement between Holmqvist[144] and the UK study, we chose to use Holmqvist[144] to estimate the rate of progression from paroxysmal to chronic AF in the economic model.

Hazard ratios for stroke and mortality risk by whether asymptomatic or not

Xiong *et al.*[124] presented a systematic review of studies comparing mortality and stroke outcomes in symptomatic and asymptomatic patients. Because symptomatic patients may receive different anticoagulation therapy and other treatments from asymptomatic patients, we present here the results from the RCTs included in the study by Xiong *et al.*,[124] in which allocation of symptomatic and asymptomatic patients across trial arms should have been balanced. The results from these two RCTs[35,140] included in the study by Xiong *et al.*[124] are summarised in *Table 31*.

TABLE 30 Summary of studies providing information on disease progression from paroxysmal (Px) to persistent(Ps)/permanent (Pm) AF

Study	Population	Mean follow-up (years)	Px, *n*	Px → Ps (but not Pm), *n*	Px → Pm, *n*	Px → Ps or Pm, *n*
Holmqvist 2015[144]	ORBIT-AF registry of outpatients aged ≥ 18 years, centres across the USA	1.5	4697	556 (rate = 0.079)	476 (rate = 0.068)	1032 (rate = 0.146)
Ruigomez 2005[143]	UK General Practice Research Database (GPRD)	2.7	418		70 (rate = 0.062)	
Nieuwlaat 2008[14]	Euro Heart Survey	1	1170			234 (rate = 0.200)
Kato 2004[145]	Patient records at a single hospital, Japan	14	171		132 (rate = 0.055)	
Al-Khatib 2000[146]	Patients of a single cardiologist, NC, USA	4	231		42 (rate = 0.045)	

TABLE 31 Summary of studies providing information on stroke and mortality risk in symptomatic compared with asymptomatic AF

Study	Population	All-cause mortality, HR symptomatic vs. asymptomatic (95% CI)	Composite outcome (mortality + events), HR symptomatic vs. asymptomatic (95% CI)
Rienstra 2014[140]	Permanent AF patients in the RACE trial, 33 centres, the Netherlands		1.96 (1.09 to 3.45)
Flaker 2005[35]	AFFIRM RCT AF patients, USA	Unadjusted: 1.29 (0.99 to 1.68); adjusted: 1.07 (0.79 to 1.46)	Unadjusted: 1.31 (1.04 to 1.65); adjusted: 1.14 (0.87 to 1.50)

AFFIRM, Atrial Fibrillation Follow-up Investigation of Rhythm Management; RACE, RAte Control versus Electrical cardioversion for persistent atrial fibrillation.

Both studies show that symptomatic patients are more likely to have a composite outcome of mortality and other events than asymptomatic patients; however, Flaker et al.[35] show that, when adjusted for other risk factors, this effect diminishes (HR 1.14, 95% CI 0.87 to 1.50). Flaker et al.[35] found a similar result for all-cause mortality, with an adjusted HR of 1.07 (95% CI 0.79 to 1.46). Because the risk of stroke and mortality in our model already accounts for risk factors (CHA$_2$DS$_2$-VASc score, previous MI, previous stroke), the adjusted estimates are of relevance to our model and, as these suggest that there is no evidence that risk is any different in asymptomatic patients from symptomatic patients, we do not allow for differential risk according to symptomatic status in our base case. In sensitivity analysis, we assumed that the event rate in screen-detected patients is lower than that for routinely detected AF using the unadjusted HR for symptomatic AF (representing non-screen-detected AF) compared with asymptomatic AF (representing screen-detected AF) found by Flaker et al.[35] of 1.31 (95% CI 1.04 to 1.65).

Hazard ratios for stroke and mortality risk by atrial fibrillation type

To obtain estimates of the HRs for events for different AF types, we need to avoid confounding by treatment, as permanent AF is more likely to be treated than paroxysmal AF. We selected studies in which treatment allocation was the same across AF type, mostly post hoc analyses of pooled RCT arms, but also one study in which the results were broken down by treatment (Table 32). Only one study[147] gave the HR for mortality for persistent relative to paroxysmal AF, but this was in close agreement with the HR for stroke in that same study, which supports an assumption that the HRs are the same for both stroke and mortality outcomes. The majority of the studies gave results for stroke comparing persistent with paroxysmal AF, showing in general an increased risk of stroke in persistent compared with paroxysmal AF. Vanassche et al.[9] also presented the HR for permanent compared with paroxysmal AF, which suggested that there is a 'dose–response relationship', with risk increasing from paroxysmal to persistent to permanent AF. We pooled the HRs from the studies reporting HRs (and CIs) for stroke in persistent compared with paroxysmal AF. Because there were only three such studies, we pooled them using a fixed-effects model. In our model we do not distinguish between persistent and permanent AF; for at least one of the studies, the result for persistent AF also included permanent AF. We therefore used the HR for persistent compared with paroxysmal AF in our model. We assumed that the HR is the same for mortality as it is for stroke.

TABLE 32 Summary of studies providing information on stroke and mortality risk for persistent (Ps) or permanent (Pm) AF compared with paroxysmal (Px) AF

Study	Population	Stroke HR, Ps vs. Px (95% CI)	Stroke HR, Pm vs. Px (95% CI)	Mortality HR, Ps vs. Px (95% CI)
Steinberg 2015[147]	AF patients in the ROCKET-AF RCT of rivaroxaban vs. warfarin; arms pooled	1.28 (1.01 to 1.64)		1.27 (1.06 to 1.49)
Vanassche 2015[9]	Two RCT AF populations; one or more risk factors for stroke, contraindicated for OACs, on aspirin	1.43 (1.04 to 1.96)	2.04 (1.6 to 2.61)	
Friberg 2010[148]	All patients treated for AF in a Swedish hospital in 2002	Not on warfarin 1.53, on warfarin 1.11		
Hohnloser 2007[149]	RCT AF population, both arms pooled; patients with one or more risk factors for stroke	1.15 (0.77 to 1.69)		

Pooled fixed-effects estimate posterior mean (95% CrI) 1.3 (1.09 to 1.54)[a]

ROCKET-AF, Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation.
a Data from Steinberg et al.[147], Vanassche et al.[9] and Hohnloser et al.[149] were pooled for the outcome of stroke in persistent relative to paroxysmal AF using a fixed-effects meta-analysis on a log HR scale in WinBUGS version 1.4.3 (convergence checked using the Brooks–Rubin–Gelman diagnostic, n = 10,000 burn-in, n = 20,000 sample). Results converted to HRs for presentation.

The proportion of screen-detected atrial fibrillation with a CHA$_2$DS$_2$-VASc score of ≥ 2

Table 33 details the results from six studies providing information on the proportion of AF patients with a CHA$_2$DS$_2$-VASc score of ≥ 2 in a screening or community population. The results are given separately for known (diagnosed) AF patients and screen-detected AF patients. We pooled the data from all four studies reporting results for screen-detected AF using a fixed-effects meta-analysis for use in our model.

The proportion of screen-detected atrial fibrillation cases with a previous history of myocardial infarction

Table 34 shows the results from studies reporting the proportion of screen-detected AF cases with a previous history of MI. We also show the results for studies of newly diagnosed AF; however, it should be noted that these patients are more likely to be symptomatic at diagnosis than screen-detected AF patients. This is reflected in the results, with a much higher proportion of newly diagnosed AF cases having a previous history of MI than screen-detected or asymptomatic AF cases. The proportion of patients with a previous MI was 5.4%[135] and 0%[138] in the two studies of screen-detected AF. We pooled the data from both studies reporting results for screen-detected AF using a fixed-effects meta-analysis for use in our model.

TABLE 33 Summary of studies reporting the proportion of screen-detected AF with a CHA$_2$DS$_2$-VASc score of ≥ 2

Study	Population	P(CHA$_2$DS$_2$-VASc score of ≥ 2), known and screen-detected AF	P(CHA$_2$DS$_2$-VASc score of ≥ 2), known AF	P(CHA$_2$DS$_2$-VASc score of ≥ 2), screen-detected AF
Claes 2012[134]	Aged ≥ 40 years, Belgium, nationwide volunteers			0.72
Kaasenbrood 2016[135]	10 GPs in the Netherlands, individuals aged ≥ 60 years invited to screening during flu vaccination		73/84 = 0.869	29/37 = 0.784
Smyth 2016[139]	37 GPs in rural Ireland, individuals aged ≥ 65 years flagged for opportunistic screening	37/39 = 0.949		
Lowres 2014[21]	Australia, individuals aged ≥ 65 years attending 10 community pharmacies invited to screening			15/15 = 1
Deif 2013[136]	Australian hospital, individuals aged ≥ 65 years attending for minor surgery invited for screening			11/12 = 0.91
Frewen 2013[126]	Irish community sample	81/119 = 0.681		

Pooled fixed-effects estimate posterior mean (95% CrI) 0.750 (0.699 to 0.797)[a]

a　Data from Kasenbrood *et al.*,[135] Smyth *et al.*,[139] Lowres *et al.*[21] and Deif *et al.*[136] were pooled using a fixed-effects meta-analysis (insufficient number of studies to estimate a random-effects model) on a log odds scale in WinBUGS version 1.4.3 (convergence checked using the Brooks–Rubin–Gelman diagnostic, *n* = 10,000 burn-in, *n* = 20,000 sample). Results converted to probabilities for presentation.

TABLE 34 Summary of studies reporting the proportion of screen-detected AF cases with a previous history of MI

Study	Population	Proportion with previous MI, diagnosed paroxysmal AF	Proportion with previous MI, diagnosed persistent/permanent AF	Proportion with previous MI, all diagnosed	Proportion with previous MI, asymptomatic diagnosed AF	Proportion with prior MI, screen-detected AF
Kaasenbrood 2016[135]	Screen-detected AF, 10 GPs in the Netherlands, invited a subset of those attending for flu vaccination					2/37 = 0.054
Sanmartín 2013[138]	Screen-detected AF, aged ≥ 65 years, three Spanish health centres					0/17 = 0
Martinez 2014[150]	UK Clinical Practice Research Datalink incidentally detected AF patients				232/5555 = 0.042	
Avgil Tsadok 2012[151]	Patients admitted to hospitals in Quebec province, Canada, with newly diagnosed AF, identified from hospital records			Women: 5826/44,115 = 0.132; men: 8192/39,398 = 0.208		
Kerr 2005[152]	Canadian Registry of Atrial Fibrillation (CARAF) – newly diagnosed AF by type	99/588 = 0.168	30/169 = 0.178			

Pooled fixed-effects estimate posterior mean (95% CrI) 0.037 (0.005 to 0.099)[a]

a Data from Kasenbrood et al.[135] and Sanmartin[138] were pooled using a fixed-effects meta-analysis (insufficient number of studies to estimate a random-effects model) on a log odds scale in WinBUGS version 1.4.3 (convergence checked using the Brooks–Rubin–Gelman diagnostic, $n = 10,000$ burn-in, $n = 20,000$ sample). Results converted to probabilities for presentation.

The proportion of screen-detected atrial fibrillation cases with a previous history of ischaemic stroke (or transient ischaemic attack)

Table 35 shows the results from studies reporting the proportion of screen-detected AF cases with a previous history of ischaemic stroke or TIA. We also show the results for studies of newly diagnosed AF. The results do not show a consistent pattern of previous stroke risk in newly diagnosed AF patients compared with screen-detected AF. The results are variable even when restricting to screen-detected AF. We pooled the data from all seven studies (eight arms) reporting results for screen-detected AF using a random-effects meta-analysis to account for heterogeneity, reported in *Table 35*. However, the most relevant study to a UK screening population is the SAFE study,[153] the estimate for which is only just

TABLE 35 Summary of studies reporting the proportion of screen-detected AF cases with a previous history of ischaemic stroke (or TIA)

Study	Population	Proportion with previous stroke, diagnosed AF	Proportion with previous stroke, asymptomatic diagnosed AF	Proportion with previous stroke, screen-detected AF
Fitzmaurice 2014[153]	50 UK GP centres, aged ≥ 65 years, systematic screen-detected AF			7/74 = 0.095
Fitzmaurice 2014[153]	50 UK GP centres, aged ≥ 65 years, opportunistic screen-detected AF			4/75 = 0.053
Svennberg 2015[40]	Two locations in Sweden, aged ≥ 65 years, systematic screen-detected AF, hand-held ECG over 2 weeks (24 hours of continuous monitoring if unclear result)	137/666 = 0.206		21/218 = 0.096
Claes 2012[134]	Aged ≥ 40 years, Belgium, nationwide volunteers, single-lead ECG screen-detected AF			27/228 = 0.118
Kaasenbrood 2016[135]	10 GPs in the Netherlands, invited a subset of those attending for flu vaccination. Results for diagnosed and screen-detected AF	9/84 = 0.107		7/37 = 0.189
Sanmartín 2013[138]	Aged ≥ 65 years, three Spanish health centres, screen-detected AF			4/17 = 0.235
Gómez-Doblas 2014[128]	Spain, general population aged ≥ 40 years, 47 hospitals across 46 provinces			7/41 = 0.171
Lowres 2014[21]	Australia, individuals aged ≥ 65 years attending 10 community pharmacies. Results for diagnosed and screen-detected AF			2/15 = 0.133
Martinez 2014[150]	UK Clinical Practice Research Datalink incidentally detected AF patients		509/5555 = 0.092	
Avgil Tsadok 2012[151]	Patients admitted to hospitals in Quebec province, Canada, with newly diagnosed AF, identified from hospital records	Women: 3511/44,115 = 0.08; men: 2733/39,398 = 0.069		

Pooled random-effects mean posterior mean (95% CrI) 0.116 (0.080 to 0.168)[a]

SAFE study only 0.074 (0.032 to 0.116)

a Data from the seven screening studies (eight arms) were pooled in a random-effects meta-analysis on a log odds scale in WinBUGS version 1.4.3 (convergence checked using the Brooks–Rubin–Gelman diagnostic, $n = 20,000$ burn-in, $n = 40,000$ sample). Results converted to probabilities for presentation.

contained in the CrI around the random-effects mean. We elected to estimate the history of previous stroke in a UK screen-detected AF population by pooling the systematic population and systematic opportunistic arms of the SAFE study for use in our economic model (reported in *Table 35*). We assumed a normal distribution for this history of previous stroke. In sensitivity analysis, we used the random-effects mean across all studies reporting screen-detected AF.

Anticoagulation therapy for screen-detected atrial fibrillation with a CHA$_2$DS$_2$-VASc score of ≥ 2

Our model required information on the proportions of screen-detected AF patients with a CHA$_2$DS$_2$-VASc score of ≥ 2 who are prescribed OACs, who are contraindicated to or who decline OACs and who simply do not have OACs prescribed. The studies providing information relevant to this are summarised in *Table 36*. Only one study[126] provided information on the HAS-BLED bleeding risk score in AF patients with a CHA$_2$DS$_2$-VASc score of ≥ 2, finding 2.49% with a HAS-BLED score of ≥ 3. It should be noted that, although the HAS-BLED score can inform clinical decision making on whether or not a patient is suitable for OACs, other factors will also be taken into account. A more practical measure of those who are not suitable for OACs is the reported proportion contraindicated to, or preferring not to take, OACs. Three studies[3,118,154] reported this information in patients with a high stroke risk (albeit using the CHADS$_2$ rather than the CHA$_2$DS$_2$-VASc risk score), with all three studies giving very similar estimates (see *Table 36*). The QOF report[3] represents the most recent evidence covering the majority of GPs in England, with an estimated 12.94% of patients contraindicated or declining OACs ('exceptions' rate from the QOF report[3]), which we used in our model.

There is a wide variation in the proportion of AF patients with a CHA$_2$DS$_2$-VASc score of ≥ 2 who are prescribed OACs (see *Table 36*). Again, we used the figures from the QOF report[3] as they provide the most contemporary relevant evidence in England, giving an estimated 74.30% of AF patients who are prescribed OACs. It should be noted that this leaves 12.76% of patients who are not contraindicated but who do not receive OACs. In our model base case, we assumed that all patients not prescribed OACs (12.94% contraindicated or declining OACs plus 12.76% not prescribed OACs for other reasons) receive aspirin.

It may be that patients with screen-detected AF are less likely to be prescribed OACs than patients with AF detected routinely, in which case the uptake of OACs may be lower than 74.30%. Alternatively, implementation activities may lead to an increase in uptake of OACs in the future. We explored the robustness of our results to this in sensitivity analyses, with the uptake of OACs reduced to 50% or increased to 87.0% (so that all who are not contraindicated or declining OACs receive OACs).

TABLE 36 Summary of studies reporting the proportion of patients contraindicated or declining OACs and the proportion prescribed OACs

Study	Population	HAS-BLED score of ≥ 3)	Contraindicated or declining OACs	Prescribed OACs
Health and Social Care Information Centre 2015[3]	England, GP reports, CHADS$_2$ score of ≥ 1		0.1294	0.743
Cowan 2013[154]	GRASP-AF registry, voluntary GPs in England, results for CHADS$_2$ score of ≥ 2		12,128/87,198 = 0.139	72,211/87,198 = 0.828
Gordon 2012[118]	Individuals aged ≥ 65 years attending for flu vaccination, 35 general practices in Sussex, screen-detected AF, CHADS$_2$ score of ≥ 2		10/84 = 0.119	41/84 = 0.488
Apenteng 2014[155]	UK subpopulation of GARFIELD-AF, an international registry of newly diagnosed AF, CHA$_2$DS$_2$-VASc score of ≥ 2			0.472
Deif 2013[136]	Patients aged ≥ 65 years attending for minor surgery, single hospital in Australia, known AF with CHADS$_2$ score of ≥ 2			52/82 = 0.634
Frewen 2013[126]	Irish community sample, known and screen-detected AF with a CHA$_2$DS$_2$-VASc score of ≥ 2	2/81 = 0.025		

GARFIELD, Global Anticoagulant Registry in the FIELD.

Uptake of DOACs has been slow, with some studies reporting that DOACs represent as little as 3.79%[156] of all OACs prescribed for AF. *Figure 20* shows the trends in prescribing of OACs in England over time. It can be seen that there has been a steady trend of increased prescribing of DOACs. *Figure 20* includes all patients prescribed OACs. We hypothesise that the trend towards increased DOAC use is driven by prescribing in newly diagnosed patients, rather than by switching treatment in patients who have historically been prescribed warfarin. This view is supported by the STROKESTOP study,[40] which found that, of those with newly diagnosed AF who were started on OACs, 73% were prescribed DOACs. For our model we therefore assumed in the base case that the proportion of OACs that are DOACs in newly diagnosed AF patients is 75%. We varied this to 50% and 100% in sensitivity analyses. We assumed that the OAC received is apixaban (most cost-effective in the DOAC model[37]) but note that the results would be similar if other DOACs were used instead.

Incremental quality-of-life benefits and costs from rate control compared with no rate control

Incremental lifetime costs and QALYs associated with rate control (compared with no rate control) are assumed to be fixed costs and benefits obtained by patients who are prescribed rate control as a result of screening. In practice, some of these patients would eventually have been detected with AF and given rate control, so our model will overestimate these benefits. We performed a sensitivity analysis assuming an annual rate of detection of AF in the absence of screening or an AF-related stroke.

The model developed by Lord *et al.*[116] provides information on the incremental costs and QALYs over a lifetime for rate control compared with no rate control for persistent AF. The authors found an incremental cost of –£260 (i.e. cost saving), which we inflated to –£274.55 at 2015 prices, and incremental QALYs of 0.0133 for rate control. We assumed these values in our model for symptomatic patients detected by screening.

Sensitivity and specificity of the screening tests

The assumed values for the sensitivity and specificity of the screening tests were the mean values from our HSROC model reported in the systematic review of DTA studies (see *Tables 4* and *5*).

Screening costs

The costs of the screening tests consist of the capital cost of any equipment required, the cost of consumables for each test performed and the cost of clinician time to administer and interpret the test results. In the case of systematic population screening there will also be costs associated with the screening invitation.
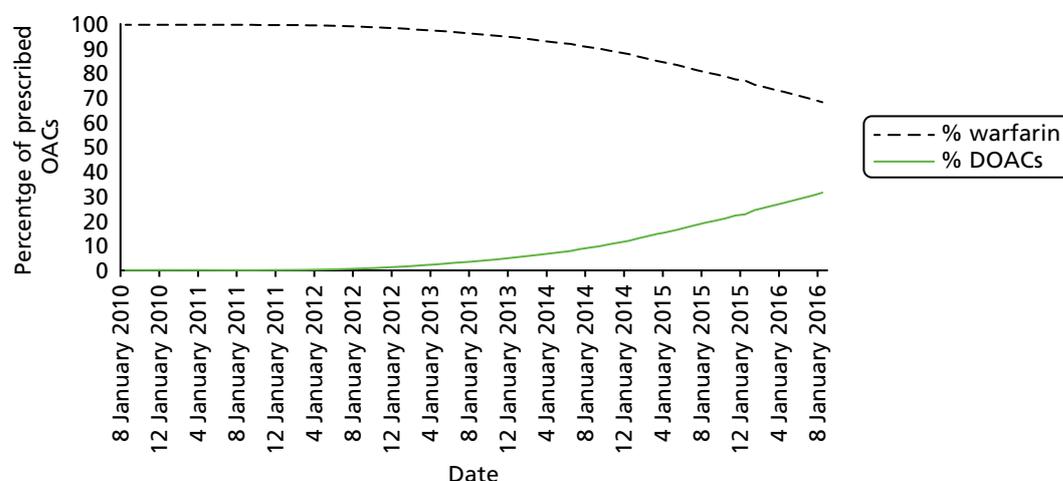
FIGURE 20 Trends over time in the proportion of OACs prescribed that are warfarin or DOACs in England, using data from Powell-Smith A, Goldacre B, OpenPrescribing.net, EBM DataLab, University of Oxford, 2017.

Equipment costs vary widely across screening tests, from approximately £1500 for a 12-lead ECG to £100 for a modified blood pressure monitor, £24 for photoplethysmography and no cost for pulse palpation. The capital cost per patient will depend on, among other factors, the lifespan of the device and the throughput of patients for AF screening. Patient throughput will, in turn, depend on practice size and demographics, screening uptake and the frequency of screening. The calculation is further complicated for multipurpose devices such as ECG machines, which will not be used solely for AF screening. In practice, even the relatively high purchase cost of an ECG machine will be spread widely over the large number of patients who use it. In our base-case analysis, we assumed that the capital cost of equipment per patient screened is zero and that the screening test costs consist of invitation costs, the costs of consumables and the costs of clinician and administrative time taken to administer and interpret the test results. We discuss the likely impact of this assumption and the issue of investment in equipment, particularly for screening strategies (e.g. photoplethysmography, modified blood pressure monitors) for which GPs do not currently have the equipment, in the discussion section of the next chapter.

We used the results from the SAFE study[19] to inform the costs of systematic population and systematic opportunistic screening. The cost of invitation to systematic population screening consists of the costs of the invitation letter, postage and 1 minute of administrator time. Estimates for these costs are based on the trial data from Hobbs et al.,[19] but we inflated from 2003 prices to quarter 1 2015 prices using the ONS Consumer Price Index for Medical Services (DKC3).[157] The quarter 1 2015 prices were chosen to maintain consistency, with costs used to generate the results of the DOAC model.[37] These costs are summarised in Table 37.

Test costs

We used PSSRU 2014/15 data[123] to estimate staff costs for administering and interpreting the screening and diagnostic tests. An hour of face-to-face contact with a nurse was assumed to cost £56 (including

TABLE 37 Summary of the costs of screening assumed in the economic model

Activity	Resource required	Source
Administrative time per individual per invitation letter	1 minute, administrator	Estimate[19]
Time to take pulse	1 minute, nurse or other health professional	Trial data[19]
Administrative time per individual attending for ECG	2 minutes, administrator	Trial data[19]
Nurse time per ECG (without interpretation)	12 lead: 7 minutes; limb lead: 3 minutes; single lead: 4 minutes	Trial data[19]
Interpretation of ECG	1 minute	Trial data[19]
Resource required	**Cost (£)**	**Source and comment**
Electrode pad	0.10	Hobbs et al.[19] (inflated from £0.3 in 2003 to 2015 prices)
Stationery and postage per ECG	0.50	Hobbs et al.[19] (inflated from £0.3 in 2003 to 2015 prices)
Invitation letter	0.50	Hobbs et al.[19] (inflated from £0.3 in 2003 to 2015 prices)
Administrator cost for invitation (1 minute)	0.30	Hobbs et al.[19] (cost per hour of £10.93 in 2003 prices inflated to £18.22 in 2015 prices giving a cost per minute of £0.30)
Total systematic invitation cost	1.31	
Rate control	−274.55	Lord et al.[116] (inflated from £260 in 2013 to 2015 prices)
Nurse (1 hour)	56	PSSRU[123]
GP (1 hour)	129	PSSRU[123]
Cardiologist (1 hour)	137	PSSRU[123]

qualification costs) (see *Table 37*). An hour of general medical services activity by a GP was assumed to cost £129 (including qualification costs but excluding direct care staff costs). An hour of cardiologist time was taken as the cost of 1 hour of 'consultant medical' time, which was £137 (including qualification costs).

We used the times for each test estimated by Hobbs *et al.*[19] for the times to complete systematic opportunistic screening. We also used their estimated interpretation time of 1 minute for all ECGs interpreted by a nurse, GP or cardiologist. Systematic population screening would require a dedicated appointment with a nurse. A typical practice nurse consultation is estimated to last approximately 15 minutes;[123] however, for routine screening the consultation time could be considerably less. Staff times assumed for each test are provided in *Table 38*.

TABLE 38 Summary of screening test costs assumed in the economic model

Test name	Opportunistic screening, staff required	Population screening, staff required	Consumables required	Total cost opportunistic screening (£)	Total cost population screening (£)
Modified blood pressure monitor	1 minute nurse	5 minutes nurse		1.54	5.27
Pulse palpation – nurse	1 minute nurse	5 minutes nurse		1.54	5.27
Photoplethysmography	1 minute nurse	5 minutes nurse		1.54	5.27
Single-lead ECG – automatic/algorithm	4 minutes nurse	8 minutes nurse	1 pad	4.44	8.17
Single-lead ECG – nurse	4 minutes nurse (administration) + 1 minute nurse (interpretation)	8 minutes nurse (administration) + 1 minute nurse (interpretation)	1 pad	5.37	9.11
Single-lead ECG – GP	4 minutes nurse (administration) + 1 minute GP (interpretation)	8 minutes nurse (administration) + 1 minute GP (interpretation)	1 pad	6.59	10.32
Single-lead ECG – cardiologist	4 minutes nurse (administration) + 1 minute cardiologist (interpretation)	8 minutes nurse (administration) + 1 minute cardiologist (interpretation)	1 pad	6.72	10.46
> 1- and < 12-lead ECG – automatic/algorithm	5.5 minutes nurse (administration)	9.5 minute nurse (administration)	2.5 pads	5.99	9.72
> 1- and < 12-lead ECG – cardiologist	5.5 minutes nurse (administration) + 1 minute cardiologist (interpretation)	9.5 minute nurse (administration) + 1 minute cardiologist (interpretation)	2.5 pads	8.27	12.01
12-lead ECG – automatic/algorithm	7 minutes nurse (administration)	11 minutes nurse (administration)	12 pads	8.34	12.07
12-lead ECG – nurse	7 minutes nurse (administration) + 1 minute nurse (interpretation)	11 minutes nurse (administration) + 1 minute nurse (interpretation)	12 pads	9.27	13.01
12-lead ECG – GP	7 minutes nurse (administration) + 1 minute GP (interpretation)	11 minutes nurse (administration) + 1 minute GP (interpretation)	12 pads	10.49	14.22
12-lead ECG – cardiologist	7 minutes nurse (administration) + 1 minute cardiologist (interpretation)	11 minutes nurse (administration) + 1 minute cardiologist (interpretation)	12 pads	10.62	14.36

If the initial screening test was a 12-lead ECG then the diagnostic test cost would be only the cost of the 1-minute interpretation by a GP or cardiologist, which is £2.15 and £2.28 respectively. If the initial screening test was something else (e.g. photoplethysmography) then the diagnostic test cost would be the full cost of a 12-lead ECG with administrator and nurse appointment time. We did not include the cost of a follow-up appointment with a GP or cardiologist for a true-positive 12-lead ECG diagnostic test as this was included in the management costs of the long-term model (see *Long-term costs and benefits*).

Following Hobbs *et al.*,[19] all test costs included 2 minutes of administrator time. Administrator time was priced at £10.93 per hour in Hobbs *et al.*[19] and was inflated to £18.22 (quarter 1 2015 prices) using the ONS Consumer Price Index for Medical Services (DKC3),[157] giving a cost of £0.61 for 2 minutes for each test.

We assumed that each lead of the ECG machine required one pad and that a > 1- and < 12-lead ECG, with an average of 2.5 leads, required 2.5 pads. Following clinical advice we assumed that no alcohol wipes were used for the ECGs.

The total estimated costs and their underlying assumptions for systematic opportunistic and systematic population screening are reported in *Table 38*.

Long-term costs and benefits

Anticoagulation therapy (directly acting oral anticoagulant model[37])
A previously developed Markov model[37] was used to estimate the long-term costs and outcomes for patients with screen-detected AF (true positives on initial screen, confirmed by 12-lead ECG). The model includes a HR for events (stroke, systemic embolism, TIA) affected by AF type (paroxysmal relative to persistent or permanent). We combined persistent and permanent AF in the model because the RCTs informing the efficacy of anticoagulation therapy predominantly combined persistent and permanent AF patients. The model depends on age, sex, previous history of ischaemic stroke or TIA and previous history of MI.

The discrete-time Markov multistate model used a cycle length of 3 months, as in other recent models.[158–160] We ran the model for a cohort for a given starting age and using a lifetime time horizon. The model structure is illustrated in *Figure 21*. Each treatment strategy has the same model structure but different costs, utilities and event probabilities. From any state, a patient can have a clinically relevant (extracranial) bleed, an intracranial haemorrhage, an ischaemic stroke or a MI, all of which have long-term consequences
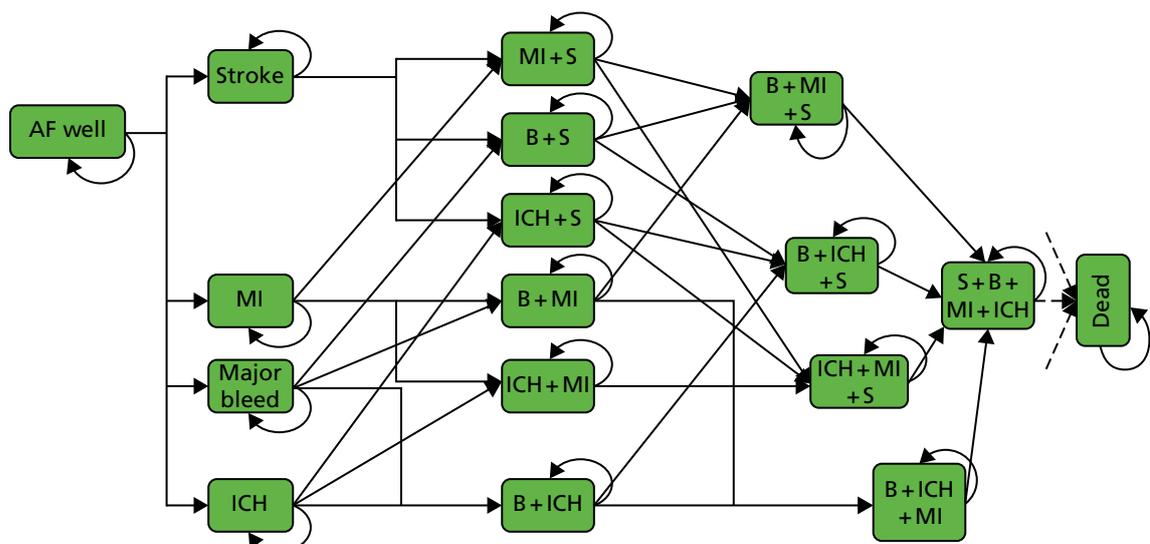
FIGURE 21 Discrete-time Markov model structure for screen-detected AF patients. B, major bleed; ICH, intracranial haemorrhage; S, stroke.

that are modelled. Patients can also experience a TIA or systemic embolism, which are transient states, or can discontinue or switch treatment because of any event, or die. Patients on apixaban who experience a bleed or intracranial haemorrhage will be switched to warfarin (INR range 2–3); patients experiencing a further bleed or intracranial haemorrhage are then switched to no treatment. Patients initially on no treatment or aspirin who experience a stroke and who are not contraindicated to DOACs and have no history of a bleed or intracranial haemorrhage are assumed to initiate apixaban treatment.

Transition probabilities were derived using HRs for apixaban and aspirin relative to warfarin estimated from a systematic literature review.[37] The warfarin arms were used to estimate baseline hazards. We relied on previous meta-analyses to estimate the relative effect of warfarin (INR range 2–3) compared with no treatment.[161] Evidence from the literature was used to estimate the effect of previous events on stroke, mortality, MI, systemic embolism, TIA and bleed risk.[162,163] Average drug costs were based on the *British National Formulary* March 2015 update,[164] and it was assumed that oral apixaban and aspirin incurred no monitoring or administration costs. Acute management costs for systemic embolism, MI, TIA and clinically relevant bleeding were obtained from 2013/14 NHS reference costs.[165] Acute and long-term management costs for ischaemic stroke and intracranial haemorrhage were taken from a study of AF patients on a UK stroke registry.[166] Utilities were identified from a previous NICE technology appraisal submission on rivaroxaban,[159] which included a systematic literature search for evidence on EQ-5D utility index scores in health states related to AF.

The model was modified to include management costs for AF reported published in the literature.[110] Costs were available for a 12-week initialisation period and annual 'maintenance' periods thereafter, summarised in *Table 39*. These costs included the costs of blood testing, medication, primary and secondary care visits, further diagnostic testing and inpatient visits. We adjusted the costs by removing the blood testing and medication costs, which were already accounted for in our model. We also inflated costs to quarter 1 2015 prices[157] to maintain consistency with the other model costs. We used a log-normal distribution with the mean and SD in *Table 39* ('adjusted total') for the initialisation and maintenance periods.

Summary of model inputs

Table 40 summarises the model parameters and assumed values and distributions and the evidence sources used for each.

Sensitivity analyses

In sensitivity analyses, we used the optimal screening test and screening method identified in the base case unless otherwise stated.

TABLE 39 Assumed non-anticoagulation therapy costs in the economic model[a]

Component	Initialisation period (£)		Maintenance period (£)	
	Mean	SD	Mean	SD
Blood testing	5.12	9.22	7.17	11.27
AF medications	9.22	10.25	50.20	50.20
Total	964.14	1120.90	436.48	611.68
Adjusted total	949.80	1120.99	379.10	613.84

a Costs are provided for a 12-week initialisation period and subsequent annual 'maintenance' time periods, based on Kassianos *et al.*[110]

TABLE 40 Summary of input parameter distributions and sources used in the cost-effectiveness model

Model Input	Value	Evidence source and distribution
Prevalence of AF (diagnosed and undiagnosed) by age and sex	See *Table 20*	Norberg *et al.*,[2] Public Health England[1]
Population size eligible for screening at each age	See *Table 21*	ONS[130]
Uptake of systematic population screening	< 75 years: 0.608 (0.589 to 0.628); ≥ 75 years: 0.429 (95% CI 0.406 to 0.453)	Hobbs *et al.*,[19] beta distribution
Proportion of initial invitees sent a reminder letter	0.45 (95% CI 0.435 to 0.465)	Hobbs *et al.*,[19] normal distribution
Proportion of flagged individuals who have their pulses checked	< 75 years: 0.699 (95% CI 0.681 to 0.716); ≥ 75 years: 0.670 (95% CI 0.648 to 0.691)	Hobbs *et al.*,[19] beta distribution
Proportion of AF undiagnosed	0.348	Public Health England[1]
Proportion of screen-detected AF that is asymptomatic	0.841 (95% CI 0.814 to 0.866)	Meta-analysis (see *Table 29*), Bayesian posterior
Proportion of undiagnosed AF that is paroxysmal	0.822 (95% CI 0.773 to 0.865)	Meta-analysis (see *Table 28*), Bayesian posterior
Proportion of screen-detected AF that is paroxysmal	0.06	Assumption based on Sanmartín *et al.*[138]
Rate of progression from paroxysmal to chronic AF	0.146 (95% CrI 0.138 to 0.156)	Holmqvist *et al.*,[144] Bayesian posterior
Proportion of 12-lead ECGs referred to a cardiologist	0.1 (95% CI 0.01 to 0.2)	Expert opinion, normal distribution
Proportion of screen-detected AF with a CHA_2DS_2-VASc score of ≥ 2	0.750 (95% CrI 0.699 to 0.797)	Meta-analysis (see *Table 33*), Bayesian posterior
Proportion of screen-detected AF cases with a previous history of MI	0.037 (95% CrI 0.005 to 0.099)	Meta-analysis (see *Table 34*), Bayesian posterior
Proportion of screen-detected AF cases with a previous history of stroke	0.074 (95% CI 0.032 to 0.116)	Fitzmaurice *et al.*,[153] normal distribution
Proportion of screen-detected AF with a CHA_2DS_2-VASc score of ≥ 2 contraindicated to OACs	0.129	QOF report[3]
Proportion of screen-detected AF with a CHA_2DS_2-VASc score of ≥ 2 prescribed OACs	0.743	QOF report[3]
Proportion of prescribed OACs that are DOACs	1	Assumption
HR for event rates in symptomatic vs. asymptomatic AF	Assumed 1 after adjusting for other risk factors	Flaker *et al.*[35]
HR for stroke/mortality in persistent vs. paroxysmal AF	1.3 (95% CrI 1.09 to 1.54)	Meta-analysis (see *Table 32*), Bayesian posterior
Incremental costs and benefits of rate control compared with no rate control in symptomatic patients	Incremental cost: −£274.55; incremental QALYs: 0.0133	Lord *et al.*[116]
Cost of screening invitation	See *Table 37*	Hobbs *et al.*[19]
Cost of initial screening tests	See *Table 38*	Hobbs *et al.*,[19] NHS reference costs
Sensitivity and specificity of tests	See *Tables 4* and *5*	DTA review meta-analysis
Cost of confirmatory ECG interpreted by a GP	See *Tables 37* and *38*	Hobbs *et al.*,[19] NHS reference costs
Cost of ECG referral to a cardiologist	See *Tables 37* and *38*	Hobbs *et al.*,[19] NHS reference costs
Long-term treatment costs and benefits from the DOAC model	Depends on age, sex and stroke risk	Sterne *et al.*[37]
Non-anticoagulation costs of AF	See *Table 39*	Kassianos *et al.*[110]

Screening test

We found that results from the evidence synthesis of the DTA studies were robust to the various sensitivity analyses that we performed. However, there was a high degree of uncertainty in the estimates of sensitivity and specificity. Also, the photoplethysmography screening test was based on a single study that did not take place in a primary care setting and the study was rated as being at high risk of bias for the index test. Furthermore, photoplethysmography and modified blood pressure monitors are not currently available in most primary care settings. We therefore performed a sensitivity analysis of the comparison between age at screening, using the most cost-effective screening test and omitting (1) photoplethysmography and (2) photoplethysmography and modified blood pressure monitors.

Repeat screening

In sensitivity analyses, we explored repeated screening every 5 years for various initial screening ages (55, 60, 65, 70, 75 and 80 years) and final screening ages (55, 60, 65, 70, 75 and 80 years), giving a total of 21 repeat screening strategies. The model structure and input parameters were as for the base case; however, the total proportion of diagnosed or undiagnosed AF at each age (see *Table 40*) was adjusted to allow for the AF detected at previous screening rounds. We considered a single-age cohort that may be screened at various ages as it gets older. We assumed that total (diagnosed and undiagnosed) prevalence increases with age in this cohort according to the patterns reported in *Table 20*, which reflects increasing incidence with age. We subtracted from this total prevalence the proportion of patients diagnosed in previous screening rounds and also the proportion of patients diagnosed in the absence of screening, assumed to be a 1% diagnosis rate per year based on the control arm of the SAFE study[19] (see *Rate of diagnosis of atrial fibrillation in the absence of screening*), which was varied to 5% in sensitivity analysis. Comparisons across different screening strategies related to this age cohort and therefore we needed to consider the cost and QALY implications for the cohort as a whole to make fair comparisons. This was achieved by using population estimates reported in *Table 21* to obtain population costs and QALYs for each screening interval, discounted at 3.5% back to the time of the initial screen. The total population costs and QALYs for each strategy were the sum of the costs and QALYs for each screening interval for a given strategy, with the last screening interval representing the remaining lifetime time horizon.

Uptake of systematic opportunistic screening

In a sensitivity analysis, we used a lower uptake of systematic opportunistic screening (0.305, 95% CI 0.282 to 0.329), as reported in Morgan and Mant,[82] which was assumed to be the same for all age groups; this compares to the base case in which we used the results from the SAFE study[19] (< 75 years age group: 0.699, 95% CI 0.681 to 0.716; ≥ 75 years age group: 0.670, 95% CI 0.648 to 0.691).

Uptake of electrocardiograms in those with a positive screening result

In the base case, we assumed that all those with a positive screening test result would be offered and would accept a diagnostic ECG to confirm the diagnosis. The SAFE study[19] found that only 72.5% of those with an irregular pulse and no previous AF diagnosis agreed to have an ECG test. We therefore conducted a sensitivity analysis assuming that only 72.5% of those with a positive screening test result will agree to have an ECG. The remaining 27.5% are assumed to be a mixture of true positives and false positives (in the same proportion as for those who agree to have an ECG).

Diagnosis in the absence of screening or atrial fibrillation-related stroke

In the base case, we included an annual rate of diagnosis in the absence of screening of 1% based on the control arm of the SAFE study.[19] We increased this to 5% in sensitivity analysis, as assumed by Aronsson *et al.*[100]

Previous history of stroke in UK screen-detected atrial fibrillation

We investigated the robustness of our conclusions to using an estimate of the previous history of stroke in UK screen-detected AF based on the random-effects meta-analysis of screening studies (0.112, 95% CrI 0.072 to 0.170); in the base case we used the estimate of the previous history of stroke from the SAFE study[19] alone (0.074, 95% CI 0.032 to 0.116).

Proportion of oral anticoagulants prescribed that are directly acting oral anticoagulants

In the base case, we assumed that 75% of OACs prescribed are DOACs. Prescribing practice is changing, with the DOACs becoming the OACs of choice for newly diagnosed AF. However, warfarin is still prescribed, with some patients/practitioners preferring the regular monitoring of warfarin. We performed a sensitivity analysis of the proportion of OACs that are DOACs, using assumed values of 50% and 100%.

Uptake of oral anticoagulants

In the base case, 74% of patients who are eligible for OACs receive them, 13% are contraindicated or do not wish to take them and the remaining 13% do not receive OACs but could do. In sensitivity analyses, we explored the impact of a reduced uptake of OACs of 50% and also an improvement in the uptake of OACs, so that all 87% of patients who are not contraindicated or who prefer not to take OACs receive OACs.

Hazard ratios for stroke and mortality risk for routine-detected atrial fibrillation compared with screen-detected atrial fibrillation

In sensitivity analysis, we assumed that the event rate in screen-detected patients was lower than that for routinely detected AF using the unadjusted HR for symptomatic AF (representing non-screen-detected AF) compared with asymptomatic AF (representing screen-detected AF) found by Flaker *et al.*[35] of 1.31 (95% CI 1.04 to 1.65), instead of a HR of 1, as assumed in the base case.

Discussion

Summary of results from the review of studies on the natural history of atrial fibrillation screening

We identified several studies that estimated the prevalence of AF by age and sex. All showed similar trends, although there was variation in the overall prevalence level, most likely because of the method of measurement used together with population differences. The studies with the most intensive ascertainment of AF diagnosis gave the highest estimates of prevalence and were in line with each other.[2,125] We used the estimates from the most comprehensive and recent of these studies[2] in our model to represent the overall prevalence of AF (both diagnosed and undiagnosed) by age and sex.

Uptake of systematic population screening was fairly consistent across studies, with a mean estimate of 64% (95% CrI 54% to 73%), although there was a strong effect of age seen in the SAFE study,[19] with a lower uptake of screening in the ≥ 75 years age group of 43% (95% CI 41% to 45%) compared with uptake in the < 75 years age group of 61% (95% CI 59% to 63%). In contrast, there was a high degree of variability across studies in the uptake of systematic opportunistic screening, ranging from 30% to 70% (regardless of age), raising concern that the high level of opportunistic screening achieved in the SAFE study[19] may not be seen in practice. Furthermore, with the introduction of computerised flags for a range of health issues, GPs may be less likely to respond to computerised flags as a result of 'alert fatigue'.[133]

The proportion of AF that is undiagnosed was estimated to be 35%; however, this estimate was very variable across screening studies, reflecting the method of screening and test used as well as differences in populations. As expected, the higher estimates came from studies with more intensive screening tests (e.g. long-term continuous monitoring). It is clear that if the objective is to detect paroxysmal AF then long-term continuous monitoring is necessary. Whether or not this is feasible in a population-based screening setting, however, is unclear.

As expected, our results suggest that a high proportion of screen-detected AF is likely to be asymptomatic. Asymptomatic AF has a lower risk of stroke and mortality; however, after adjusting for other risk factors, there is no evidence that the risk of stroke or mortality depends on whether AF is asymptomatic or not.[35] We explored the robustness of our conclusions to this assumption in a sensitivity analysis.

We estimated that paroxysmal AF progresses to chronic (persistent or permanent) AF at a rate of 0.15 per year. There is evidence of a 'dose–response' relationship, with stroke and mortality risk increasing as AF progresses from paroxysmal to persistent to permanent AF.

Strengths and limitations

Our review of the natural history of AF and of AF screening covered a very broad set of review questions to identify evidence relevant to our economic model. Our review focused on evidence relevant to screening populations in a primary care setting in the UK. When appropriate, we pooled results from studies in meta-analysis. We also took care to account for 'confounding by treatment' in estimating mortality and stroke risk according to type of AF and whether symptomatic or not.

Because of the scale and scope of this review, it was only possible for a single reviewer (NJW) to review the literature, and so it was not possible to discuss inclusion with another reviewer. We ran searches from 2000 onwards only, and so we may have missed relevant evidence published before this date. However, our focus was on understanding the epidemiology of AF and screening for AF in 2016, and so we wished to find the most contemporary evidence available.

Many inputs for the economic model relied on a single trial (the SAFE study[19]). Our model tracks AF diagnosed as a result of acute events, such as stroke, but does not estimate serendipitous detection of AF in the absence of screening in the base case. There was no evidence on the proportion of 12-lead ECGs interpreted by a GP that are referred to a cardiologist to help with interpretation, and we had to rely on expert opinion. There was scant information on patient characteristics of screen-detected AF cases, and there was no data on the joint distribution of different risk factors; we therefore had to assume that they were independent (which is unlikely, as, for example, those with a previous history of MI may be more likely to have a previous history of stroke as well).

We assumed that the HR for stroke comparing chronic AF with paroxysmal AF was the same as that for mortality, although there was evidence that this assumption was reasonable.[147] We assumed that the stroke and mortality rates were the same in screen-detected AF as in routinely detected AF, once risk factors have been accounted for, based on the results from Flaker *et al.*[35] We explored the robustness of our conclusions to this assumption in a sensitivity analysis.

If primary care screening were to be adopted by the NHS, then it would be necessary for all practices to acquire and maintain an ECG machine. We did not include this capital cost in the model, but we note that this would be a consideration in the decision to introduce a national screening programme. We took a NHS perspective, excluding the costs of participants, carers and other agencies. Attending for screening is likely to have a financial impact on participants in terms of travel and time to attend screening, particularly if they have to take time off work to attend.

Findings in the context of previous research

The Irish HIQA study[25] used a German study[125] to estimate AF prevalence, whereas we used a Swedish study;[2] however; as noted, the results from the two studies were very similar (see *Prevalence of atrial fibrillation in the UK by age and sex*).

Previous systematic reviews were identified for only two of our review questions (prevalence of AF[90] and cardiovascular risk for symptomatic vs. asymptomatic AF[124]). Our findings were in line with the results from these reviews.

Chapter 6 Results of the economic evaluation

Base-case results

We ran our economic evaluation using 10,000 simulation samples for a single screen invitation at one of various different ages: 55, 60, 65, 70, 75 and 80 years. The results are presented for each of these ages in *Tables 41–46* respectively. These tables show, for each of the screening strategies and screening tests, the mean incremental costs, incremental benefits (QALYs), INB and ICER (see *Outcomes of the economic evaluation*). All results are presented per 1000 individuals eligible for screening. The results are ordered according to mean INB, so that strategies that are closer to the top of the table are more cost-effective than strategies lower down the table. In nearly all cases, screening for AF has a positive INB at a willingness-to-pay threshold of £20,000, suggesting that a population-based screening programme is likely to be a cost-effective use of resources compared with no screening. The exception was using 12-lead ECG as a systematic population screening test in 55-year-olds. Regardless of age and screening test, systematic opportunistic screening has lower mean incremental costs and higher mean incremental QALYs and hence a higher INB than systematic population screening. As age at screening increases, incremental costs increase, but this is outweighed by the increase in incremental benefits, leading to an increase in INB and a reduction in the ICER. By age 70 years, all CrIs around the INB are positive, indicating very little uncertainty that screening is cost-effective compared with current practice (whatever format the screening takes). The CIs around the incremental costs and QALYs for individual screening strategies are quite wide. However, the INB is highest for photoplethysmography, closely followed by modified blood pressure monitors and then nurse pulse palpation and 12-lead ECG (as a screening test).

We explored the uncertainty in our findings using CEACs, which display the probability of different screening strategies being most cost-effective at a range of willingness to pay per QALY thresholds. *Figures 22–24* show the CEACs for each age at invitation to screening. Uncertainty in the results can also be seen in the cost-effectiveness planes presented in *Appendix 10*.

For all age groups and for all willingness-to-pay thresholds > £10,000, systematic opportunistic screening has a higher probability of being cost-effective than systematic population screening, and is also superior to no screening. Systematic opportunistic screening using photoplethysmography as the screening test had the highest probability of being cost-effective at all thresholds > £10,000, and this result becomes more certain (higher probability) as the age at screening increases.

To inform a decision about which age to screen (if we were to choose a single age at which to screen) we considered a specific age cohort (e.g. 55-year-olds in 2016) and computed the total population incremental costs and QALYs for this cohort by multiplying the incremental costs and QALYs by the population size at each candidate screening age (those still alive and eligible for screening). *Table 47* shows the age cohort population attributable mean incremental costs, QALYs and net benefits for different screening ages (all under opportunistic screening with photoplethysmography). Mean incremental costs increase with age until age 75 years and begin to fall at age 80 years. Mean incremental QALYs increase with age at screening, as does the INB. *Figure 25* shows the CEACs comparing different ages at screening for a single one-off screen. The cost per screen-detected case decreases from £55,637 to £31,135 as the age of screening increases from 55 to 80 years (see *Table 47*). Note that these costs include lifetime treatment costs and lifetime event costs, as well as the costs of the screening strategy, We should be cautious in interpreting results for the older and younger age groups, as many of the model inputs are based on studies with individuals aged 65–75 years, and we have assumed that some model inputs (such as the proportion undiagnosed) apply across all age ranges. Of course, these results assume that we screen only once, whereas the format of many screening programmes is to screen at repeated intervals. We present the results for repeated screening in the sensitivity analyses (see *Screening test*).

TABLE 41 Base-case results for a single screen at age 55 years: mean incremental costs, incremental QALYs, INB for a willingness to pay per QALY threshold of £20,000 and ICER[a]

Test/interpreter	Incremental costs (£) (95% CrI)	Incremental QALYs (95% CrI)	INB at £20,000 (£) (95% CrI)	ICER (£) (95% CrI)
Systematic opportunistic photoplethysmography	8078 (4232 to 15,401)	0.89 (0.43 to 1.5)	9700 (1835 to 19,135)	9511 (6000 to 17,053)
Systematic opportunistic modified blood pressure monitor	7459 (3900 to 14,485)	0.85 (0.41 to 1.5)	9546 (2313 to 18,741)	9155 (5990 to 15,927)
Systematic opportunistic pulse palpation – nurse	8129 (4329 to 15,012)	0.81 (0.38 to 1.4)	8152 (1026 to 17,007)	10,509 (6635 to 18,043)
Systematic opportunistic single-lead ECG – automatic/algorithm	9607 (5997 to 16,549)	0.86 (0.41 to 1.5)	7603 (375 to 16,797)	11,824 (7751 to 19,385)
Systematic opportunistic single-lead ECG – nurse	9845 (6264 to 16,512)	0.83 (0.39 to 1.4)	6661 (–445 to 15,784)	12,710 (8246 to 20,850)
Systematic opportunistic single-lead ECG – GP	10,326 (6698 to 17,119)	0.83 (0.38 to 1.5)	6353 (–943 to 15,739)	13,284 (8609 to 22,020)
Systematic opportunistic single-lead ECG – cardiologist	10,884 (7378 to 17,717)	0.85 (0.41 to 1.5)	6187 (–1029 to 15,361)	13,594 (8895 to 22,016)
Systematic opportunistic > 1- and < 12-lead ECG – cardiologist	11,659 (8055 to 18,769)	0.89 (0.43 to 1.5)	6129 (–1272 to 15,646)	13,985 (9168 to 22,536)
Systematic opportunistic > 1- and < 12-lead ECG – automatic/algorithm	9176 (5848 to 15,376)	0.73 (0.29 to 1.3)	5520 (–1278 to 14,137)	13,563 (8620 to 23,626)
Systematic opportunistic 12-lead ECG – automatic/algorithm	11,153 (7867 to 17,655)	0.8 (0.39 to 1.4)	4942 (–1762 to 13,479)	14,838 (9665 to 24,157)
Systematic opportunistic 12-lead ECG – nurse	12,389 (8821 to 19,339)	0.86 (0.41 to 1.5)	4840 (–2465 to 14,017)	15,451 (9997 to 25,220)
Systematic opportunistic 12-lead ECG – GP	13,295 (9685 to 20,413)	0.89 (0.43 to 1.5)	4493 (–2936 to 14,076)	16,032 (10,388 to 26,067)
Systematic population photoplethysmography	8882 (6256 to 13,741)	0.57 (0.28 to 0.98)	2517 (–2747 to 8664)	16,740 (10,565 to 28,320)
Systematic population modified blood pressure monitor	8405 (6050 to 12,989)	0.55 (0.27 to 0.95)	2498 (–2166 to 8430)	16,559 (10,594 to 27,166)
Systematic population pulse palpation – nurse	9040 (6418 to 13,506)	0.52 (0.25 to 0.91)	1400 (–3321 to 7078)	18,721 (11,633 to 31,544)
Systematic population single-lead ECG – automatic/algorithm	9813 (7431 to 14,313)	0.55 (0.27 to 0.95)	1222 (–3461 to 7121)	19,200 (12,086 to 31,768)
Systematic population single-lead ECG – nurse	9933 (7532 to 14,356)	0.53 (0.25 to 0.92)	651 (–3948 to 6524)	20,396 (12,661 to 34,726)
Systematic population single-lead ECG – GP	10,154 (7798 to 14,608)	0.53 (0.24 to 0.94)	540 (–4176 to 6596)	20,770 (12,763 to 35,776)
Systematic population > 1- and < 12-lead ECG – cardiologist	10,969 (8624 to 15,539)	0.57 (0.28 to 0.98)	436 (–4312 to 6530)	20,804 (13,104 to 34,228)
Systematic population single-lead ECG – cardiologist	10,587 (8248 to 14,985)	0.55 (0.27 to 0.94)	359 (–4306 to 6220)	20,939 (13,164 to 34,802)
Systematic population > 1- and < 12-lead ECG – automatic/algorithm	9401 (7238 to 13,381)	0.47 (0.19 to 0.85)	22 (–4334 to 5553)	22,244 (13,222 to 41,982)
Systematic population 12-lead ECG – automatic/algorithm	10,644 (8496 to 14,872)	0.52 (0.25 to 0.89)	–324 (–4617 to 5210)	22,393 (13,942 to 37,455)
Systematic population 12-lead ECG – nurse	11,437 (9111 to 15,908)	0.55 (0.27 to 0.96)	–390 (–5064 to 5479)	22,536 (13,940 to 37,489)
Systematic population 12-lead ECG – GP	12,018 (9662 to 16,578)	0.57 (0.28 to 0.98)	–612 (–5352 to 5494)	22,850 (14,240 to 37,953)

a All results are per 1000 eligible individuals. Results are ordered by the highest INB.

TABLE 42 Base-case results for a single screen at age 60 years: mean incremental costs, incremental QALYs, INB for a willingness to pay per QALY threshold of £20,000 and ICER[a]

Test/interpreter	Incremental costs (£) (95% CrI)	Incremental QALYs (95% CrI)	INB at £20,000 (£) (95% CrI)	ICER (£) (95% CrI)
Systematic opportunistic photoplethysmography	11,797 (5750 to 24,270)	1.4 (0.68 to 2.3)	15,638 (2774 to 30,431)	8991 (5492 to 17,158)
Systematic opportunistic modified blood pressure monitor	11,004 (5313 to 22,795)	1.3 (0.65 to 2.3)	15,238 (2891 to 29,365)	8755 (5472 to 16,976)
Systematic opportunistic pulse palpation – nurse	11,566 (5764 to 23,047)	1.3 (0.61 to 2.2)	13,567 (1685 to 27,470)	9667 (5993 to 18,138)
Systematic opportunistic single-lead ECG – automatic/algorithm	13,285 (7552 to 25,169)	1.3 (0.66 to 2.3)	13,273 (908 to 27,424)	10,555 (6755 to 19,031)
Systematic opportunistic single-lead ECG – nurse	13,398 (7662 to 24,914)	1.3 (0.6 to 2.2)	12,073 (101 to 26,218)	11,153 (7092 to 19,897)
Systematic opportunistic > 1- and < 12-lead ECG – cardiologist	15,537 (9705 to 27,601)	1.4 (0.68 to 2.3)	11,913 (–948 to 26,533)	12,015 (7794 to 20,997)
Systematic opportunistic single-lead ECG – GP	13,934 (8208 to 25,553)	1.3 (0.59 to 2.2)	11,800 (–566 to 26,271)	11,540 (7380 to 20,672)
Systematic opportunistic single-lead ECG – cardiologist	14,591 (8940 to 26,301)	1.3 (0.65 to 2.3)	11,751 (–504 to 25,843)	11,751 (7599 to 20,567)
Systematic opportunistic 12-lead ECG – nurse	16,193 (10,468 to 28,095)	1.3 (0.65 to 2.3)	10,394 (–2126 to 24,573)	13,002 (8362 to 22,447)
Systematic opportunistic > 1- and < 12-lead ECG – automatic/algorithm	12,359 (6991 to 22,974)	1.1 (0.46 to 2.0)	10,315 (–883 to 23,836)	11,741 (7368 to 21,430)
Systematic opportunistic 12-lead ECG – GP	17,243 (11,388 to 29,314)	1.4 (0.68 to 2.3)	10,207 (–2657 to 24,809)	13,395 (8669 to 22,881)
Systematic opportunistic 12-lead ECG – automatic/algorithm	14,688 (9402 to 25,639)	1.2 (0.61 to 2.1)	10,149 (–1429 to 23,486)	12,589 (8133 to 21,762)
Systematic population photoplethysmography	14,727 (9256 to 25,626)	1.2 (0.59 to 2.1)	9160 (–2219 to 22,214)	13,132 (8305 to 22,912)
Systematic population modified blood pressure monitor	13,925 (8904 to 24,257)	1.1 (0.56 to 2.0)	8924 (–1840 to 21,240)	12,981 (8337 to 22,279)
Systematic population pulse palpation – nurse	14,704 (9450 to 24,990)	1.1 (0.53 to 1.9)	7180 (–3178 to 19,283)	14,401 (9015 to 24,951)
Systematic population single-lead ECG – automatic/algorithm	15,954 (10,885 to 26,343)	1.2 (0.57 to 2)	7170 (–3613 to 19,633)	14,772 (9429 to 25,119)
Systematic population > 1- and < 12-lead ECG – cardiologist	17,690 (12,571 to 28,279)	1.2 (0.59 to 2.1)	6211 (–4991 to 18,974)	15,882 (10,165 to 26,777)
Systematic population single-lead ECG – nurse	16,007 (10,912 to 25,963)	1.1 (0.52 to 1.9)	6171 (–4273 to 18,549)	15,539 (9754 to 26,641)
Systematic population single-lead ECG – GP	16,350 (11,316 to 26,385)	1.1 (0.51 to 2.0)	6057 (–4735 to 18,665)	15,793 (9904 to 27,341)
Systematic population single-lead ECG – cardiologist	17,028 (12,039 to 27,279)	1.1 (0.57 to 2.0)	5909 (–4779 to 18,249)	15,942 (10,150 to 26,877)
Systematic population 12-lead ECG – nurse	18,260 (13,230 to 28,603)	1.2 (0.57 to 2.0)	4889 (–5975 to 17,305)	17,020 (10,730 to 28,633)
Systematic population > 1- and < 12-lead ECG – automatic/algorithm	14,956 (10,255 to 24,287)	0.99 (0.4 to 1.8)	4786 (–4941 to 16,659)	16,684 (10,125 to 30,394)
Systematic population 12-lead ECG – GP	19,174 (14,031 to 29,851)	1.2 (0.59 to 2.1)	4727 (–6460 to 17,477)	17,262 (10,984 to 28,884)
Systematic population 12-lead ECG – automatic/algorithm	16,950 (12,291 to 26,468)	1.1 (0.53 to 1.9)	4676 (–5409 to 16,341)	16,873 (10,720 to 28,289)

a All results are per 1000 eligible individuals. Results are ordered by the highest INB.

TABLE 43 Base-case results for a single screen at age 65 years: mean incremental costs, incremental QALYs, INB for a willingness to pay per QALY threshold of £20,000 and ICER[a]

Test/interpreter	Incremental costs (£) (95% CrI)	Incremental QALYs (95% CrI)	INB at £20,000 (£) (95% CrI)	ICER (£) (95% CrI)
Systematic opportunistic photoplethysmography	18,325 (8870 to 36,261)	2.3 (1.2 to 4.0)	28,623 (9404 to 52,829)	8065 (5253 to 14,282)
Systematic opportunistic modified blood pressure monitor	17,250 (8250 to 34,719)	2.2 (1.1 to 3.8)	27,656 (9279 to 50,560)	7929 (5202 to 14,037)
Systematic opportunistic single-lead ECG – automatic/algorithm	19,605 (10,561 to 37,225)	2.3 (1.2 to 3.9)	25,846 (7228 to 49,288)	8971 (6002 to 15,292)
Systematic opportunistic pulse palpation – nurse	17,535 (8496 to 34,159)	2.1 (1.1 to 3.7)	25,463 (7515 to 48,022)	8451 (5567 to 14,713)
Systematic opportunistic > 1- and < 12-lead ECG – cardiologist	22,077 (12,819 to 39,987)	2.3 (1.2 to 4.0)	24,898 (5692 to 49,023)	9822 (6657 to 16,325)
Systematic opportunistic single-lead ECG – cardiologist	20,864 (11,921 to 37,995)	2.3 (1.2 to 3.8)	24,216 (5829 to 47,351)	9666 (6540 to 16,171)
Systematic opportunistic single-lead ECG – nurse	19,458 (10,534 to 36,527)	2.2 (1.1 to 3.8)	24,121 (5883 to 47,147)	9319 (6225 to 15,790)
Systematic opportunistic single-lead ECG – GP	20,061 (10,945 to 37,391)	2.2 (1.0 to 3.8)	23,979 (5150 to 47,476)	9546 (6385 to 16,181)
Systematic opportunistic 12-lead ECG – GP	23,780 (14,513 to 41,664)	2.3 (1.2 to 4.0)	23,194 (4068 to 47,338)	10,621 (7207 to 17,385)
Systematic opportunistic 12-lead ECG – nurse	22,524 (13,407 to 40,188)	2.3 (1.1 to 3.9)	22,974 (4226 to 46,757)	10,392 (7013 to 17,177)
Systematic opportunistic 12-lead ECG – automatic/algorithm	20,602 (12,146 to 37,173)	2.1 (1.1 to 3.6)	21,900 (4447 to 43,917)	10,153 (6886 to 16,774)
Systematic opportunistic > 1- and < 12-lead ECG – automatic/algorithm	17,764 (9168 to 33,310)	1.9 (0.82 to 3.5)	21,034 (3363 to 43,398)	9654 (6419 to 16,660)
Systematic population photoplethysmography	20,410 (12,045 to 36,194)	2 (1.0 to 3.5)	20,469 (3666 to 41,835)	10,466 (6992 to 17,201)
Systematic population modified blood pressure monitor	19,364 (11,487 to 34,514)	2 (0.99 to 3.3)	19,738 (3680 to 40,068)	10,380 (7018 to 17,069)
Systematic population single-lead ECG – automatic/algorithm	21,456 (13,502 to 36,907)	2 (1.0 to 3.4)	18,119 (1885 to 38,581)	11,416 (7684 to 18,469)
Systematic population pulse palpation – nurse	19,898 (11,904 to 34,611)	1.9 (0.92 to 3.2)	17,542 (1710 to 37,206)	11,193 (7444 to 18,464)
Systematic population > 1- and < 12-lead ECG – cardiologist	23,385 (15,280 to 39,067)	2 (1.1 to 3.5)	17,517 (881 to 38,814)	12,065 (8129 to 19,308)
Systematic population single-lead ECG – cardiologist	22,489 (14,718 to 37,535)	2 (1.0 to 3.4)	16,764 (734 to 37,071)	12,096 (8119 to 19,446)
Systematic population single-lead ECG – nurse	21,283 (13,365 to 36,257)	1.9 (0.93 to 3.3)	16,663 (774 to 36,763)	11,862 (7932 to 19,369)
Systematic population single-lead ECG – GP	21,686 (13,729 to 36,749)	1.9 (0.91 to 3.3)	16,662 (266 to 37,170)	12,014 (7957 to 19,798)
Systematic population 12-lead ECG – GP	24,867 (16,775 to 40,497)	2 (1.1 to 3.5)	16,035 (–611 to 37,308)	12,864 (8634 to 20,516)
Systematic population 12-lead ECG – nurse	23,773 (15,784 to 39,085)	2 (0.99 to 3.4)	15,843 (–506 to 36,866)	12,722 (8477 to 20,413)
Systematic population 12-lead ECG – automatic/algorithm	22,099 (14,713 to 36,553)	1.9 (0.93 to 3.2)	14,908 (–221 to 34,253)	12,638 (8493 to 20,272)
Systematic population > 1- and < 12-lead ECG – automatic/algorithm	19,663 (12,203 to 33,342)	1.7 (0.71 to 3)	14,120 (–1270 to 33,500)	12,513 (8172 to 21,451)

a All results are per 1000 eligible individuals. Results are ordered by the highest INB.

TABLE 44 Base-case results for a single screen at age 70 years: mean incremental costs, incremental QALYs, INB for a willingness to pay per QALY threshold of £20,000 and ICER[a]

Test/interpreter	Incremental costs (£) (95% CrI)	Incremental QALYs (95% CrI)	INB at £20,000 (£) (95% CrI)	ICER (£) (95% CrI)
Systematic opportunistic photoplethysmography	28,603 (13,079 to 61,452)	4 (2 to 6.7)	50,562 (16,224 to 92,951)	7523 (4637 to 14,429)
Systematic opportunistic modified blood pressure monitor	27,083 (12,211 to 58,475)	3.8 (1.9 to 6.4)	48,640 (15,746 to 88,859)	7444 (4593 to 14,418)
Systematic opportunistic single-lead ECG – automatic/algorithm	29,558 (14,518 to 61,314)	3.8 (1.9 to 6.5)	47,076 (13,852 to 87,578)	8063 (5109 to 15,059)
Systematic opportunistic > 1- and < 12-lead ECG – cardiologist	32,373 (16,989 to 65,313)	4 (2 to 6.7)	46,836 (12,335 to 89,163)	8572 (5518 to 15,663)
Systematic opportunistic pulse palpation – nurse	26,945 (12,448 to 57,840)	3.6 (1.8 to 6.3)	45,573 (13,829 to 85,258)	7750 (4803 to 14,705)
Systematic opportunistic single-lead ECG – cardiologist	30,738 (15,872 to 62,674)	3.8 (1.9 to 6.4)	45,278 (12,253 to 85,955)	8478 (5436 to 15,531)
Systematic opportunistic 12-lead ECG – GP	34,074 (18,690 to 67,139)	4 (2 to 6.7)	45,135 (10,685 to 87,471)	9047 (5880 to 16,183)
Systematic opportunistic single-lead ECG – GP	29,701 (14,709 to 61,343)	3.7 (1.8 to 6.4)	44,555 (10,831 to 85,574)	8407 (5364 to 15,407)
Systematic opportunistic single-lead ECG – nurse	29,012 (14,214 to 60,355)	3.7 (1.8 to 6.3)	44,479 (11,909 to 84,926)	8271 (5259 to 15,343)
Systematic opportunistic 12-lead ECG – nurse	32,497 (17,416 to 64,724)	3.8 (1.9 to 6.5)	44,225 (10,579 to 85,598)	8910 (5760 to 16,043)
Systematic opportunistic 12-lead ECG – automatic/algorithm	29,920 (15,938 to 60,096)	3.6 (1.8 to 6.1)	41,743 (10,790 to 79,834)	8768 (5663 to 15,893)
Systematic population photoplethysmography	29,357 (15,791 to 57,907)	3.4 (1.7 to 5.8)	39,574 (9632 to 76,378)	8951 (5757 to 16,068)
Systematic opportunistic > 1- and < 12-lead ECG – automatic/algorithm	26,273 (11,922 to 54,546)	3.3 (1.4 to 5.8)	39,167 (7475 to 77,335)	8472 (5367 to 15,506)
Systematic population modified blood pressure monitor	27,924 (14,994 to 55,165)	3.3 (1.6 to 5.6)	38,010 (9498 to 72,987)	8902 (5757 to 16,059)
Systematic population > 1- and < 12-lead ECG – cardiologist	32,351 (18,920 to 61,152)	3.4 (1.7 to 5.9)	36,617 (6693 to 73,385)	9907 (6472 to 17,182)
Systematic population single-lead ECG – automatic/algorithm	30,120 (16,922 to 57,932)	3.3 (1.7 to 5.7)	36,606 (7765 to 71,902)	9517 (6185 to 16,806)
Systematic population 12-lead ECG – GP	33,829 (20,408 to 62,675)	3.4 (1.7 to 5.9)	35,139 (5246 to 71,843)	10381 (6800 to 17,820)
Systematic population single-lead ECG – cardiologist	31,085 (18,078 to 58,648)	3.3 (1.7 to 5.6)	35,103 (6365 to 70,172)	9923 (6486 to 17,217)
Systematic population pulse palpation – nurse	28,086 (15,288 to 55,043)	3.2 (1.5 to 5.4)	35,056 (7396 to 69,459)	9380 (6046 to 16,712)
Systematic population single-lead ECG – GP	30,080 (16,933 to 57,832)	3.2 (1.5 to 5.6)	34,577 (5182 to 70,523)	9874 (6408 to 17,364)
Systematic population single-lead ECG – nurse	29,600 (16,701 to 56,854)	3.2 (1.6 to 5.5)	34,390 (6009 to 69,425)	9783 (6339 to 17,104)
Systematic population 12-lead ECG – nurse	32,457 (19,308 to 60,424)	3.3 (1.7 to 5.7)	34,346 (5096 to 70,369)	10,296 (6713 to 17,750)
Systematic population 12-lead ECG – automatic/algorithm	30,213 (17,969 to 56,659)	3.1 (1.6 to 5.3)	32,185 (5184 to 65,284)	10,246 (6691 to 17,608)
Systematic population > 1- and < 12-lead ECG – automatic/algorithm	27,073 (14,583 to 51,723)	2.8 (1.2 to 5.1)	29,908 (2297 to 63,258)	10,173 (6515 to 18,189)

a All results are per 1000 eligible individuals. Results are ordered by the highest INB.

TABLE 45 Base-case results for a single screen at age 75 years: mean incremental costs, incremental QALYs, INB for a willingness to pay per QALY threshold of £20,000 and ICER[a]

Test/interpreter	Incremental costs (£) (95% CrI)	Incremental QALYs (95% CrI)	INB at £20,000 (£) (95% CrI)	ICER (£) (95% CrI)
Systematic opportunistic photoplethysmography	42,351 (18,662 to 91,816)	5.8 (2.9 to 9.6)	72,655 (21,084 to 130,621)	7596 (4657 to 15,271)
Systematic opportunistic modified blood pressure monitor	40,246 (17,756 to 87,546)	5.5 (2.8 to 9.3)	69,759 (20,180 to 125,238)	7543 (4626 to 15,097)
Systematic opportunistic > 1- and < 12-lead ECG – cardiologist	46,146 (22,584 to 95,939)	5.8 (2.9 to 9.6)	68,924 (17,348 to 126,966)	8319 (5284 to 16,001)
Systematic opportunistic single-lead ECG – automatic/algorithm	42,870 (19,997 to 91,137)	5.6 (2.8 to 9.4)	68,466 (18,876 to 124,651)	7967 (4985 to 15,581)
Systematic opportunistic 12-lead ECG – GP	47,842 (24,296 to 97,625)	5.8 (2.9 to 9.6)	67,228 (15,678 to 125,343)	8643 (5540 to 16,323)
Systematic opportunistic single-lead ECG – cardiologist	43,947 (21,232 to 91,680)	5.5 (2.8 to 9.3)	66,488 (17,033 to 122,485)	8252 (5231 to 15,872)
Systematic opportunistic pulse palpation – nurse	39,527 (17,382 to 85,450)	5.3 (2.6 to 9.0)	65,815 (17,555 to 121,716)	7749 (4803 to 15,381)
Systematic opportunistic 12-lead ECG – nurse	45,808 (22,816 to 94,635)	5.6 (2.8 to 9.4)	65,642 (15,075 to 123,303)	8549 (5455 to 16,254)
Systematic opportunistic single-lead ECG – GP	42,614 (19,850 to 89,237)	5.4 (2.6 to 9.2)	65,259 (15,371 to 122,622)	8205 (5178 to 15,864)
Systematic opportunistic single-lead ECG – nurse	41,762 (19,219 to 88,078)	5.3 (2.6 to 9)	64,990 (16,451 to 120,311)	8109 (5108 to 15,768)
Systematic opportunistic 12-lead ECG – automatic/algorithm	42,379 (20,950 to 87,750)	5.2 (2.6 to 8.7)	61,734 (15,339 to 114,630)	8453 (5402 to 16,109)
Systematic population photoplethysmography	41,320 (20,660 to 84,674)	5 (2.5 to 8.4)	58,817 (13,680 to 109,183)	8574 (5469 to 16,333)
Systematic opportunistic > 1- and < 12-lead ECG – automatic/algorithm	37,634 (15,917 to 80,767)	4.8 (2.0 to 8.4)	57,416 (11,266 to 111,330)	8249 (5200 to 15,889)
Systematic population modified blood pressure monitor	39,380 (19,666 to 80,591)	4.8 (2.4 to 8.1)	56,403 (13,322 to 104,820)	8542 (5459 to 16,262)
Systematic population > 1- and < 12-lead ECG – cardiologist	44,342 (23,815 to 87,886)	5 (2.5 to 8.4)	55,851 (10,916 to 106,209)	9234 (6003 to 17,016)
Systematic population single-lead ECG – automatic/algorithm	41,705 (21,744 to 83,833)	4.8 (2.4 to 8.2)	55,234 (11,827 to 104,566)	8963 (5789 to 16,707)
Systematic population 12-lead ECG – GP	45,814 (25,266 to 89,292)	5 (2.5 to 8.4)	54,378 (9342 to 104,651)	9559 (6236 to 17,374)
Systematic population single-lead ECG – cardiologist	42,582 (22,755 to 83,758)	4.8 (2.4 to 8.1)	53,574 (10,533 to 102,432)	9242 (5996 to 17,053)
Systematic population 12-lead ECG – nurse	44,044 (24,042 to 86,123)	4.9 (2.4 to 8.2)	52,997 (8900 to 102,859)	9500 (6184 to 17,344)
Systematic population pulse palpation – nurse	39,032 (19,626 to 79,148)	4.6 (2.2 to 7.8)	52,691 (10,635 to 100,920)	8866 (5690 to 16,613)
Systematic population single-lead ECG – GP	41,321 (21,383 to 82,040)	4.7 (2.2 to 8.1)	52,607 (8892 to 102,716)	9212 (5931 to 17,047)
Systematic population single-lead ECG – nurse	40,697 (21,021 to 80,639)	4.6 (2.3 to 7.9)	52,253 (10,245 to 100,588)	9146 (5902 to 16,953)
Systematic population 12-lead ECG – automatic/algorithm	41,058 (22,397 to 80,180)	4.5 (2.3 to 7.6)	49,593 (8894 to 95,581)	9467 (6158 to 17,244)
Systematic population > 1- and < 12-lead ECG – automatic/algorithm	36,963 (18,056 to 75,050)	4.1 (1.7 to 7.3)	45,799 (5526 to 92,680)	9415 (6027 to 17,387)

a All results are per 1000 eligible individuals. Results are ordered by the highest INB.

TABLE 46 Base-case results for a single screen at age 80 years: mean incremental costs, incremental QALYs, INB for a willingness to pay per QALY threshold of £20,000 and ICER[a]

Test/interpreter	Incremental costs (£) (95% CrI)	Incremental QALYs (95% CrI)	INB at £20,000 (£) (95% CrI)	ICER (£) (95% CrI)
Systematic opportunistic photoplethysmography	53,753 (23,995 to 112,770)	8 (4.0 to 14)	106,818 (39,910 to 191,259)	6954 (4317 to 13,150)
Systematic opportunistic > 1- and < 12-lead ECG – cardiologist	57,418 (27,857 to 116,241)	8 (4.0 to 14)	103,243 (36,064 to 187,797)	7453 (4777 to 13,709)
Systematic opportunistic modified blood pressure monitor	51,182 (22,811 to 108,233)	7.7 (3.9 to 13)	102,405 (38,226 to 183,424)	6918 (4297 to 13,065)
Systematic opportunistic 12-lead ECG – GP	59,039 (29,484 to 117,796)	8 (4.0 to 14)	101,623 (34,536 to 186,152)	7676 (4967 to 13,965)
Systematic opportunistic single-lead ECG – automatic/algorithm	53,837 (25,023 to 110,826)	7.8 (3.9 to 13.0)	101,607 (36,514 to 183,878)	7210 (4554 to 13,395)
Systematic opportunistic single-lead ECG – cardiologist	54,758 (26,262 to 111,401)	7.7 (3.9 to 13.0)	99,427 (35,025 to 180,487)	7406 (4742 to 13,637)
Systematic opportunistic 12-lead ECG – nurse	56,675 (27,746 to 115,373)	7.8 (3.9 to 13.0)	98,947 (32,867 to 182,406)	7611 (4901 to 13,916)
Systematic opportunistic single-lead ECG – GP	53,199 (24,162 to 110,982)	7.5 (3.5 to 13.0)	97,416 (31,285 to 180,723)	7374 (4698 to 13,630)
Systematic opportunistic pulse palpation – nurse	49,926 (22,118 to 105,397)	7.4 (3.6 to 13.0)	97,139 (34,351 to 177,397)	7057 (4409 to 13,201)
Systematic opportunistic single-lead ECG – nurse	52,261 (24,142 to 107,827)	7.5 (3.6 to 13.0)	96,824 (32,285 to 177,826)	7308 (4644 to 13,552)
Systematic opportunistic 12-lead ECG – automatic/algorithm	52,559 (25,599 to 106,267)	7.3 (3.7 to 12.0)	92,805 (32,174 to 169,458)	7545 (4863 to 13,805)
Systematic opportunistic > 1- and < 12-lead ECG – automatic/algorithm	46,999 (19,703 to 100,172)	6.6 (2.7 to 12.0)	85,733 (23,800 to 163,137)	7406 (4718 to 13,591)
Systematic population photoplethysmography	38,163 (18,936 to 76,220)	5.1 (2.6 to 8.7)	64,791 (21,700 to 119,110)	7746 (5012 to 14,112)
Systematic population > 1- and < 12-lead ECG – cardiologist	40,317 (21,265 to 78,333)	5.2 (2.6 to 8.7)	62,695 (19,757 to 116,918)	8203 (5373 to 14,545)
Systematic population modified blood pressure monitor	36,440 (18,215 to 73,099)	4.9 (2.5 to 8.4)	62,037 (21,019 to 113,899)	7730 (5000 to 14,022)
Systematic population 12-lead ECG – GP	41,352 (22,289 to 79,374)	5.2 (2.6 to 8.7)	61,660 (18,688 to 115,840)	8425 (5536 to 14,868)
Systematic population single-lead ECG – automatic/algorithm	38,170 (19,605 to 75,072)	5 (2.5 to 8.5)	61,497 (19,796 to 113,886)	8018 (5224 to 14,341)
Systematic population single-lead ECG – cardiologist	38,718 (20,457 to 74,633)	4.9 (2.5 to 8.4)	60,141 (18,872 to 112,148)	8212 (5394 to 14,571)
Systematic population 12-lead ECG – nurse	39,836 (21,259 to 76,963)	5 (2.5 to 8.5)	59,942 (17,652 to 112,823)	8389 (5509 to 14,917)
Systematic population single-lead ECG – GP	37,649 (19,028 to 74,793)	4.8 (2.3 to 8.4)	58,923 (16,328 to 112,385)	8196 (5357 to 14,675)
Systematic population pulse palpation – nurse	35,826 (17,980 to 71,026)	4.7 (2.3 to 8.1)	58,471 (18,349 to 110,096)	7955 (5161 to 14,228)
Systematic population single-lead ECG – nurse	37,130 (19,033 to 72,996)	4.8 (2.3 to 8.2)	58,461 (16,966 to 110,431)	8152 (5316 to 14,528)

TABLE 46 Base-case results for a single screen at age 80 years: mean incremental costs, incremental QALYs, INB for a willingness to pay per QALY threshold of £20,000 and ICER[a] (continued)

Test/interpreter	Incremental costs (£) (95% CrI)	Incremental QALYs (95% CrI)	INB at £20,000 (£) (95% CrI)	ICER (£) (95% CrI)
Systematic population 12-lead ECG – automatic/algorithm	37,198 (19,847 to 71,715)	4.7 (2.3 to 7.9)	56,007 (16,998 to 105,243)	8375 (5516 to 14,807)
Systematic population > 1- and < 12-lead ECG – automatic/algorithm	33,656 (16,073 to 67,842)	4.3 (1.8 to 7.5)	51,448 (11,798 to 101,716)	8359 (5411 to 14,964)

a All results are per 1000 eligible individuals. Results are ordered by the highest INB.

(a)

(b)
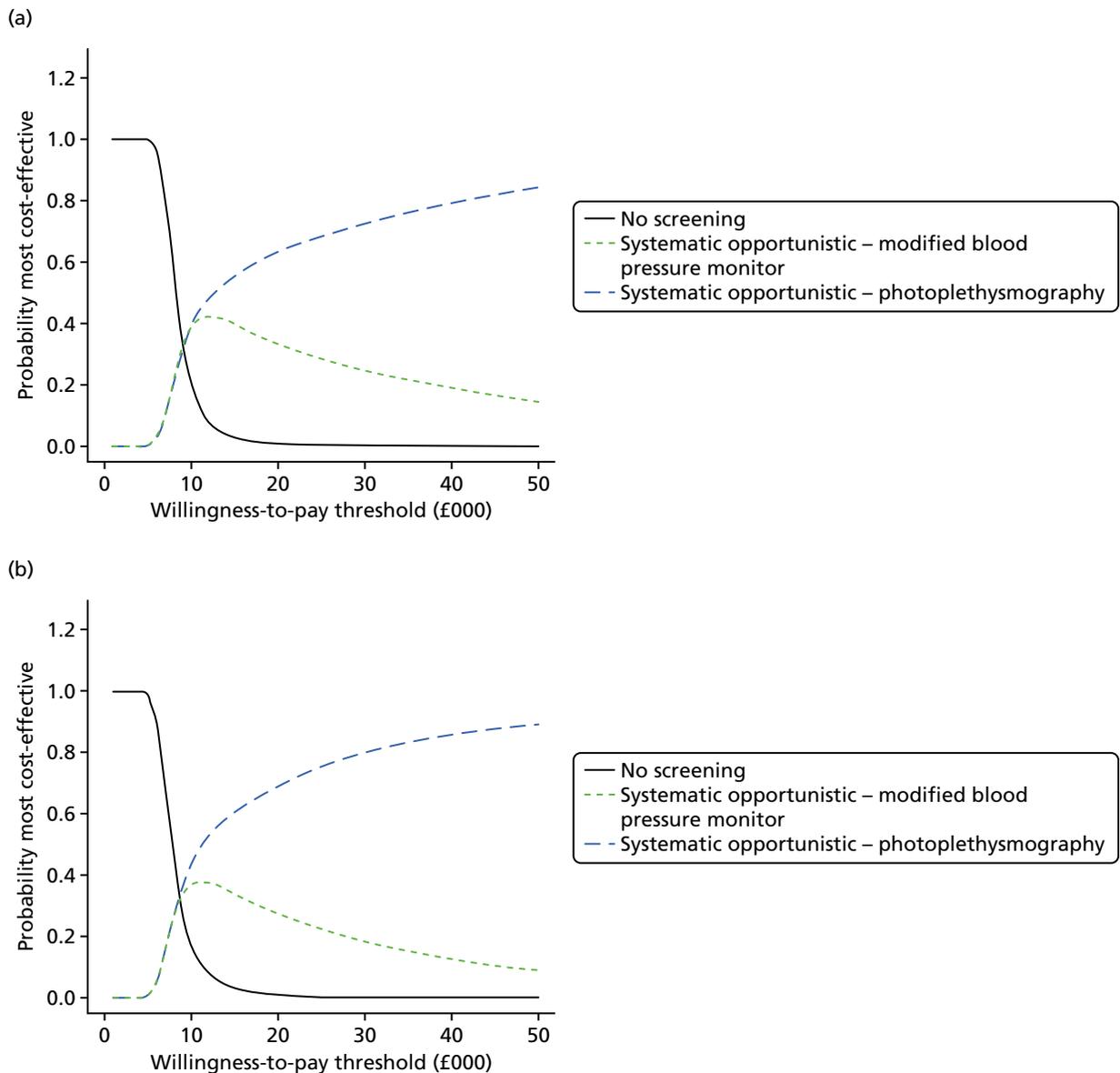
FIGURE 22 Cost-effectiveness acceptability curves for a single screen at ages (a) 55 years and (b) 60 years. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

(a)

(b)
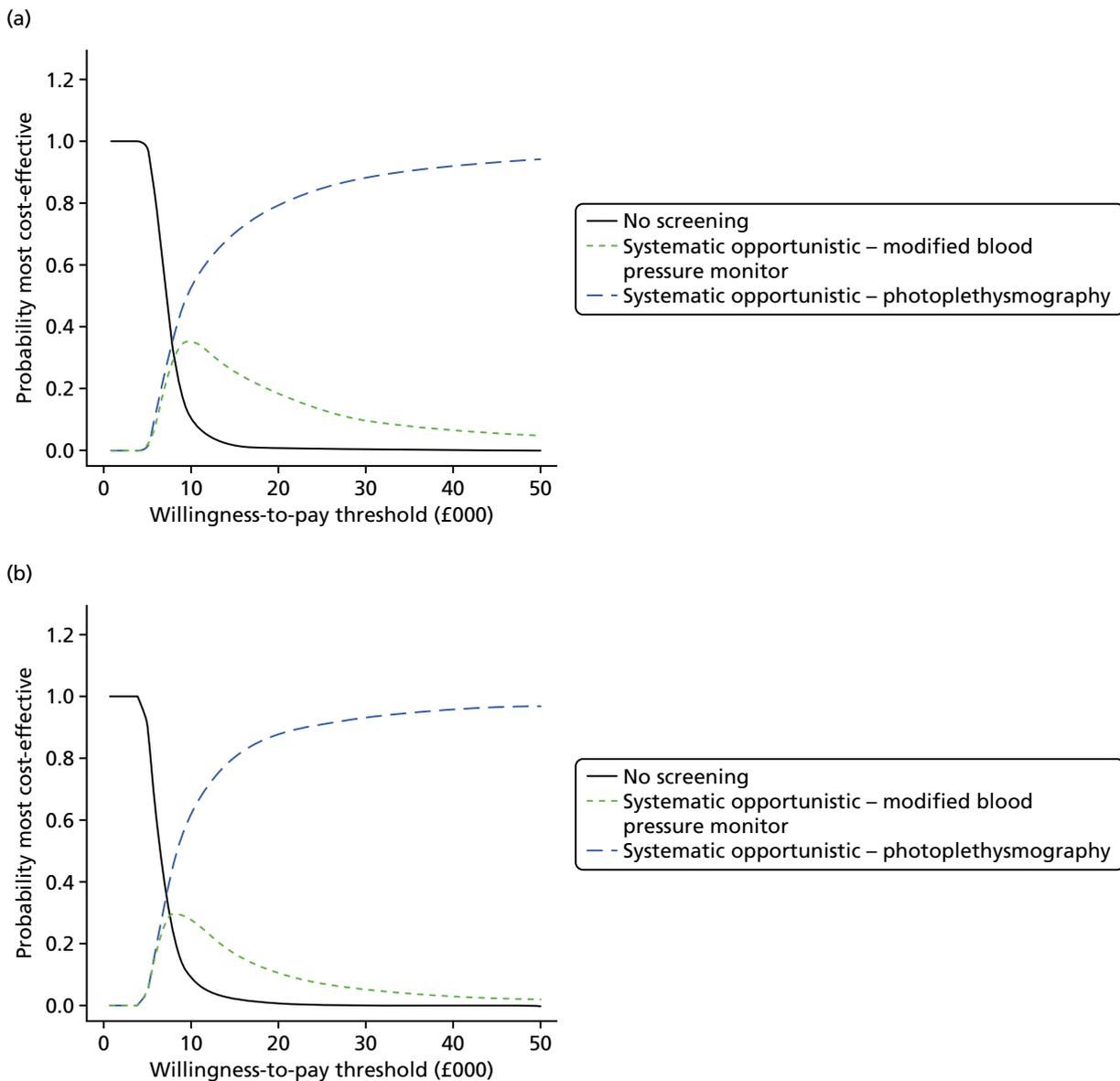
FIGURE 23 Cost-effectiveness acceptability curves for a single screen at ages (a) 65 years and (b) 70 years. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Results from sensitivity analyses

Screening test

Because the photoplethysmography screening test was based on a single study[69] that did not take place in a primary care setting, and the study was rated as being at high risk of bias for the index test, we performed a sensitivity analysis of the comparison of different ages at screening for a single one-off screen using the most cost-effective screening test omitting photoplethysmography, which was the modified blood pressure monitor. The results were very similar (see *Appendix 11*, *Using a modified blood pressure monitor as the screening test*) and the conclusions were unchanged.

Photoplethysmography and modified blood pressure monitors are not currently available in most primary care settings. Therefore, we conducted a further sensitivity analysis using the screening test that was most cost-effective (pulse palpation interpreted by a nurse) when restricting to tests that are currently typically available in primary care (i.e. excluding photoplethysmography and modified blood pressure monitors).

(a)

(b)
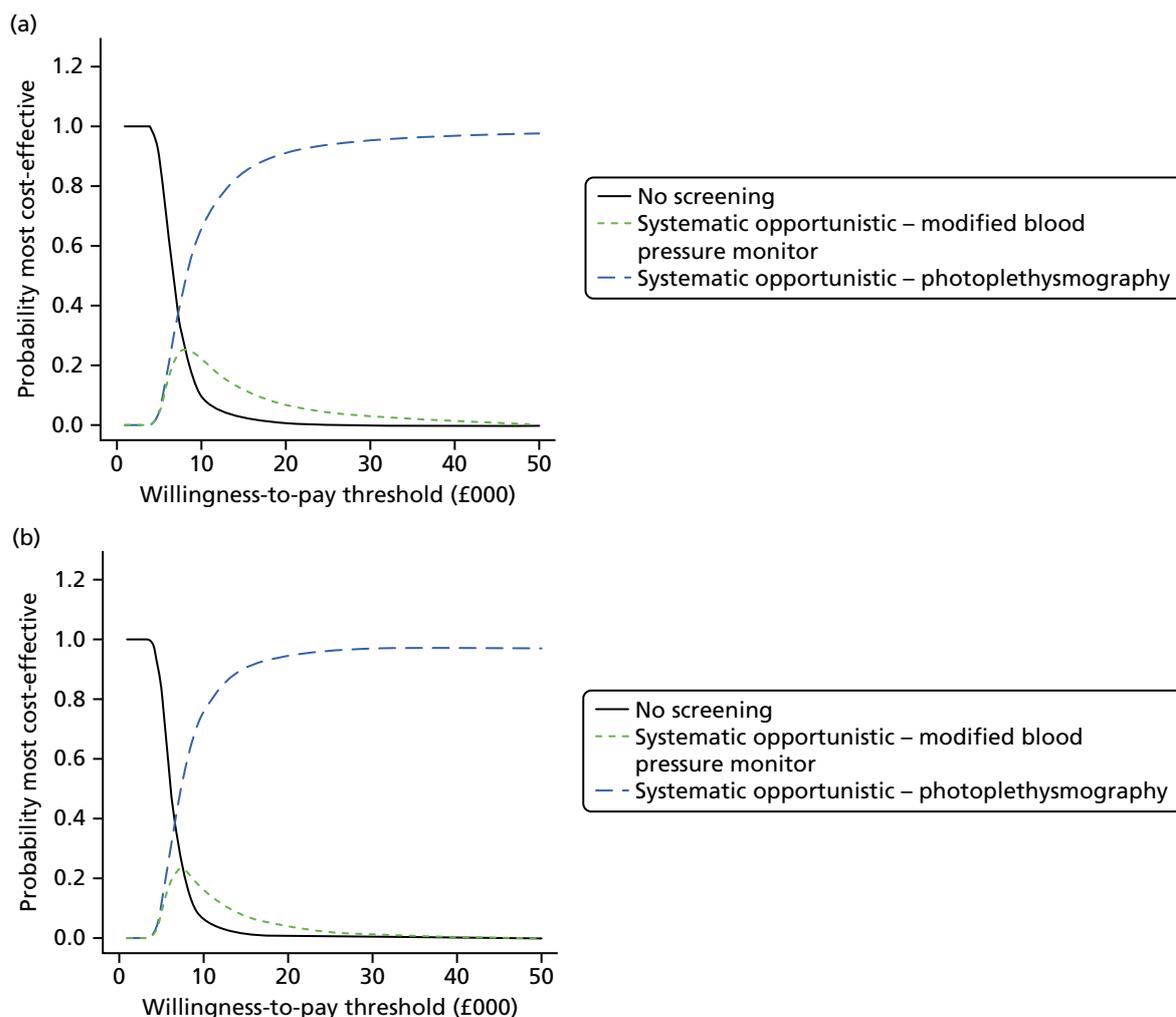
FIGURE 24 Cost-effectiveness acceptability curves for a single screen at ages (a) 75 years and (b) 80 years. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

TABLE 47 Age cohort population mean incremental costs, incremental QALYs, net benefit for a willingness to pay per QALY threshold of £20,000, number of screen-detected cases and total cost per screen-detected case for different screening ages, all using opportunistic screening with photoplethysmography relative to no screening[a]

Screening age (years)	Population size of age cohort for given ages	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Screen-detected cases (95% CrI)	Cost per screen-detected case (£) (95% CrI)
80	279,437	15,021 (6705 to 31,512)	2243 (1130 to 3799)	29,849 (11,152 to 53,445)	445 (244 to 716)	55,637 (28,215 to 97,765)
75	379,913.4	16,090 (7090 to 34,882)	2185 (1105 to 3666)	27,603 (8010 to 49,625)	449 (249 to 723)	48,371 (24,759 to 85,931)
70	466,605.4	13,346 (6103 to 28,674)	1847 (930 to 3134)	23,593 (7570 to 43,372)	345 (190 to 560)	41,349 (19,006 to 80,786)
65	633,658.8	11,612 (5621 to 22,977)	1487 (763 to 2539)	18,137 (5959 to 33,476)	278 (152 to 448)	36,949 (15,774 to 72,064)
60	620,292.8	7317 (3567 to 15,055)	851 (421 to 1450)	9700 (1721 to 18,876)	153 (84 to 250)	34,224 (13,916 to 67,342)
55	675,855	5460 (2860 to 10,409)	601 (292 to 1031)	6556 (1240 to 12,932)	104 (56 to 173)	31,135 (12,525 to 59,556)

a Age cohort population represents England and Wales 2014.
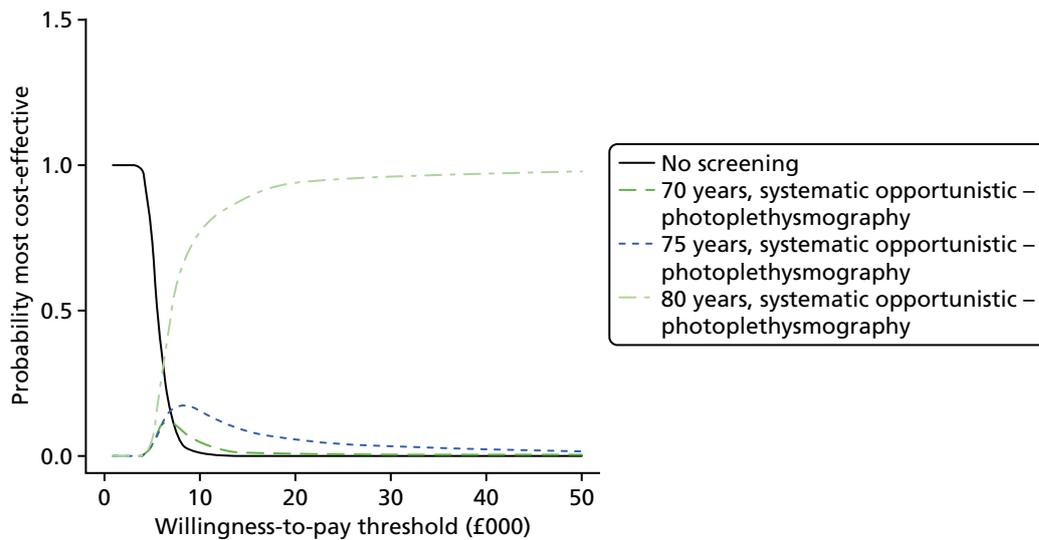
FIGURE 25 Cost-effectiveness acceptability curves comparing different ages for a single one-off screen in a given age cohort population. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

The results were very similar (see *Appendix 11*, *Using pulse palpation interpreted by a nurse as the screening test*) and, although INBs were slightly lower than in the base case, the overall conclusions were unchanged.

Repeated screening

We explored the impact of repeated screening on a given age cohort by varying the initial and final screening ages, and assuming screening every 5 years, reporting all results for the age cohort population (see *Chapter 5*, *Repeat screening*). We found that screening every 5 years from age 65 years to age 80 years had the highest probability of being cost-effective, with screening every 5 years from age 70 years to age 80 years having the second highest probability of being cost-effective (*Figure 26*). These strategies also have the highest expected INB (*Table 48*). The optimal strategy from the base case (initial screen at

FIGURE 26 Cost-effectiveness acceptability curves for screening strategies with repeated screening every 5 years for various initial and final screening ages, all using opportunistic pulse palpation by a nurse as the screening test. No screening represents no screening at any age. The results are based on incremental costs and QALYs for a single age cohort population. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

TABLE 48 Age cohort population mean incremental costs, incremental QALYs, net benefit for willingness to pay per QALY thresholds of £20,000, number of screen-detected cases and total cost per screen-detected case for repeat screening at different initial and final screening ages using opportunistic pulse palpation by a nurse as the screening test[a]

Initial and final age for repeated screening every 5 years[b]	Age cohort population incremental costs (£million) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£million) (95% CrI)	Screen-detected cases (95% CrI)	Cost per screen-detected case (£) (95% CrI)
Initial age 65 years, final age 80 years	38 (19 to 72)	4910 (2478 to 8286)	60 (22 to 107)	1241 (666 to 2026)	31,193 (20,886 to 50,491)
Initial age 70 years, final age 80 years	34 (17 to 66)	4495 (2265 to 7640)	56 (21 to 101)	1048 (562 to 1706)	32,684 (21,417 to 55,016)
Initial age 60 years, final age 80 years	38 (20 to 70)	4685 (2387 to 7903)	56 (20 to 101)	1316 (704 to 2129)	29,175 (19,732 to 46,589)
Initial age 55 years, final age 80 years	36 (19 to 65)	4307 (2197 to 7312)	50 (18 to 91)	1349 (724 to 2189)	27,240 (18,659 to 42,579)
Initial age 65 years, final age 75 years	31 (15 to 58)	3848 (1925 to 6516)	46 (16 to 84)	890 (475 to 1451)	34,963 (23,189 to 57,823)
Initial age 75 years, final age 80 years	27 (13 to 52)	3578 (1788 to 6112)	45 (16 to 82)	782 (419 to 1278)	34,374 (21,781 to 59,434)
Initial age 60 years, final age 75 years	32 (16 to 59)	3833 (1952 to 6502)	45 (15 to 82)	981 (522 to 1591)	32,892 (22,073 to 53,111)
Initial age 55 years, final age 75 years	31 (16 to 57)	3624 (1858 to 6160)	41 (14 to 76)	1030 (553 to 1671)	30,899 (20,986 to 48,669)
Initial age 70 years, final age 75 years	25 (12 to 49)	3170 (1584 to 5405)	39 (13 to 71)	678 (362 to 1106)	36,516 (23,330 to 64,354)
Initial age 60 years, final age 70 years	24 (12 to 45)	2843 (1444 to 4840)	33 (11 to 61)	641 (341 to 1046)	38,082 (25,262 to 62,211)
Initial age 55 years, final age 70 years	25 (13 to 46)	2829 (1444 to 4830)	31 (9.7 to 59)	706 (379 to 1148)	36,124 (24,342 to 57,393)
Initial age 65 years, final age 70 years	21 (10 to 41)	2613 (1301 to 4457)	31 (9.9 to 58)	533 (283 to 874)	40,076 (25,976 to 68922)
Initial age 80 years, final age 80 years	15 (6.5 to 30)	2063 (994 to 3549)	27 (8.7 to 49)	408 (218 to 666)	36,000 (21,583 to 66,120)
Initial age 75 years, final age 75 years	15 (6.7 to 32)	1925 (928 to 3329)	23 (5.9 to 44)	394 (210 to 645)	38,291 (22,648 to 74,239)
Initial age 55 years, final age 65 years	19 (9.9 to 34)	1991 (1001 to 3383)	21 (5.4 to 41)	445 (237 to 723)	42,506 (28,155 to 68,113)
Initial age 70 years, final age 70 years	13 (5.8 to 27)	1628 (783 to 2799)	20 (5.2 to 38)	303 (161 to 497)	41,676 (25,219 to 78,271)
Initial age 60 years, final age 65 years	16 (8.2 to 30)	1797 (899 to 3076)	20 (5.1 to 38)	367 (195 to 602)	44,250 (28,517 to 75,451)
Initial age 65 years, final age 65 years	11 (5.4 to 22)	1309 (628 to 2269)	15 (2.8 to 29)	244 (128 to 402)	45,785 (28,021 to 86,277)
Initial age 55 years, final age 60 years	11 (6 to 21)	1151 (570 to 1997)	12 (1.9 to 24)	223 (118 to 367)	51,966 (33,535 to 86,197)
Initial age 60 years, final age 60 years	7.1 (3.6 to 14)	749 (355 to 1302)	7.9 (0.54 to 16)	134 (71 to 224)	53,624 (32,829 to 97,800)
Initial age 55 years, final age 55 years	5.6 (2.9 to 11)	552 (256 to 978)	5.4 (0.02 to 12)	96 (49 to 160)	60,477 (37,044 to 106,296)

a Age cohort population represents England and Wales 2014.
b Each strategy assumes 5-yearly screening from the initial age to the final age or a single screen at the initial age (when the initial age is equal to the final age).

80 years with no repeat screening) had a low probability of being the most cost-effective strategy and a relatively low INB when repeated screening strategies were included in the decision. Because the results of a single screen were very similar for photoplethysmography, modified blood pressure monitors and pulse palpation by a nurse (see *Screening test*), we present only the results for pulse palpation by a nurse when comparing different repeated screening strategies, as nurse palpation has no capital costs and is currently available in primary care. The conclusions were unaltered using different screening tests.

Uptake of systematic opportunistic screening

In this sensitivity analysis we assumed a lower uptake of systematic opportunistic screening, as reported in Morgan and Mant[82] (0.305, 95% CI 0.282 to 0.329), assuming that uptake was the same for all age groups. The main impact was to change the optimal screening method from systematic opportunistic screening to systematic population screening. All other conclusions were unchanged. This suggests that our finding that systematic opportunistic screening is the optimal screening approach is sensitive to the uptake of systematic opportunistic screening. The CEACs for different screening ages are shown in *Figure 27*.

Uptake of an electrocardiogram in those with a positive screening result

We conducted a sensitivity analysis using the assumption that only 72.5% of those with a positive screening test result will agree to have an ECG, compared with 100% in the base case. We found that incremental costs, incremental QALYs and INBs were all lower than for the base case, but overall the conclusions were robust to reducing the uptake of a diagnostic ECG (see *Appendix 11*, *Uptake of an electrocardiogram in those with a positive screening result*).

Diagnosis in the absence of screening or an atrial fibrillation-related stroke

For the repeated screening strategies, we conducted a sensitivity analysis assuming an annual rate of diagnosis in the absence of screening or an AF-related stoke of 5%, as assumed by Aronsson *et al.*,[100] compared with 1% based on the control arm of the SAFE study[19] used in the base case. As one would expect, the INB decreases as the diagnosis rate in the absence of screening or AF-related stroke increases (*Table 49*). When the diagnosis rate in the absence of screening or AF-relative stroke is 5%, it is unclear whether it is more cost-effective to initiate repeated screening at age 65 years or age 70 years
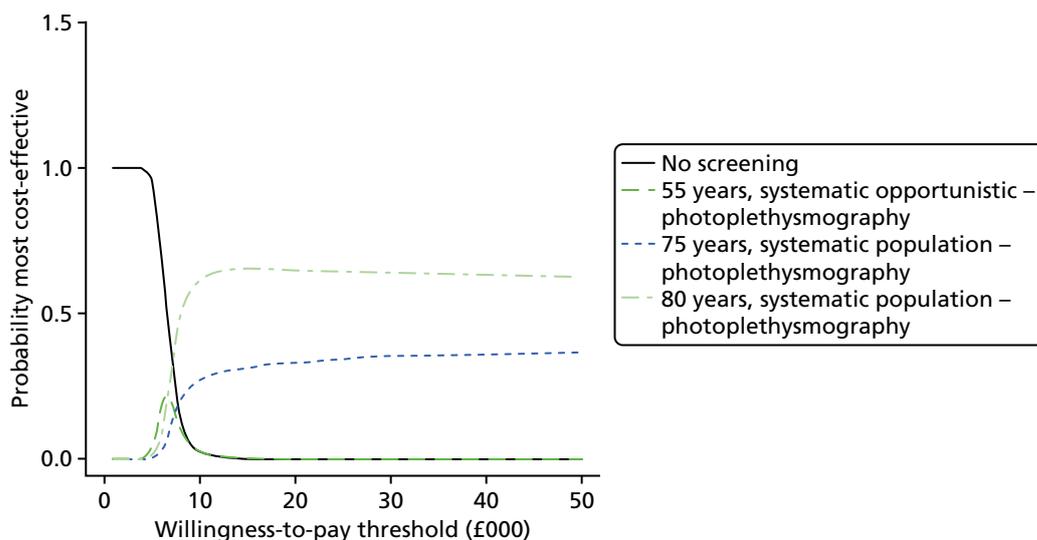
FIGURE 27 Sensitivity analysis assuming a lower uptake of systematic opportunistic screening: CEACs comparing different ages for a single one-off screen in an age cohort population. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

TABLE 49 Mean INBs for an age cohort population at the £20,000 willingness-to-pay threshold, comparing different repeat screening strategies using opportunistic pulse palpation by a nurse as the screening test[a]

Diagnosis rate in the absence of screening: 1% (base case)		Diagnosis rate in the absence of screening: 5%	
Initial and final age for repeated screening every 5 years[b]	Age cohort population INB at £20,000 (£million)	Initial and final age for repeated screening every 5 years	Age cohort population INB at £20,000 (£million)
Initial age 65 years, final age 80 years	60 (22 to 107)	Initial age 70 years, final age 80 years	47 (17 to 85)
Initial age 70 years, final age 80 years	56 (21 to 101)	Initial age 65 years, final age 80 years	45 (16 to 83)
Initial age 60 years, final age 80 years	56 (20 to 101)	Initial age 75 years, final age 80 years	41 (14 to 74)
Initial age 55 years, final age 80 years	50 (18 to 91)	Initial age 65 years, final age 75 years	38 (13 to 70)
Initial age 65 years, final age 75 years	46 (16 to 84)	Initial age 60 years, final age 80 years	37 (12 to 68)
Initial age 75 years, final age 80 years	45 (16 to 82)	Initial age 70 years, final age 75 years	35 (12 to 65)
Initial age 60 years, final age 75 years	45 (15 to 82)	Initial age 60 years, final age 75 years	32 (10 to 60)
Initial age 55 years, final age 75 years	41 (14 to 76)	Initial age 55 years, final age 80 years	29 (8.5 to 54)
Initial age 70 years, final age 75 years	39 (13 to 71)	Initial age 65 years, final age 70 years	28 (8.7 to 52)
Initial age 60 years, final age 70 years	33 (11 to 61)	Initial age 80 years, final age 80 years	27 (8.7 to 49)
Initial age 55 years, final age 70 years	31 (9.7 to 59)	Initial age 60 years, final age 70 years	26 (7.5 to 49)
Initial age 65 years, final age 70 years	31 (9.9 to 58)	Initial age 55 years, final age 75 years	26 (7.3 to 49)
Initial age 80 years, final age 80 years	27 (8.7 to 49)	Initial age 75 years, final age 75 years	23 (6.1 to 44)
Initial age 75 years, final age 75 years	23 (5.9 to 44)	Initial age 55 years, final age 70 years	21 (5.7 to 41)
Initial age 55 years, final age 65 years	21 (5.4 to 41)	Initial age 70 years, final age 70 years	20 (5.3 to 38)
Initial age 70 years, final age 70 years	20 (5.2 to 38)	Initial age 60 years, final age 65 years	17 (4.3 to 34)
Initial age 60 years, final age 65 years	20 (5.1 to 38)	Initial age 55 years, final age 65 years	16 (3.5 to 32)
Initial age 65 years, final age 65 years	15 (2.8 to 29)	Initial age 65 years, final age 65 years	15 (2.8 to 29)
Initial age 55 years, final age 60 years	12 (1.9 to 24)	Initial age 55 years, final age 60 years	10 (1.4 to 21)
Initial age 60 years, final age 60 years	7.9 (0.54 to 16)	Initial age 60 years, final age 60 years	7.9 (0.45 to 16)
Initial age 55 years, final age 55 years	5.4 (0.02 to 12)	Initial age 55 years, final age 55 years	5.4 (0.036 to 12)

a Results shown for different diagnosis rates in the absence of screening or AF-related stroke: 1% (base case) and 5%.
b Each strategy assumes 5-yearly screening from the initial to the final age or a single screen at the initial age (when the initial age is equal to the final age).

(*Figure 28*); this compares with the base case, in which there was a reasonably high probability that the optimal screening strategy was to initiate repeated screening at age 65 years (see *Figure 26*).

Meta-analysis of previous history of stroke in UK screen-detected atrial fibrillation

In a sensitivity analysis, we used the results of a random-effects meta-analysis for the proportion of screen-detected AF patients with a previous history of stroke (0.112, 95% CrI 0.072 to 0.170); this compares with 0.074 (95% CI 0.032 to 0.116) based on the SAFE study[19] used in the base case. We found no change in the conclusions and only a slight decrease in the INBs (see *Appendix 11, Meta-analysis of previous history of stroke in UK screen-detected atrial fibrillation*).
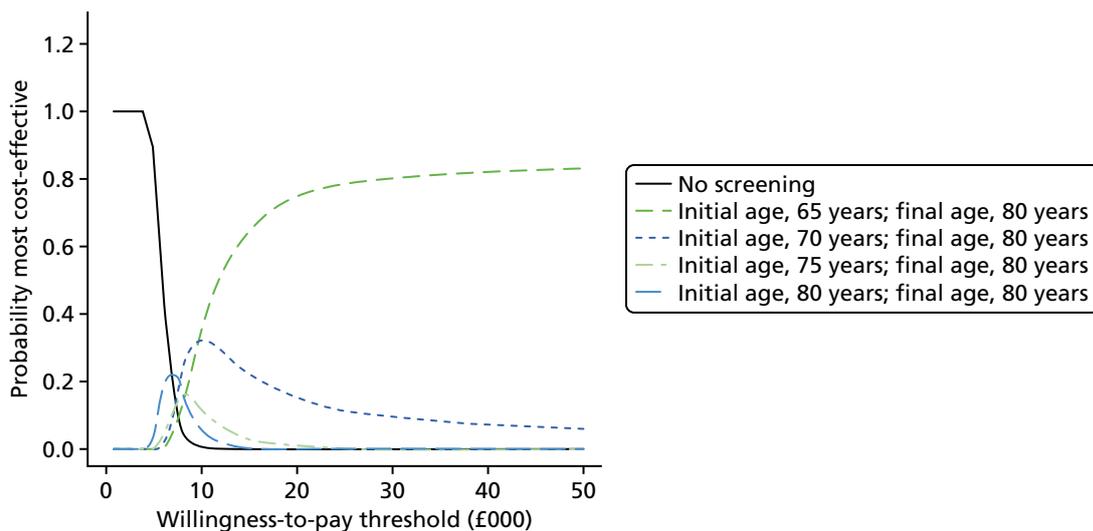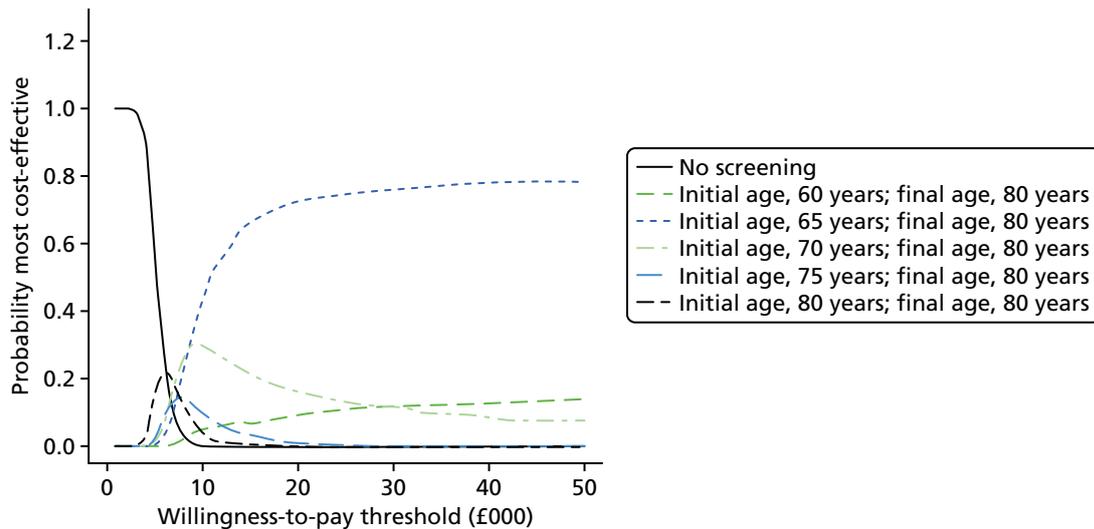
FIGURE 28 Cost-effectiveness acceptability curves for screening strategies with repeated screening every 5 years for various initial and final screening ages, all using opportunistic pulse palpation by a nurse as the screening test, for the sensitivity analysis with a 5% diagnosis rate in the absence of screening or AF-related stroke. No screening represents no screening at any age. Results are based on incremental costs and QALYs for a single age cohort population. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Proportion of anticoagulant patients receiving directly acting oral anticoagulants instead of warfarin

In the base case, we assumed that 75% of patients prescribed OACs would be prescribed DOACs. We ran sensitivity analyses on the proportion of OAC patients prescribed DOACs instead of warfarin (INR range 2–3), assuming that 50% and 100% of patients were prescribed DOACs. We found no difference in the conclusions and almost no difference in absolute results from the base case (see *Appendix 11*, *Proportion of anticoagulant patients receiving directly acting oral anticoagulants instead of warfarin*).

Uptake of oral anticoagulants

In the base case, 74% of patients eligible for OACs receive them, 13% of patients are contraindicated or do not wish to take OACs and the remaining 13% do not receive OACs but could do so. In sensitivity analyses we explored the impact of (1) a reduced OAC uptake of 50% and (2) improving the uptake of OACs so that all 87% of patients who are not contraindicated or who prefer not to take OACs receive OACs. The overall INB was lower if OAC uptake was lower, but this did not alter our conclusions (see *Appendix 11*, *Uptake of oral anticoagulants*).

Hazard ratios for stroke and mortality risk for routine-detected atrial fibrillation compared with screen-detected atrial fibrillation

In sensitivity analysis, we assumed that the event rate in screen-detected patients was lower than that for routinely detected AF using the unadjusted HR for symptomatic AF (representing non-screen-detected AF) compared with asymptomatic AF (representing screen-detected AF) found by Flaker *et al.*[35] of 1.31 (95% CI 1.04 to 1.65), instead of a HR of 1, as assumed in the base case. Our conclusions were unchanged from the base case (see *Appendix 11*, *Hazard ratios for stroke and mortality risk for routine-detected atrial fibrillation compared with screen-detected atrial fibrillation*).

Discussion

Summary of the results

Our results indicate that systematic screening for AF is likely to be cost-effective. This is because of the relatively large net benefit resulting from anticoagulation therapy for patients identified with AF and the

relatively low cost of a screening programme. We found that cheaper screening tests (such as photoplethysmography, blood pressure monitors and pulse palpation by a nurse) were more likely to be cost-effective than other screening tests, which is probably because these tests have good sensitivity as well as because they are cheaper than alternative tests. It should be noted, however, that photoplethysmography and blood pressure monitor devices are currently not typically available in primary care, and so adoption of these tests would incur a substantial upfront capital investment (see *Strengths and limitations*). Furthermore, the evidence informing the DTA of photoplethysmography was rated as being at high risk of bias for the index test, and there was no evidence available from a primary care context. The next most cost-effective screening test was found to be pulse palpation by a nurse. All of these findings rely on the use of a 12-lead ECG diagnostic test interpreted by a GP (referred to a cardiologist when the diagnosis is unclear) for individuals with a positive screening test result.

Both systematic opportunistic and systematic population screening methods were found to be cost-effective compared with no screening (current practice); however, systematic opportunistic screening was more likely to be cost-effective than systematic population screening in our base-case analyses. This result was found to be sensitive to assumptions on the uptake of opportunistic screening, which, if lower than in the base case, changes the optimal method of screening to systematic population screening.

For a given age cohort we found that, if we were to conduct a single screen, then strategies that use a higher age at screening were more likely to be cost-effective. However, when allowing for the possibility of repeated screening strategies with 5-year intervals, single screens were no longer found to be cost-effective. Instead, an initial screen at age 65 years followed by repeat screens every 5 years until age 80 years was found most likely to be cost-effective. If we assume a high (5%) diagnosis rate in the absence of screening or AF-related stroke, then it may be more cost-effective to initiate screening at age 70 years with repeated screens at 75 and 80 years. It should be noted, however, that the majority of the evidence informing the model was based on individuals aged 65–75 years, and so we are less confident of our results outside this age range.

Strengths and limitations

Our model is based on a comprehensive set of reviews of DTA studies (see *Chapters 2* and *3*), screening RCTs (see *Chapters 2* and *4*), the natural history of AF and AF screening (see *Chapter 5*) and a systematic review, network meta-analysis and cost-effectiveness model of OACs.[37] This allowed us to explore a wide range of screening tests and to incorporate net benefits of recently developed DOACs as well as warfarin.

We explored the optimal age range at which to conduct a single 'prevalent' screen and also assessed the cost-effectiveness of repeat screening rounds for a given age cohort. We considered only 5-year screening intervals for repeat screens and could have varied the screening interval in further sensitivity analyses. However, previous studies that have investigated this have found that cost-effectiveness depends on the interval of screening, with an interval of around 5 years found to be acceptable.[25,96,99]

We assumed that uptake of screening varies with age by crudely dichotomising into < 75 years and ≥ 75 years age ranges, as reported in the SAFE study,[19] which found that uptake of systematic population screening was greater in younger than older individuals. No other studies reported uptake of screening by age. Although we incorporated the effect of age for prevalence and the uptake of screening, many other model inputs were based on studies that primarily included those in the age range 65–70 years. We are therefore more confident in our results for this age range than for other age ranges.

The sensitivity analysis exploring repeated screening is not as robust as the base-case analysis. We did not directly model the incidence of new cases of AF or the progression from paroxysmal to persistent/permanent AF or from asymptomatic to symptomatic AF in the model of repeat screening – we simply used information on population size, prevalence and previous diagnosis of AF at each age. This assumes that the patterns of prevalence seen in the study by Norberg *et al.*[2] would apply to a single age cohort (i.e. no cohort or period

effects). We assumed a 1% annual diagnosis rate in the absence of screening and assumed that this did not change with increasing age. When we allowed a higher diagnosis rate in the absence of screening of 5%, there was a suggestion that initiating a repeated screening strategy at 70 years rather than at 65 years might be cost-effective. These limitations should be considered when making final decisions about repeat screening strategies.

We considered only screening programmes based on age, rather than on other risk factors. The SAFE study,[19] however, found that systematic screening targeted at high-risk individuals had a similar rate of identification of new cases of AF as the control (no screening), suggesting that targeted systematic screening is unlikely to be cost-effective. An ongoing controlled before-and-after study (due to report in 2016) in Scotland has trialled the AliveCor app used in primary care to screen high-risk individuals,[167] which will provide more evidence on this question.

We have averaged over men and women in our model to enable a single decision to be made for both sexes. The SAFE study[19] found that screening was more effective in men than in women. For women, only systematic opportunistic screening was effective, whereas for men both systematic opportunistic screening and systematic population screening were effective. Women were less likely to accept an ECG than men. If a systematic screening programme is introduced, individuals need to be fully informed of the consequences of receiving an ECG to confirm the diagnosis if the screening test is positive.

We assumed that the screening test would be conducted at a single visit to primary care and therefore would detect only a small proportion of cases of paroxysmal AF (compared with more intensive monitoring). We made this assumption because of a lack of evidence on prolonged screening tests identified in our systematic review of DTA studies. The proportion of paroxysmal AF detected by population-based screening would increase if long-term monitoring devices were used, and costs might therefore increase too. It is not clear, however, whether or not this would represent a cost-effective use of resources, as paroxysmal AF patients may be less likely to have a CHA_2DS_2-VASc score of > 1 and be eligible to receive anticoagulation therapy. Our review of natural history studies found that stroke risk is lower in patients with paroxysmal AF than in patients with persistent or permanent AF.

It should be noted that some false positives may be diagnosed with another condition (e.g. atrial flutter), and thus receive some benefit from an AF screening programme. We did not capture this benefit in our model; however, incorporating this would increase the cost-effectiveness of a population-based systematic screening programme.

We assumed that the DOAC used is apixaban because this was found to be most cost-effective in the DOAC model.[37] However, several DOACs are currently prescribed in UK practice and all licensed DOACs were found to be similar in terms of cost and benefits.[37] Our conclusions would therefore be robust to the actual DOAC that was prescribed and we envisage that a variety of DOACs would be prescribed for screen-detected AF, as is currently the case for incidentally detected AF. In the base case, we also assumed that 75% of patients prescribed OACs received DOACs, but in practice more patients may receive warfarin for various reasons, including comorbidity and tolerability. Also, in the future a higher proportion of OACs prescribed may be DOACs. Varying the proportion of OACs that were DOACs from 75% to 50% or 100% did not alter our conclusions on the cost-effectiveness of screening strategies.

We have not included capital costs of the screening test devices in our analyses. However, based on plausible assumptions about the purchase costs, lifespan and patient throughput of ECG, photoplethysmography and modified blood pressure monitor equipment (*Table 50*), we believe that the inclusion of the capital costs of equipment within the screening test cost would have little impact on our conclusions. The capital cost of the most expensive equipment (£1495 for a 12-lead ECG) is 'spread out' over the lifespan of the equipment (e.g. 5 years) among patients undergoing screening for AF and patients for whom an ECG is used for other clinical indications (*Table 51*). The estimated additional capital cost of an ECG per individual screened would be approximately £1, increasing slightly with screening age as the number of patients screened declines.

TABLE 50 Assumptions and sources for the purchase costs, lifespan and patient throughput of screening and diagnostic tests for AF

Parameter	Value	Source/comment
Uptake of systematic opportunistic screening	0.692	See *Table 40*, proportion of flagged individuals who have their pulses checked
Uptake of systematic population screening	0.641	See *Table 40*
Number of general practices	7835	Health and Social Care Information Centre[168]
Proportion of general practices with an ECG machine	0.81	Taggar *et al.*[44]
Non-AF patients using an ECG machine per 1000	34.7	Wolff *et al.*[169]
Average patients per practice	7267	Health and Social Care Information Centre[168]
ECG machine cost	£1495.00	Numed Healthcare[170]
Photoplethysmography cost	£23.88	Bonanzamarket[171]
Modified blood pressure monitor cost	£99.95	Omron[172]
ECG lifespan	5 years	Assumption
Photoplethysmography lifespan	5 years	Assumption
Modified blood pressure monitor lifespan	5 years	Assumption
Resale value (all equipment)	£0	Assumption

TABLE 51 Capital cost per patient of systematic opportunistic (Opp) and systematic population (Syst) screening strategies using photoplethysmography (PP), modified blood pressure (MBP) monitor or ECG testing

Screen age (years)	Eligible patients per practice, *n*	Opp patients screened per practice, *n*	Syst patients per practice, *n*	Non-AF screening patients using ECG, *n*	Opp PP cost per patient (£)	Syst PP cost per patient (£)	Opp MBP monitor cost per patient (£)	Syst MBP monitor cost per patient (£)	Opp ECG cost per patient (£)
55	86.26	59.69	55.43	252.16	0.08	0.09	0.33	0.36	0.96
60	79.17	54.79	50.87	252.16	0.09	0.09	0.36	0.39	0.97
65	80.88	55.97	51.97	252.16	0.09	0.09	0.36	0.39	0.97
70	59.55	41.21	38.27	252.16	0.12	0.13	0.49	0.52	1.02
75	48.49	33.55	31.16	252.16	0.14	0.15	0.60	0.64	1.05
80	35.67	24.68	22.92	252.16	0.19	0.21	0.81	0.87	1.08

The capital costs per AF patient screened for a modified blood pressure monitor (from £0.33 to £0.87) and photoplethysmography (from £0.08 to £0.21) were lower still if each practice purchased one device dedicated to AF screening. Alternatively, practices might eventually choose to replace existing blood pressure monitors with modified monitors in each consultation and treatment room and use them for screening AF and other patients.

If a screening strategy using equipment not available in practices is selected, then investment in equipment will be required to achieve nationwide primary care screening. This may act as a barrier to initiating screening. A recent review found that 81% of practices have a 12-lead ECG machine available.[44] Investment would therefore be required for the 19% of practices that do not have a machine available. This would be required under all screening strategies, because we assume that all positive screening tests would be followed by a diagnostic 12-lead ECG. Modified blood pressure monitors or photoplethysmography devices are not typically available in primary care and if these screening tests were adopted then this would represent a large capital

investment. In particular, photoplethysmography has been asessed only in a single study, which was not carried out in a primary care context and which was at high risk of bias for the index test. Investment in such a technology would be unwise without further research to better understand its diagnostic performance in a primary care screening setting. We estimate that, for photoplethysmography and modified blood pressure monitor screening strategies, for which all practices would have to purchase equipment and some practices would also have to purchase 12-lead ECG equipment for confirmatory testing, the investment needed would be approximately £2.4M and £3.0M respectively (*Table 52*). In contrast, for strategies that require only some practices to purchase 12-lead ECG equipment for screening or confirmatory testing (e.g. pulse palpation, ECG) the initial investment would be £2.2M.

Findings in the context of previous research

In line with other studies,[19,25,96,100] we found that population-based screening is likely to be cost-effective. We found that systematic opportunistic screening was more cost-effective than systematic population screening, which, again, is in line with previous studies.[19] The SAFE[19] study is the only other study we are aware of that looked at different screening tests. The DTA results from the SAFE study[19] were included in our review of DTA studies, along with the results from 14 other studies. Our model is therefore based on a more comprehensive evidence base and covers a wider range of screening tests than the SAFE model.[19]

TABLE 52 Investment costs for photoplethysmography (PP), modified blood pressure (MBP) monitor or ECG screening strategies

Screening strategy	Investment cost (£)
PP	187,100
MBP monitor	783,108
12-lead ECG	2,225,532
PP + 12-lead ECG	2,412,632
MBP monitor + 12-lead ECG	3,008,640

Chapter 7 Discussion

Summary of findings

Review of diagnostic accuracy studies

We identified 15 studies of screening tests for detecting AF, including 12-lead ECG, single-lead ECG, between 1- and 12-lead ECG, pulse palpation, modified blood pressure monitors, photoplethysmography and two-stage testing. Screening tests varied in whether they were interpreted by a cardiologist, a GP, a nurse or an automatic algorithm, although evidence was not available for every test and interpreter combination. There was a high degree of variability between studies and a high level of uncertainty in the estimates of DTA. In general, most tests had a high sensitivity, in excess of 0.9. Specificity was, in general, lower than sensitivity for all of the tests, and was lowest for pulse palpation by a nurse (specificity 0.79), a 12-lead ECG interpreted by a nurse (specificity 0.84) and photoplethysmography (specificity 0.87). Tests with the highest DOR were the 12-lead ECG (regardless of interpreter), the between 1- and 12-lead ECG (automatic or cardiologist interpretation), two-stage tests and the single-lead ECG interpreted by a GP; all of these tests had similar DORs.

In general, for a given interpreter, the results for single-lead ECGs were less accurate and more variable than ECGs with more than one lead. Nurse interpretation of single-lead ECGs performed similarly to single-lead ECGs with other interpretation methods, but nurse interpretation of 12-lead ECGs did not perform as well as 12-lead ECGs with other interpretation methods. Automatic interpretation did not have a consistent impact on test accuracy, with automatic interpretation of single-lead ECGs having a high sensitivity but variable specificity. In contrast, automatic interpretation of ECGs with more leads had good specificity but variable sensitivity. The different two-stage screening strategies all had very high specificity, but sensitivity was high only when a 12-lead ECG was used as the second-stage test.

Review of randomised controlled trials comparing screening strategies

We identified five RCTs comparing screening strategies for AF; however, only two of these provided data that could be included in our review, and only one was included in our primary outcome (the number of new AF diagnoses). The SAFE study[19] therefore remains the main source of evidence on the comparative efficacy of different screening strategies for AF. Systematic population (a 12-lead ECG interpreted by a cardiologist) and systematic opportunistic (pulse palpation followed by a 12-lead ECG interpreted by a cardiologist) screening strategies were found to be similarly effective, with an estimated 170 individuals needed to screen to detect one additional AF case compared with no screening. There was no evidence that systematic screening targeted to high-risk individuals was effective compared with no screening.

Uptake of systematic population screening was typically around 50%, although it was as high as 70% in one study. There was variability in uptake between practices (between 22% and 70% in the SAFE study[19]). Reasons for not attending for screening were varied, although older age and decreased mobility were common themes. The proportion of individuals having their pulses checked under systematic opportunistic screening varied across studies (between 30% and 66%) and between practices within studies (from 8% to 93%). The proportion of individuals consulting with their GP was not reported, so it is unclear how much these uptake rates are driven by consultation rates, GPs offering pulse palpation and uptake of pulse palpation by individuals. Of those with an irregular pulse who did not have a previous diagnosis of AF, approximately 18% did not attend for an ECG test, although, again, this varied across practices.

Subgroup analyses based on the SAFE study[19,84] indicated that both systematic population screening and systematic opportunistic screening were more effective for men than for women, and this was especially the case for systematic population screening. The efficacy of screening was not found to vary with age, despite AF prevalence being strongly associated with age.

Review of the natural history of atrial fibrillation and atrial fibrillation screening

We searched for studies to inform the natural history of AF and screening for AF for use in our economic evaluation. This comprehensive search identified 48 studies that provided data relevant to our model. The prevalence of AF was variable across studies but showed similar trends with age and sex. The studies with the most intensive ascertainment of AF gave the highest estimates of prevalence and were in line with each other.[2,125] The proportion of AF that is undiagnosed was estimated to be 35%; however, this estimate was variable across screening studies, reflecting the method of screening and test used as well as differences in populations. As expected, the higher estimates came from studies with more intensive screening tests (e.g. long-term continuous monitoring). It is clear that, if the objective is to detect paroxysmal AF, then long-term continuous monitoring is necessary.

Uptake of systematic population screening was fairly consistent across studies, with a mean estimate of 64% (95% CrI 54% to 73%), although this was shown to depend on age in the SAFE study.[19] In contrast, there was a high degree of variability in the uptake of systematic opportunistic screening, ranging from 30% to 70% (regardless of age), raising concern that the high level of opportunistic screening achieved in the SAFE study[19] may not be seen in practice.

As expected, our results suggest that a high proportion of screen-detected AF is likely to be asymptomatic. Asymptomatic AF has a lower risk of stroke and mortality; however, after adjusting for other risk factors, there is no evidence that the risk of stroke or mortality depends on whether AF is asymptomatic or not.[35] We found that our results were robust to this assumption. We estimated that paroxysmal AF progresses to chronic (persistent or permanent) AF at a rate of 0.15 per year. There is evidence of a 'dose–response' relationship, with stroke and mortality risk increasing as AF progresses from paroxysmal to persistent to permanent.

Economic evaluation

Our results indicate that both systematic opportunistic screening and systematic population screening followed by DOAC therapy when indicated are likely to be cost-effective compared with no screening (current practice). However, systematic opportunistic screening was more likely to be cost-effective than systematic population screening as long as the proportion of flagged individuals who have their pulses checked observed in the SAFE study[19] is realised in practice.

We found that photoplethysmography, modified blood pressure monitors and pulse palpation by a nurse were more likely to be cost-effective than other screening tests because they are cheaper than other screening tests while have adequate test sensitivity. This finding relies on the use of a 12-lead ECG diagnostic test interpreted by a trained GP (with referral to a cardiologist when the diagnosis is unclear) in individuals with a positive screening test result.

For a single screen of a given age cohort, we found that strategies that use a higher age of screening were more likely to be cost-effective. However, when allowing for the possibility of repeated screening strategies with 5-year intervals, single screens were no longer found to be cost-effective. Instead, an initial screen at age 65 or 70 years followed by repeat screens every 5 years until age 80 years was found to be most likely to be cost-effective, provided that compliance with screening and treatment does not decline with increasing age.

Strengths and limitations

We have conducted a comprehensive set of reviews of DTA studies (see *Chapters 2* and *3*), screening RCTs (see *Chapters 2* and *4*) and the natural history of AF and screening for AF (see *Chapter 5*), with inclusion criteria designed specifically for relevance to a general population screening programme for AF. In our economic model we also make use of a recent systematic review, network meta-analysis and cost-effectiveness

model of OACs[37] (including both DOACs and warfarin) to fully reflect the costs and benefits of anticoagulation therapy in screen-detected AF patients.

Screening tests

Although we were able to explore screening test and interpreter combinations, we were restricted by the evidence that we identified in our systematic review of DTA studies. Only a small number of studies met our inclusion criteria, not all screening test and interpreter combinations were included and, when present, few observations were reported, leading to a lack of statistical power to detect meaningful differences. None of the included studies was assessed to be at low risk of bias across all domains and to have low concerns with regard to applicability. In particular, only four studies were conducted in primary care. Although we had preplanned a wide range of subgroup analyses and sensitivity analyses, enough data were available to perform only some of these.

Although photoplethysmography had very high test sensitivity, this estimate was based on a single study[69] that was not based in a primary care setting, and for which the index test was rated as being at high risk of bias and the applicability of the patient selection was unclear. Similarly, a 12-lead ECG interpreted by a GP was evaluated in only a single study,[70] which was at high risk of bias according to patient selection. We are therefore cautious in interpreting the results from these screening tests.

Setting

We restricted our analyses to a primary care setting, but opportunistic screening could potentially be conducted in other settings, for example community centres, such as pharmacies,[104] or in secondary care such as in elderly patients attending for minor surgery.[136]

Comparative evidence

For systematic opportunistic screening, a greater proportion of the 75 newly identified cases were diagnosed outside the screening programme (44/75, 59%) than within it (31/75, 41%). One explanation for this is that GPs in the systematic opportunistic screening arm changed their usual practice to check for AF more frequently (the 'Hawthorne effect', in which behaviour changes when under observation). If this is the case, then the full benefits seen in the systematic opportunistic arm may not be realised outside the context of a RCT.

None of the RCTs identified in our review considered repeated screening, only a one-off screen. We therefore do not have any comparative evidence on the effectiveness of repeated strategies.

We have considered screening programmes based only on age, rather than on other risk factors. The SAFE study,[19] however, found that systematic screening targeted at high-risk individuals had a similar rate of identification of new AF cases as no screening, suggesting that targeted systematic screening is unlikely to be cost-effective. A controlled before-and-after study evaluating the AliveCor ECG app used in primary care to screen high-risk individuals is ongoing in Scotland (due to report in 2016),[167] which will provide more evidence on this question.

Other arrhythmias

Some false-positive patients may be diagnosed with another condition (e.g. atrial flutter), and therefore receive some benefit from an AF screening programme. We did not capture this benefit in our model; however, incorporating this would increase the cost-effectiveness of a population-based systematic screening programme.

Paroxysmal atrial fibrillation

Continuous monitoring devices are showing promise for detecting paroxysmal AF.[22] Engdahl et al.[22] found that, in patients who screened negative on a 12-lead ECG but who had two or more risk factors based on the CHADS$_2$ score, a period of 2-week monitoring using a hand-held monitoring device identified 7.4% of paroxysmal AF cases. This study specifically targeted high-risk individuals, however, and so it is unclear how

large the yield would be in a general population. In the general population, paroxysmal AF patients may be less likely to have a CHA$_2$DS$_2$-VASc score of > 1 and may be eligible to receive anticoagulation therapy. Also, our review of natural history studies found that stroke risk is lower in patients with paroxysmal AF than in those with persistent or permanent AF. If such devices were to be used in a screening context, there would need to be clear instructions for their use, and there would also be concerns about the devices not being returned.

Screening age

The studies informing many model inputs were primarily carried out with those in the age range 65–70 years, and we are therefore more confident about our results in this age range than in other age ranges.

We did not directly model the incidence of new cases of AF or the progression from paroxysmal AF to persistent/permanent AF or from asymptomatic AF to symptomatic AF in the model of repeat screening – we simply used information on population size, prevalence and previous diagnosis of AF at each age. This assumes that the patterns of prevalence seen in the study by Norberg et al.[2] would apply to a single age cohort (i.e. no cohort or period effects). We assumed a 1% annual diagnosis rate in the absence of screening, and that this did not change with increasing age. When we allowed a higher diagnosis rate in the absence of screening of 5%, there was a suggestion that initiating a repeated screening strategy at 70 years rather than 65 years might be cost-effective.

Economic model

Our review of the literature for inputs to the economic model was the most comprehensive search of the literature to date. We combined the results from multiple studies using meta-analysis when possible. However, despite this, many inputs to the economic model relied on the results from a single trial.[19] There was no evidence on the proportion of 12-lead ECGs interpreted by a GP that are referred to a cardiologist to help with interpretation, and so we had to rely on expert opinion. There was scant information on patient characteristics of screen-detected AF cases and no data on the joint distribution of different risk factors, so we had to assume that they were independent (which is unlikely, as, for example, those with a previous history of MI may be more likely to have a previous history of stroke as well).

We took a NHS perspective, excluding the costs of participants, carers and other agencies. Attending for screening is likely to have a financial impact on participants in terms of travel and time to attend screening, particularly if they have to take time off work to attend.

If a nationwide primary care screening strategy using equipment not available in practices is adopted, then investment in equipment will be required. This may act as a barrier to initiating screening. A recent review found that 81% of practices have a 12-lead ECG machine available.[44] Investment would therefore be required for the 19% of practices that do not have a machine available. This would be required under all screening strategies, because we assume that all positive screening tests would be followed by a diagnostic 12-lead ECG. Modified blood pressure monitors or photoplethysmography devices are not typically available in primary care and, if these screening tests were adopted, this would represent a large capital investment. Investment in photoplethysmography would be unwise without further research to better understand its diagnostic performance in a primary care screening setting. We estimate that, for photoplethysmography and modified blood pressure monitor screening strategies, for which all practices would have to purchase equipment and some practices would also have to purchase 12-lead ECG equipment for confirmatory testing, the investment needed would be approximately £2.4M and £3.0M respectively (see Table 52). For strategies that require only some practices to purchase 12-lead ECG equipment for screening or confirmatory testing (e.g. pulse palpation, ECG), the initial investment would be £2.2M.

Research needs

Screening test accuracy

Our study was restricted to comparing screening tests that had been evaluated in DTA studies that used a 12-lead ECG interpreted by a cardiologist as the reference standard. Only 15 studies met our inclusion criteria, and of those only four were carried out in a primary care or community setting, which is where we envisage screening taking place. When replicates existed of the same screening test/interpreter combination, heterogeneity in results was observed. This suggests a need to replicate the results based on a single study or few studies. In particular, the DTA results for photoplethysmography were based on a single study[69] rated as being at high risk of bias for the index test and conducted outside a primary care setting. Robust evidence on the performance of this device in a primary care screening setting, for a range of cut-off values, would be of merit. Such a study should also record information on hygiene, use in people with skin conditions, ease of use, durability, length of test time and sensation.

We found that a 12-lead ECG interpreted by a GP had a very good diagnostic performance compared with a 12-lead ECG interpreted by a cardiologist/specialist. However, this was based on a single study[70] that was at high risk of bias with regard to patient selection. A study to replicate this result in a screening population in a format that reflects how screening is likely to be implemented (i.e. ECGs interpreted by a trained GP with referral to a cardiologist/specialist in cases that are unclear) may be of value, although resources may be better directed towards training GPs in the interpretation of ECGs for diagnosing AF. Such a study would also be able to provide information on the proportion of ECGs that are unclear and that would thus be referred to a cardiologist/specialist, which was an assumption in our economic model.

The development of new devices to detect AF is a fast-moving area, and a horizon-scanning exercise to identify new literature as it becomes available, and to update our review, would be of value. It should be noted that we assumed that only a small proportion of paroxysmal AF would be detected by the screening strategies that we considered. However, the availability of newer technologies that allow for intensive testing over a period of time would allow a higher proportion of paroxysmal AF cases to be detected. More intensive screening tests are likely to be more costly and to incur initial capital investment, although this would not necessarily be the case, for example in the case of smartphone apps. Studies comparing the DTA of smartphone apps with that of a 12-lead ECG interpreted by a cardiologist/specialist reference standard are needed. The benefits of detecting paroxysmal AF through screening also need to be better understood, to help evaluate more intensive screening tests.

Comparative effectiveness of screening strategies for atrial fibrillation

The primary comparative evidence of effectiveness of screening strategies remains the SAFE study.[19] This is a well-conducted study, but it evaluates only a single screen for a given age range and does not follow up to measure long-term outcomes of screening or the value of repeat screening. To fully evaluate the benefits of a screening programme and change practice, there is a need for a comparative study that measures long-term outcomes (stroke and mortality) for a screening strategy compared with a no screening strategy. The ongoing STROKESTOP study should provide evidence on this over a longer follow-up period than previous studies;[26,40] however, the results are specific to the format of screening used (intermittent ECG over 2 weeks). Because an intensive screening test was used in the STROKESTOP study, a large proportion of those with newly detected AF will have paroxysmal AF. This study should therefore provide valuable evidence on the impact of screening on long-term outcomes for paroxysmal AF patients. The benefits of early detection of AF through screening will depend on the subsequent medical care and anticoagulation therapy used. The generalisability of the results of the STROKESTOP study will depend in part on how similar the use of anticoagulation therapies for AF patients is between Sweden and the UK.

Given the huge number of potential screening strategies (test, interpreter, age at initial screen, screening interval, age at final screen, systematic opportunistic screening, systematic population screening, systematic targeted screening, setting) it is not possible to design a primary research study to evaluate all combinations. Some form of modelling/simulation is therefore necessary to evaluate different options to identify those that

are most likely to be effective so that they can be evaluated further in a primary research study. The design of such a primary research study may take the form of a cluster randomised factorial trial and make use of linked electronic health records. This is likely to be challenging because of the fast-moving development of screening test technologies. Uncertainties in the economic model that a RCT could inform include the uptake of screening by age and sex, patient characteristics of those with screen-detected AF (including CHA_2DS_2-VASc score), the proportion of patients diagnosed in the absence of screening and reason for diagnosis (event, symptoms or other) and anticoagulation therapies used (including the proportion of OACs that are DOACs).

We have focused on screening in a primary care setting. Further research into alternative settings including secondary care (e.g. outpatient clinics) and community settings (e.g. pharmacies) would be of value.

Repeat screening

To thoroughly assess repeat screening strategies, good evidence is required on the incidence of new cases of AF (and the proportion of these that are undiagnosed), progression from paroxysmal to persistent/permanent AF, progression from asymptomatic to symptomatic AF and the diagnosis rate in the absence of screening that is not a result of an AF-related stroke. A large epidemiological study with linked electronic health records may be the most efficient way to provide some of this evidence; however, such a study may not identify undiagnosed AF that could be identified by screening. A nested follow-on study to a cluster RCT of screening could be conducted in those who were screened and who tested negative, by randomising individuals to a repeat screen after 5 years or not. This would provide evidence on the incidence of AF that is diagnosed and undiagnosed and the yield of a repeat screen. All previous models of repeat screening (including ours) use routinely collected data on the prevalence/incidence of AF, which likely underestimates undiagnosed AF cases.

Implementation of screening

The economic model results were sensitive to the proportion of flagged individuals who have their pulses checked. This proportion was estimated to be very different (30% and 66%) in the two studies that reported it[19,82] and a high degree of variability was seen across practices.[19] Research on the effectiveness of implementation strategies to improve the proportion of pulses that are taken among flagged individuals would be of value if a systematic opportunistic screening strategy were recommended.

Implications for practice

Screening method

Our findings support the use of systematic opportunistic screening for AF, with evidence that a screening strategy with an initial screen at age 65 years and repeated screens every 5 years until age 80 years is likely to be cost-effective. Operationalising opportunistic screening would require some training of clinicians, possibly using age-triggered prompts in the electronic medical record and simple practice-level protocols to ensure that patients in whom AF is suspected based on the screening test are offered an ECG. The practice protocol, based on national guidance, would need to specify who to invite, operationalised through the Read coding of patients to exclude individuals with a previous diagnosis of AF, as it can be distressing to patients to be invited to screening for a condition that they know they have. Similar systems are in place for other screening programmes (e.g. breast and cervical screening), suggesting that such a system could be implemented for AF screening.

Wright *et al.*[173] found that a multifaceted intervention programme aimed at health professionals in primary care led to an increase in detection of AF cases, highlighting the need for an implementation strategy to embed screening in general practices. Several studies performed screening alongside influenza vaccination,[117,118] but yield was fairly low. Another study[174] looked at changes in AF detection rates in practices that implemented the NHS Health Checks programme compared with practices that did not implement the NHS Health Checks programme and found no difference in detection rates. This is perhaps not surprising, as AF was not an aim of the NHS Health Checks programme, although it could be in the future.

Screening test

As discussed earlier, we do not consider photoplethysmography to be appropriate for a screening study until further evidence on its diagnostic performance has been collected. The next most cost-effective screening tests were modified blood pressure monitors and nurse pulse palpation. Modified blood pressure monitors are not available in most practices. If modified blood pressure monitors were to be recommended as the screening test in a national screening programme, an investment in the devices would be required (see *Chapter 6*, *Strengths and limitations*). This could be implemented over a period of time by replacing existing blood pressure monitors at the end of their lifespan with modified devices. As well as the capital cost, training would be required in their use, although the automatic interpretation algorithm would potentially be simpler to interpret than pulse palpation. A screening programme could therefore allow both screening tests as options, with practices without a modified blood pressure monitor using pulse palpation by a nurse for screening. Potentially, a nationwide screening programme could be rolled out in a stepped or otherwise randomised fashion, enabling additional evidence to be collected on the optimal screening strategies.

A recent review[44] found enthusiasm for screening in primary care but a recognised need for the training of nurses. Barriers to AF screening were listed as lack of time, workload, lack of appointments, staffing levels, access to equipment, funding, waiting times at anticoagulation clinics and the cost of anticoagulation therapies. These resource issues would need to be addressed if a national screening programme were to be adopted. Taggar *et al.*[44] found that health-care assistants and nurses were mostly responsible for conducting ECGs, with interpretation by GPs, but these staff reported that they would benefit from training. In the future, pulse palpation and use of modified blood pressure monitors and ECG machines could become a core part of training and continued professional development for nurses and health-care assistants. The review[44] found that only 81% of practices had an ECG machine. Under all of our screening strategies, we assumed that a diagnostic 12-lead ECG interpreted by a trained GP (with referral to a cardiologist/specialist in cases that are unclear) would be conducted in those who screen positive, and so an additional investment in an ECG machine would be required in the 19% of practices that do not already have one. This investment would also be used for other indications as well as AF, and so would bring additional benefits.

Other arrhythmias

A consequence of screening for AF is that other arrhythmias, such as atrial flutter, may be detected. This is not captured in our economic model. Assuming that routine care for patients with other arrhythmias is cost-effective, then the detection of these patients through screening would only increase the cost-effectiveness of a screening programme. It is important that individuals who are invited to be screened understand that other arrhythmias, as well as AF, may be detected and that, if they are diagnosed, they are fully informed of the risks and treatment options.

Future developments

New devices to detect AF are becoming available and are under evaluation (patches, smartphone/watch devices, iPads, hand-held devices), as well as devices used for other reasons that can also detect AF (pacemakers, ICDs and implantable loop recorder devices). These innovations may be of relevance to a screening programme when their diagnostic performance is better understood.

Patient perspective

To validate an AF screening programme, it is essential to raise awareness and understanding of the condition: its symptoms, impact and outcomes (stroke risk), lifestyle changes, how it can be managed and surgical implications and drugs to correct heart rate and rhythm, as well as the need for anticoagulation. Many people do not know what AF is unless they or someone close to them has personal experience of it. Therefore, any screening programme must aim to increase awareness, diagnosis, treatment and prevention. It is important that patients understand that anticoagulation therapy aims to prevent stroke and does not

provide a cure for AF or reduce symptoms. Receiving a diagnosis of AF creates anxiety for patients, and therefore appropriate counselling is important.

If a screening programme were introduced, a screening pathway would need to be developed with step-by-step guidance with regard to the process and, most importantly, health-care professionals conducting any screening process must have specialist knowledge of all aspects of the condition, treatments, symptoms and risks.

Both the screening test and follow-up diagnostic test would need to be timely and convenient to the individual. Lack of mobility and age were listed as reasons for not attending screening in the SAFE study.[19] Community screening (e.g. pharmacy or similar) may be an option to increase uptake in the population and raise awareness at the same time. Provision of transport to and from practices may also increase participation. Another possibility would be to include a screening test for AF alongside another health programme (e.g. NHS Health Check) to minimise visits to the GP. A one-stop shop approach with screening test, diagnostic test, counselling and onward referral undertaken in one session may also be more convenient for individuals and would minimise individuals being 'lost' in the screening process.

Conclusions

A national screening programme for AF is likely to represent a cost-effective use of resources. Systematic opportunistic screening is more likely to be cost-effective than systematic population screening. Nurse pulse palpation or modified blood pressure monitors (if available) would be appropriate screening tests, followed by a diagnostic 12-lead ECG interpreted by a trained GP in those who screen positive, with referral to a specialist in cases in which diagnosis is unclear. Implementation strategies to operationalise the uptake of opportunistic screening in primary care should accompany any screening recommendations.

Acknowledgements

Contributions of authors

Nicky J Welton managed the project, contributed to the development of the methods for the systematic review, conducted the review of natural history relevant to AF screening, contributed to the development of the economic model and supervised the statistical analysis and economic modelling work. She drafted the introduction, economic model methods and discussion sections of the report and contributed to all other sections of the report.

Alexandra McAleenan was responsible for the systematic review of DTA studies, developed the review protocol, screened studies for inclusion, extracted data from the studies, produced descriptive results and drafted the sections of the report describing the methods and results of the DTA review. She also assisted with the screening studies review as a second reviewer and contributed to the drafting of the sections of the report describing the methods and results of the screening studies review. She commented on all other sections of the report.

Howard HZ Thom performed the statistical analysis for the DTA review, contributed to the development of the economic model, identified cost inputs to the economic model, linked the economic model to a previous model of anticoagulation therapy, conducted the economic evaluation, drafted the statistical analysis results section of the DTA review, drafted the cost inputs to the economic model and the economic evaluation results sections of the report, and contributed to the economic model methods section.

Philippa Davies provided advice on the systematic reviews, developed the data extraction database for the DTA review, acted as the third reviewer in both systematic reviews to resolve conflicts, contributed to the risk-of-bias assessments and contributed to the writing of the systematic review sections of the report.

Will Hollingworth contributed to the development of the economic model, the identification of model inputs and the writing of the economic model and results sections of the report.

Julian PT Higgins supervised the systematic reviews, provided advice on the statistical analysis and contributed to the writing of the systematic review and results sections of the report.

George Okoli was responsible for the systematic review of screening studies, developed the review protocol, screened studies for inclusion, extracted data from the studies, produced descriptive tables of the results and drafted the sections of the report on the screening study review describing the methods and results of descriptive analyses. He also assisted with the DTA review as a second reviewer.

Jonathan AC Sterne provided advice on the statistical analysis of the DTA studies and advice on project management.

Gene Feder provided input on the review methods and model assumptions from a clinical primary care perspective and contributed to the writing of the introduction and discussion sections of the report.

Data sharing statement

All extracted data are described in the report.

References

1. Public Health England. *Atrial Fibrillation Prevalence Estimates in England: Application of Recent Population Estimates of AF in Sweden*. London: Public Health England; 2015.

2. Norberg J, Bäckström S, Jansson JH, Johansson L. Estimating the prevalence of atrial fibrillation in a general population using validated electronic health data. *Clin Epidemiol* 2013;**5**:475–81.

3. Primary Care Domain, Health and Social Care Information Centre. *Quality and Outcomes Framework – Prevalence, Achievements and Exceptions Report. England, 2014–15*; 2015. URL: http://content.digital.nhs.uk/catalogue/PUB18887/qof-1415-Report%20vl.l.pdf

4. Wilke T, Groth A, Mueller S, Pfannkuche M, Verheyen F, Linder R, *et al.* Incidence and prevalence of atrial fibrillation: an analysis based on 8.3 million patients. *Europace* 2013;**15**:486–93. http://dx.doi.org/10.1093/europace/eus333

5. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;**22**:983–8. https://doi.org/10.1161/01.STR.22.8.983

6. Raju NC, Hankey GJ. Dabigatran etexilate in people with atrial fibrillation. *BMJ* 2010;**341**:c3784. http://dx.doi.org/10.1136/bmj.c3784

7. Lin HJ, Wolf PA, Kelly-Hayes M, Beiser AS, Kase CS, Benjamin EJ, D'Agostino RB. Stroke severity in atrial fibrillation. The Framingham Study. *Stroke* 1996;**27**:1760–4. https://doi.org/10.1161/01.STR.27.10.1760

8. Healey JS, Connolly SJ, Gold MR, Israel CW, Van Gelder IC, Capucci A, *et al.* Subclinical atrial fibrillation and the risk of stroke. *N Engl J Med* 2012;**366**:120–9. http://dx.doi.org/10.1056/NEJMoa1105575

9. Vanassche T, Lauw MN, Eikelboom JW, Healey JS, Hart RG, Alings M, *et al.* Risk of ischaemic stroke according to pattern of atrial fibrillation: analysis of 6563 aspirin-treated patients in ACTIVE-A and AVERROES. *Eur Heart J* 2015;**36**:281–7a. http://dx.doi.org/10.1093/eurheartj/ehu307

10. Smolina K, Wright FL, Rayner M, Goldacre MJ. Incidence and 30-day case fatality for acute myocardial infarction in England in 2010: national-linked database study. *Eur J Public Health* 2012;**22**:848–53. http://dx.doi.org/10.1093/eurpub/ckr196

11. National Institute for Health and Care Excellence. *Atrial Fibrillation: the Management of Atrial Fibrillation*. Clinical guideline CG180. London: NICE; 2014.

12. Canadian Agency for Drugs and Technologies in Health. *Oral Anticoagulants in Nun-valvular Atrial Fibrillation: CADTH Recommendations and Clinical Practice Guidelines*; 2016. URL: www.cadth.ca/oral-anticoagulants-non-valvular-atrial-fibrillation-cadth-recommendations-and-clinical-practice (accessed 27 April 2017).

13. Kirchhof P, Benussi S, Koteeha D, Ahlsson A, Atar D, Casadei B, *et al.* 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016;**37**:2893–962. https://doi.org/10.1093/eur-heartj/ehw210

14. Nieuwlaat R, Prins MH, Le Heuzey JY, Vardas PE, Aliot E, Santini M, *et al.* Prognosis, disease progression, and treatment of atrial fibrillation patients during 1 year: follow-up of the Euro Heart Survey on atrial fibrillation. *Eur Heart J* 2008;**29**:1181–9. http://dx.doi.org/10.1093/eurheartj/ehn139

15. Lee S, Shafe AC, Cowie MR. UK stroke incidence, mortality and cardiovascular risk management 1999–2008: time-trend analysis from the General Practice Research Database. *BMJ Open* 2011;**1**:e000269. http://dx.doi.org/10.1136/bmjopen-2011-000269

16. Ben Freedman S, Lowres N. Asymptomatic atrial fibrillation: the case for screening to prevent stroke. *JAMA* 2015;**314**:1911–2. http://dx.doi.org/10.1001/jama.2015.9846

17. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest* 2010;**137**:263–72. http://dx.doi.org/10.1378/chest.09-1584

18. Lip GY. The CHA₂DS₂-VASc score for stroke risk stratification in patients with atrial fibrillation: a brief history. *Eur Heart J* 2015;**36**:2880–5. http://dx.doi.org/10.1093/eurheartj/ehv431

19. 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 Technol Assess* 2005;**9**(40). https://doi.org/10.3310/hta9400

20. Andrade JG, Field T, Khairy P. Detection of occult atrial fibrillation in patients with embolic stroke of uncertain source: a work in progress. *Front Physiol* 2015;**6**:100. http://dx.doi.org/10.3389/fphys.2015.00100

21. 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. *Thromb Haemost* 2014;**111**:1167–76. http://dx.doi.org/10.1160/TH14-03-0231

22. 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**:930–7. http://dx.doi.org/10.1161/CIRCULATIONAHA.112.126656

23. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. Geneva: WHO; 1968.

24. Allaby M. *Screening for Atrial Fibrillation in People Aged 65 and Over. A Report for the National Screening Committee*. Oxford: Solutions for Public Health; 2014.

25. Moran P, Teljeur C, Harrington P, Ryan M. *Health Technology Assessment (HTA) of a National Screening Programme for Atrial Fibrillation in Primary Care: Health Information and Quality Authority*; 2015. URL: www.hiqa.ie/sites/default/files/2017-02/HTA-of-screening-for-Atrial-Fibrillation.pdf

26. Friberg L, Engdahl J, Frykman V, Svennberg E, Levin LÅ, Rosenqvist M. Population screening of 75- and 76-year-old men and women for silent atrial fibrillation (STROKESTOP). *Europace* 2013;**15**:135–40. http://dx.doi.org/10.1093/europace/eus217

27. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation* 1998;**98**:946–52. https://doi.org/10.1161/01.CIR.98.10.946

28. Friberg L, Hammar N, Pettersson H, Rosenqvist M. Increased mortality in paroxysmal atrial fibrillation: report from the Stockholm Cohort-Study of Atrial Fibrillation (SCAF). *Eur Heart J* 2007;**28**:2346–53. https://doi.org/10.1093/eurheartj/ehm308

29. Hannon N, Sheehan O, Kelly L, Marnane M, Merwick A, Moore A, *et al.* Stroke associated with atrial fibrillation – incidence and early outcomes in the north Dublin population stroke study. *Cerebrovasc Dis* 2010;**29**:43–9. http://dx.doi.org/10.1159/000255973

30. Xie J, Wu EQ, Zheng ZJ, Croft JB, Greenlund KJ, Mensah GA, Labarthe DR. Impact of stroke on health-related quality of life in the noninstitutionalized population in the United States. *Stroke* 2006;**37**:2567–72. https://doi.org/10.1161/01.STR.0000240506.34616.10

31. Schnabel RB, Yin X, Gona P, Larson MG, Beiser AS, McManus DD, *et al.* 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet* 2015;**386**:154–62. http://dx.doi.org/10.1016/S0140-6736(14)61774-8

32. Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, *et al.* Guidelines for the management of atrial fibrillation: the task force for the management of atrial fibrillation of the European Society of Cardiology (ESC). *Eur Heart J* 2010;**31**:2369–429. https://doi.org/10.1093/eurheartj/ehq278

33. Furberg CD, Psaty BM, Manolio TA, Gardin JM, Smith VE, Rautaharju PM. Prevalence of atrial fibrillation in elderly subjects (the Cardiovascular Health Study). *Am J Cardiol* 1994;**74**:236–41. https://doi.org/10.1016/0002-9149(94)90363-8

34. Savelieva I, Camm AJ. Clinical relevance of silent atrial fibrillation: prevalence, prognosis, quality of life, and management. *J Interv Card Electrophysiol* 2000;**4**:369–82. https://doi.org/10.1023/A:1009823001707

35. Flaker GC, Belew K, Beckman K, Vidaillet H, Kron J, Safford R, *et al.* Asymptomatic atrial fibrillation: demographic features and prognostic information from the Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) study. *Am Heart J* 2005;**149**:657–63. https://doi.org/10.1016/j.ahj.2004.06.032

36. 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**:380. https://doi.org/10.1136/bmj.39227.551713.AE

37. Sterne J, Bodalia P, Bryden P, Davies P, Lopez-Lopez J, Okoli G, *et al.* Oral anticoagulants for primary prevention, treatment and secondary prevention of venous thromboembolic disease, and for prevention of stroke in atrial fibrillation: systematic review, network meta-analysis and cost-effectiveness analysis. *Health Technol Assess* 2017;**21**(9).

38. Mant J, Hobbs FDR, Fletcher K, Roalfe A, Fitzmaurice D, Lip GYH, *et al.* Warfarin versus aspirin for stroke prevention in an elderly community population with atrial fibrillation (the Birmingham Atrial Fibrillation Treatment of the Aged Study, BAFTA): a randomised controlled trial. *Lancet* 2007;**370**:493–503. https://doi.org/10.1016/S0140-6736(07)61233-1

39. NICE Implementation Collaborative. *Supporting Local Implementation of NICE Guidance on Use of the Novel (Non-Vitamin K Antagonist) Oral Anticoagulants in Non-Valvular Atrial Fibrillation*; 2014. URL: www.nice.org.uk/guidance/cg180/resources/nic-consensus-statement-on-the-use-of-noacs-243733501

40. Svennberg E, Engdahl J, Al-Khalili F, Friberg L, Frykman V, Rosenqvist M. Mass screening for untreated atrial fibrillation: the STROKESTOP study. *Circulation* 2015;**131**:2176–84. http://dx.doi.org/10.1161/CIRCULATIONAHA.114.014343

41. Lowres N, Neubeck L, Freedman SB. Can screening for atrial fibrillation be implemented at scale? *Europace* 2016;**18**:1449–51. https://doi.org/10.1093/europace/euw030

42. Christie B. People over 65 should be screened for atrial fibrillation, say stroke specialists. *BMJ* 2012;**344**:e1644. http://dx.doi.org/10.1136/bmj.e1644

43. Stott DJ, Dewar RI, Garratt W, Griffith KE, Harding NJ, James MA, *et al.* RCPE UK Consensus Conference on 'Approaching the comprehensive management of atrial fibrillation: evolution or revolution?'. *J R Coll Physicians Edinb* 2012;**42**(Suppl. 18):3–4. http://dx.doi.org/10.4997/JRCPE.2012.S01

44. Taggar JS, Coleman T, Lewis S, Jones M. Screening for atrial fibrillation – a cross-sectional survey of healthcare professionals in primary care. *PLOS ONE* 2016;**11**:e0152086. http://dx.doi.org/10.1371/journal.pone.0152086

45. National Institute for Health and Care Excellence. *Patient Decision Aid. Atrila Fibrillation: Medicines to Help Reduce Your Risk of a Stroke – What are the Options?* URL: www.nice.org.uk/guidance/cg180/resources/patient-decision-aid-243734797 (accessed 1 December 2016).

46. Man-Son-Hing M, Gage BF, Montgomery AA, Howitt A, Thomson R, Devereaux PJ, *et al.* Preference-based antithrombotic therapy in atrial fibrillation: implications for clinical decision making. *Med Decis Making* 2005;**25**:548–59. https://doi.org/10.1177/0272989X05280558

47. Thomson RG, Eccles MP, Steen IN, Greenaway J, Stobbart L, Murtagh MJ, May CR. A patient decision aid to support shared decision-making on anti-thrombotic treatment of patients with atrial fibrillation: randomised controlled trial. *Qual Saf Health Care* 2007;**16**:216–23. https://doi.org/10.1136/qshc.2006.018481

48. NICE. *Atrial fibrillation and heart valve disease: self-monitoring coagulation status using point-of-care coagulometers (ythe Coagu Check XS system and the INRatio2 PT/INR monitor)*; 2014. URL: www.nice.org.uk/guidance/dg14

49. Moran PS, Flattery MJ, Teljeur C, Ryan M, Smith SM. Effectiveness of systematic screening for the detection of atrial fibrillation. *Cochrane Database Syst Rev* 2013;**4**:CD009586. https://doi.org/10.1002/14651858.cd009586.pub2

50. Willis BH. Spectrum bias – why clinicians need to be cautious when applying diagnostic test studies. *Fam Pract* 2008;**25**:390–6. http://dx.doi.org/10.1093/fampra/cmn051

51. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529–36. http://dx.doi.org/10.7326/0003-4819-155-8-201110180-00009

52. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;**58**:882–93. https://doi.org/10.1016/j.jclinepi.2005.01.016

53. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;**20**:2865–84. https://doi.org/10.1002/sim.942

54. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: evolution, critique and future directions. *Stat Med* 2009;**28**:3049–67. https://doi.org/10.1002/sim.3680

55. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Statist Soc B* 2002;**64**:583–639. https://doi.org/10.1111/1467-9868.00353

56. Welton NJ. *Screening strategies for atrial fibrillation: a systematic review and cost-effectiveness analysis*. 2014. PROSPERO 2014:CRD42014013739. URL: www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014013739 (accessed 1 December 2016).

57. Feinberg WM, Blackshear JL, Laupacis A, Kronmal R, Hart RG. Prevalence, age distribution, and gender of patients with atrial fibrillation. Analysis and implications. *Arch Intern Med* 1995;**155**:469–73. https://doi.org/10.1001/archinte.1995.00430050045005

58. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928. http://dx.doi.org/10.1136/bmj.d5928

59. Effective Practice and Organisation of Care (EPOC). Summary assessments of the risk of bias. In *EPOC Resources for Review Authors*. Oslo: Norwegian Knowledge Centre for the Health Services; 2013.

60. Lee HO, Lee SJ, Jeong SH, Cho YH, Jang J, Kim DW, *et al.* A development and clinical evaluation of automated diagnostic algorithm for atrial fibrillation using 12-lead EKG. World Congress on Medical Physics and Biomedical Engineering 2006. In Kim SI, Suh TS, editors. *IFMBE Proceedings*, volume **14/1**; 2007. pp. 1195–8.

61. 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) [ISRCTN19633732]. *BMC Cardiovasc Disord* 2004;**4**:12. http://dx.doi.org/10.1186/1471-2261-4-12

62. Marazzi G, Iellamo F, Volterrani M, Lombardo M, Pelliccia F, Righi D, *et al.* Erratum to: Comparison of Microlife BP A200 Plus and Omron M6 blood pressure monitors to detect atrial fibrillation in hypertensive patients. *Adv Ther* 2014;**31**:1317. https://doi.org/10.1007/s12325-014-0172-2

63. 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. *Adv Ther* 2012;**29**:64–70. http://dx.doi.org/10.1007/s12325-011-0087-0

64. 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. *J Telemed Telecare* 2012;**18**:104–8. http://dx.doi.org/10.1258/jtt.2011.110613

65. Gregg RE, Zhou SH, Lindauer JM, Feild DQ, Helfenbein ED. Where do derived precordial leads fail? *J Electrocardiol* 2008;**41**:546–52. http://dx.doi.org/10.1016/j.jelectrocard.2008.07.018

66. 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* 2009;**11**:1362–8. http://dx.doi.org/10.1093/europace/eup262

67. Kearley K, Selwood M, Van den Bruel A, Thompson M, Mant D, Hobbs FR, *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**:e004565. http://dx.doi.org/10.1136/bmjopen-2013-004565

68. 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. *Med Eng Phys* 2012;**34**:1441–7. http://dx.doi.org/10.1016/j.medengphy.2012.02.002

69. Lewis M, Parker D, Weston C, Bowes M. Screening for atrial fibrillation: sensitivity and specificity of a new methodology. *Br J Gen Pract* 2011;**61**:38–9. http://dx.doi.org/10.3399/bjgp11X548956

70. Somerville S, Somerville J, Croft P, Lewis M. Atrial fibrillation: a comparison of methods to identify cases in general practice. *Br J Gen Pract* 2000;**50**:727–9.

71. Stergiou GS, Karpettas N, Protogerou A, Nasothimiou EG, Kyriakidis M. Diagnostic accuracy of a home blood pressure monitor to detect atrial fibrillation. *J Hum Hypertens* 2009;**23**:654–8. http://dx.doi.org/10.1038/jhh.2009.5

72. Vaes B, Stalpaert S, Tavernier K, Thaels B, Lapeire D, Mullens W, Degryse J. The diagnostic accuracy of the MyDiagnostick to detect atrial fibrillation in primary care. *BMC Fam Pract* 2014;**15**:113. http://dx.doi.org/10.1186/1471-2296-15-113

73. Slocum J, Sahakian A, Swiryn S. Diagnosis of atrial fibrillation from surface electrocardiograms based on computer-detected atrial activity. *J Electrocardiol* 1992;**25**:1–8. https://doi.org/10.1016/0022-0736(92)90123-H

74. Lau JK, Lowres N, Neubeck L, Brieger DB, Sy RW, Galloway CD, *et al.* iPhone ECG application for community screening to detect silent atrial fibrillation: a novel technology to prevent stroke. *Int J Cardiol* 2013;**165**:193–4. http://dx.doi.org/10.1016/j.ijcard.2013.01.220

75. Harris K, Edwards D, Mant J. How can we best detect atrial fibrillation? *J R Coll Physicians Edinb* 2012;**42**(Suppl. 18):5–22. http://dx.doi.org/10.4997/JRCPE.2012.S02

76. National Institute for Health and Care Excellence. *WatchBP Home A for Opportunistically Detecting Atrial Fibrillation during Diagnosis and Monitoring of Hypertension*. Medical technologies guidance (MTG13). London: NICE; 2013.

77. 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. *Eur J Prev Cardiol* 2016;**23**:1330–8. http://dx.doi.org/10.1177/2047487315611347

78. Taggar JS, Coleman T, Lewis S, Heneghan C, Jones M. Accuracy of methods for diagnosing atrial fibrillation using 12-lead ECG: a systematic review and meta-analysis. *Int J Cardiol* 2015;**184**:175–83. http://dx.doi.org/10.1016/j.ijcard.2015.02.014

79. 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**:e000362. http://dx.doi.org/10.1136/openhrt-2015-000362

80. Verberk WJ, Omboni S, Kollias A, Stergiou GS. Screening for atrial fibrillation with automated blood pressure measurement: research evidence and practice recommendations. *Int J Cardiol* 2016;**203**:465–73. http://dx.doi.org/10.1016/j.ijcard.2015.10.182

81. Pérula-de-Torres LA, Martinez-Adell MÁ, González-Blanco V, Baena-Diez JM, Martin-Rioboó E, Parras-Rejano JM, *et al.* Opportunistic detection of atrial fibrillation in subjects aged 65 years or older in primare care: a randomised clinical trial of efficacy. DOFA-AP study protocol. *BMC Fam Pract* 2012;**13**:106. https://doi.org/10.1186/1471-2296-13-106

82. Morgan S, Mant D. Randomised trial of two approaches to screening for atrial fibrillation in UK general practice. *Br J Gen Pract* 2002;**52**:373–4, 377–80.

83. Benito L, Coll-Vinent B, Gómez E, Martí D, Mitjavila J, Torres F, *et al.* EARLY: a pilot study on early diagnosis of atrial fibrillation in a primary healthcare centre. *Europace* 2015;**17**:1688–93. http://dx.doi.org/10.1093/europace/euv146

84. Fitzmaurice DA, Hobbs FD, Jowett S, Mant J, Murray ET, Holder R, *et al.* Screening versus routine practice in detection of atrial fibrillation in patients aged 65 or over: cluster randomised controlled trial. *BMJ* 2007;**335**:383. https://doi.org/10.1136/bmj.39280.660567.55

85. ClinicalTrials.gov. *Improving DEtection of Atrial fibriLlation in primary care with the MyDiagnostick (IDEAL-MD)*. 4 September 2014. URL: https://clinicaltrials.gov/ct2/show/NCT02270151 (accessed 16 January 2017).

86. Uittenbogaart SB, Verbiest-van Gurp N, Erkens PM, Lucassen WA, Knottnerus JA, Winkens B, *et al.* Detecting and Diagnosing Atrial Fibrillation (D2AF): study protocol for a cluster randomised controlled trial. *Trials* 2015;**16**:478. https://doi.org/10.1186/s13063-015-1006-5

87. Tucker J, Consumer Price Inflation time series dataset (MM23); 2017. URL: www.ons.gov.uk/economy/inflationandpriceindices/timeseries/chaw/mm23

88. EuroQol Group. *EQ-5D*. URL: www.euroqol.org

89. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State–Trait Anxiety Inventory (STAI). *Br J Clin Psychol* 1992;**31**:301–6. http://dx.doi.org/10.1111/j.2044-8260.1992.tb00997.x

90. Lowres N, Neubeck L, Redfern J, Freedman SB. Screening to identify unknown atrial fibrillation. A systematic review. *Thromb Haemost* 2013;**110**:213–22. http://dx.doi.org/10.1160/TH13-02-0165

91. Briggs AH, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press; 2006.

92. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013*. URL: www.nice.org.uk/process/pmg9/chapter/foreword (accessed 1 December 2016).

93. HM Treasury. *The Green Book: Appraisal and Evaluation in Central Government.* London: TSO; 2003.

94. Lavenson GS, Pantera RL, Garza RM, Neff T, Rothwell SD, Cisneros J. Development and implementation of a rapid, accurate, and cost-effective protocol for national stroke prevention screening. *Am J Surg* 2004;**188**:638–43. https://doi.org/10.1016/j.amjsurg.2004.08.055

95. McClennen S, Zimetbaum PJ, Ho KK, Goldberger AL. Holter monitoring: are two days better than one? *Am J Cardiol* 2000;**86**:562–4, A9.

96. Maeda K, Shimbo T, Fukui T. Cost-effectiveness of a community-based screening programme for chronic atrial fibrillation in Japan. *J Med Screen* 2004;**11**:97–102. http://dx.doi.org/10.1258/096914104774061092

97. Levin LÅ, Husberg M, Sobocinski PD, Kull VF, Friberg L, Rosenqvist M, Davidson T. A cost-effectiveness analysis of screening for silent atrial fibrillation after ischaemic stroke. *Europace* 2015;**17**:207–14. http://dx.doi.org/10.1093/europace/euu213

98. Centre for Reviews and Dissemination. *ISSG Search Filters Resource. Filters to Identify Economic Evaluations*. URL: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/filters-to-find-i (accessed 11 June 2016).

99. Aronsson M, Levin L. Finding the optimal screening program for unknown atrial fibrillation using simulation models. *Value Health* 2015;**18**:A354. http://dx.doi.org/10.1016/j.jval.2015.09.658

100. Aronsson M, Svennberg E, Rosenqvist M, Engdahl J, Al-Khalili F, Friberg L, *et al.* Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ECG recording. *Europace* 2015;**17**:1023–9. http://dx.doi.org/10.1093/europace/euv083

101. Diamantopoulos A, Sawyer L, Sungher D, Lip G, Witte K, Reynolds MR, *et al.* Use of an insertable cardiac monitor to detect atrial fibrillation is cost effective in patients with cryptogenic stroke. *Stroke* 2015;**46**:A14.

102. Limone BL, Baker WL, Mearns ES, White CM, Kluger J, Coleman CI. Common flaws exist in published cost-effectiveness models of pharmacologic stroke prevention in atrial fibrillation. *J Clin Epidemiol* 2014;**67**:1093–102. http://dx.doi.org/10.1016/j.jclinepi.2014.05.013

103. Lorenzoni G, Folino F, Soriani N, Iliceto S, Gregori D. Cost-effectiveness of early detection of atrial fibrillation via remote control of implanted devices. *J Eval Clin Pract* 2014;**20**:570–7. http://dx.doi.org/10.1111/jep.12132

104. Lowres N, Neubeck L, Redfern J, McLachlan A, Krass I, Bennett A, *et al.* Screening Education And Recognition in Community pHarmacies of Atrial Fibrillation to prevent stroke (SEARCH-AF stroke prevention study). *Heart Lung Circ* 2013;**22**:S223. https://doi.org/10.1016/j.hlc.2013.05.532

105. Moran P, Teljeur C, Harrington P, Smith S, Normand C, Ryan M. Opportunistic screening for atrial fibrillation in primary care – a clinical and cost-effectiveness analysis. *Value Health* 2015;**18**:A391. http://dx.doi.org/10.1016/j.jval.2015.09.868

106. 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. *Appl Health Econ Health Policy* 2014;**12**:255–65. https://doi.org/10.1007/s40258-014-0096-7

107. Becker C. Cost-of-illness studies of atrial fibrillation: methodological considerations. *Expert Rev Pharmacoecon Outcomes Res* 2014;**14**:661–84. http://dx.doi.org/10.1586/14737167.2014.940904

108. Fak AS, Küçükoğlu MS, Fak NA, Demir M, Ağir AA, Demirtaş M, *et al.* Expert panel on cost analysis of atrial fibrillation. *Anadolu Kardiyol Derg* 2013;**13**:26–38. http://dx.doi.org/10.5152/akd.2013.004

109. Turakhia MP, Shafrin J, Bognar K, Goldman DP, Mendys PM, Abdulsattar Y, *et al.* Economic burden of undiagnosed nonvalvular atrial fibrillation in the United States. *Am J Cardiol* 2015;**116**:733–9. http://dx.doi.org/10.1016/j.amjcard.2015.05.045

110. Kassianos G, Arden C, Hogan S, Baldock L, Fuat A. The non-anticoagulation costs of atrial fibrillation management: findings from an observational study in NHS primary care. *Drugs Context* 2014;**3**:212254. http://dx.doi.org/10.7573/dic.212254

111. Mayer F, Stahrenberg R, Gröschel K, Mostardt S, Biermann J, Edelmann F, *et al.* Cost-effectiveness of 7-day-Holter monitoring alone or in combination with transthoracic echocardiography in patients with cerebral ischemia. *Clin Res Cardiol* 2013;**102**:875–84. http://dx.doi.org/10.1007/s00392-013-0601-2

112. Krummen DE, Patel M, Nguyen H, Ho G, Kazi DS, Clopton P, *et al.* Accurate ECG diagnosis of atrial tachyarrhythmias using quantitative analysis: a prospective diagnostic and cost-effectiveness study. *J Cardiovasc Electrophysiol* 2010;**21**:1251–9. http://dx.doi.org/10.1111/j.1540-8167.2010.01809.x

113. Suter D, Saczynski J, Floyd K, Browning C, Rosenthal L, McManus D. Quality of life in symptomatic patients with atrial fibrillation: preliminary data from the InRhythm study. *Crit Path Cardiol* 2014;**13**:128–9.

114. Zhang L, Gallagher R, Neubeck L. Health-related quality of life in atrial fibrillation patients over 65 years: a review. *Eur J Prev Cardiol* 2015;**22**:987–1002. http://dx.doi.org/10.1177/2047487314538855

115. Grönefeld GC, Lilienthal J, Kuck K-H, Hohnloser S, Pharmacological Intervention in Atrial Fibrillation (PIAF) study investigators. Impact of rate versus rhythm control on quality of life in patients with persistent atrial fibrillation. Results from a prospective randomized study. *Eur Heart J* 2003;**24**:1430–6. https://doi.org/10.1016/S0195-668X(03)00261-6

116. Lord J, Willis S, Eatock J, Tappenden P, Trapero-Bertran M, Miners A, *et al.* Economic modelling of diagnostic and treatment pathways in National Institute for Health and Care Excellence clinical guidelines: the Modelling Algorithm Pathways in Guidelines (MAPGuide) project. *Health Technol Assess* 2013;**17**(58). http://dx.doi.org/10.3310/hta17580

117. 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. *Qual Prim Care* 2013;**21**:131–40.

118. Gordon S, Hickman M, Pentney V. Screening for asymptomatic atrial fibrillation at seasonal influenza vaccination. *Primary Care Cardiovasc J* 2012;**5**:161–4.

119. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, *et al.* Consolidated Health Economic Evaluation Reporting Standards (CHEERS) – explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health* 2013;**16**:231–50. http://dx.doi.org/10.1016/j.jval.2013.02.002

120. Trinity College Dublin. *The Irish Longitudinal Study on Ageing (TILDA)*. 2017. URL: http://tilda.tcd.ie/

121. National Institute for Health and Care Excellence. *Atrial Fibrillation: The Management of Atrial Fibrillation (CG36)*. 2006. URL: www.nice.org.uk/guidance/cg36

122. Department of Health. *NHS Reference Costs 2014–15*. London: Department of Health; 2014.

123. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: PSSRU, University of Kent; 2015.

124. Xiong Q, Proietti M, Senoo K, Lip GY. Asymptomatic versus symptomatic atrial fibrillation: a systematic review of age/gender differences and cardiovascular outcomes. *Int J Cardiol* 2015;**191**:172–7. http://dx.doi.org/10.1016/j.ijcard.2015.05.011

125. Schnabel RB, Wilde S, Wild PS, Munzel T, Blankenberg S. Atrial fibrillation: its prevalence and risk factor profile in the German general population. *Dtsch Arztebl Int* 2012;**109**:293–9. http://dx.doi.org/10.3238/arztebl.2012.0293

126. Frewen J, Finucane C, Cronin H, Rice C, Kearney PM, Harbison J, Kenny RA. Factors that influence awareness and treatment of atrial fibrillation in older adults. *QJM* 2013;**106**:415–24. http://dx.doi.org/10.1093/qjmed/hct060

127. Krijthe BP, Kunst A, Benjamin EJ, Lip GY, Franco OH, Hofman A, *et al.* Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *Eur Heart J* 2013;**34**:2746–51. http://dx.doi.org/10.1093/eurheartj/eht280

128. Gómez-Doblas JJ, Muñiz J, Martin JJ, Rodríguez-Roca G, Lobos JM, Awamleh P, *et al.* Prevalence of atrial fibrillation in Spain. OFRECE study results. *Rev Esp Cardiol* 2014;**67**:259–69. http://dx.doi.org/10.1016/j.rec.2013.07.014

129. Baena-Díez JM, Grau M, Forés R, Fernández-Bergés D, Elosua R, Sorribes M, *et al.* Prevalence of atrial fibrillation and its associated factors in Spain: an analysis of 6 population-based studies. DARIOS Study. *Rev Clin Esp* 2014;**214**:505–12. http://dx.doi.org/10.1016/j.rce.2014.06.006

130. Office for National Statistics. *Statistical Bulletin: Annual Mid-Year Population Estimates: 2014*. 2015. URL: www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/2015-06-25 (accessed 16 January 2017).

131. Davis RC, Hobbs FD, Kenkre JE, Roalfe AK, Iles R, Lip GY, Davies MK. Prevalence of atrial fibrillation in the general population and in high-risk groups: the ECHOES study. *Europace* 2012;**14**:1553–9. http://dx.doi.org/10.1093/europace/eus087

132. Bury G, Swan D, Cullen W, Keane D, Tobin H, Egan M, *et al.* Screening for atrial fibrillation in general practice: a national, cross-sectional study of an innovative technology. *Int J Cardiol* 2015;**178**:247–52. http://dx.doi.org/10.1016/j.ijcard.2014.10.037

133. Kesselheim AS, Cresswell K, Phansalkar S, Bates DW, Sheikh A. Clinical decision support systems could be modified to reduce 'alert fatigue' while still minimizing the risk of litigation. *Health Aff* 2011;**30**:2310–17. http://dx.doi.org/10.1377/hlthaff.2010.1111

134. Claes N, Van Laethem C, Goethals M, Goethals P, Mairesse G, Schwagten B, *et al.* Prevalence of atrial fibrillation in adults participating in a large-scale voluntary screening programme in Belgium. *Acta Cardiol* 2012;**67**:273–8. http://dx.doi.org/10.2143/AC.67.3.2160714

135. 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* 2016;**18**:1514–20. https://doi.org/10.1093/europace/euv426

136. Deif B, Lowres N, Freedman SB. Screening for atrial fibrillation above age 65 detects an asymptomatic subset at high risk of stroke. *Int J Cardiol* 2013;**164**:371–2. http://dx.doi.org/10.1016/j.ijcard.2012.08.002

137. Panisello-Tafalla A, Clua-Espuny JL, Gil-Guillen VF, González-Henares A, Queralt-Tomas ML, López-Pablo C, *et al.* Results from the Registry of Atrial Fibrillation (AFABE): gap between undiagnosed and registered atrial fibrillation in adults – ineffectiveness of oral anticoagulation treatment with VKA. *Biomed Res Int* 2015;**2015**:134756. http://dx.doi.org/10.1155/2015/134756

138. Sanmartín M, Fraguela Fraga F, Martín-Santos A, Moix Blázquez P, García-Ruiz A, Vázquez-Caamaño M, Vilar M, Grupo RIA (Registros en Insuficiencia Cardiaca y Aterosclerosis). A campaign for information and diagnosis of atrial fibrillation: 'pulse week'. *Rev Esp Cardiol* 2013;**66**:34–8. http://dx.doi.org/10.1016/j.recesp.2012.05.012

139. Smyth B, Marsden P, Corcoran R, Walsh R, Brennan C, McSharry K, *et al.* Opportunistic screening for atrial fibrillation in a rural area. *QJM* 2016;**109**:539–43. http://dx.doi.org/10.1093/qjmed/hcw011

140. Rienstra M, Vermond RA, Crijns HJ, Tijssen JG, Van Gelder IC, RACE Investigators. Asymptomatic persistent atrial fibrillation and outcome: results of the RACE study. *Heart Rhythm* 2014;**11**:939–45. http://dx.doi.org/10.1016/j.hrthm.2014.03.016

141. Frykman V, Frick M, Jensen-Urstad M, Ostergren J, Rosenqvist M. Asymptomatic versus symptomatic persistent atrial fibrillation: clinical and noninvasive characteristics. *J Intern Med* 2001;**250**:390–7. https://doi.org/10.1046/j.1365-2796.2001.00893.x

142. Senoo K, Suzuki S, Otsuka T, Sagara K, Matsuno S, Kano H, *et al.* Progression to the persistent form in asymptomatic paroxysmal atrial fibrillation. *Circ J* 2014;**78**:1121–6. https://doi.org/10.1253/circj.CJ-13-1272

143. Ruigómez A, Johansson S, Wallander MA, García Rodríguez LA. Predictors and prognosis of paroxysmal atrial fibrillation in general practice in the UK. *BMC Cardiovasc Disord* 2005;**5**:20. https://doi.org/10.1186/1471-2261-5-20

144. Holmqvist F, Kim S, Steinberg BA, Reiffel JA, Mahaffey KW, Gersh BJ, *et al.* Heart rate is associated with progression of atrial fibrillation, independent of rhythm. *Heart* 2015;**101**:894–9. http://dx.doi.org/10.1136/heartjnl-2014-307043

145. Kato T, Yamashita T, Sagara K, Iinuma H, Fu LT. Progressive nature of paroxysmal atrial fibrillation. Observations from a 14-year follow-up study. *Circ J* 2004;**68**:568–72. https://doi.org/10.1253/circj.68.568

146. Al-Khatib SM, Wilkinson WE, Sanders LL, McCarthy EA, Pritchett EL. Observations on the transition from intermittent to permanent atrial fibrillation. *Am Heart J* 2000;**140**:142–5. https://doi.org/10.1067/mhj.2000.107547

147. Steinberg BA, Hellkamp AS, Lokhnygina Y, Patel MR, Breithardt G, Hankey GJ, *et al.* Higher risk of death and stroke in patients with persistent vs. paroxysmal atrial fibrillation: results from the ROCKET-AF trial. *Eur Heart J* 2015;**36**:288–96. http://dx.doi.org/10.1093/europace/euv390

148. Friberg L, Hammar N, Rosenqvist M. Stroke in paroxysmal atrial fibrillation: report from the Stockholm Cohort of Atrial Fibrillation. *Eur Heart J* 2010;**31**:967–75. http://dx.doi.org/10.1093/eurheartj/ehn599

149. Hohnloser SH, Pajitnev D, Pogue J, Healey JS, Pfeffer MA, Yusuf S, Connolly SJ, ACTIVE W Investigators. Incidence of stroke in paroxysmal versus sustained atrial fibrillation in patients taking oral anticoagulation or combined antiplatelet therapy: an ACTIVE W substudy. *J Am Coll Cardiol* 2007;**50**:2156–61. https://doi.org/10.1016/j.jacc.2007.07.076

150. Martinez C, Katholing A, Freedman SB. Adverse prognosis of incidentally detected ambulatory atrial fibrillation. A cohort study. *Thromb Haemost* 2014;**112**:276–86. http://dx.doi.org/10.1160/TH4-04-0383

151. Avgil Tsadok M, Jackevicius CA, Rahme E, Humphries KH, Behlouli H, Pilote L. Sex differences in stroke risk among older patients with recently diagnosed atrial fibrillation. *JAMA* 2012;**307**:1952–8. http://dx.doi.org/10.1001/jama.2012.3490

152. Kerr CR, Humphries KH, Talajic M, Klein GJ, Connolly SJ, Green M, *et al.* Progression to chronic atrial fibrillation after the initial diagnosis of paroxysmal atrial fibrillation: results from the Canadian Registry of Atrial Fibrillation. *Am Heart J* 2005;**149**:489–96. https://doi.org/10.1016/j.ahj.2004.09.053

153. Fitzmaurice DA, McCahon D, Baker J, Murray ET, Jowett S, Sandhar H, *et al.* Is screening for AF worthwhile? Stroke risk in a screened population from the SAFE study. *Fam Pract* 2014;**31**:298–302. http://dx.doi.org/10.1093/fampra/cmu011

154. Cowan C, Healicon R, Robson I, Long WR, Barrett J, Fay M, *et al.* The use of anticoagulants in the management of atrial fibrillation among general practices in England. *Heart* 2013;**99**:1166–72. http://dx.doi.org/10.1136/heartjnl-2012-303472

155. Apenteng PN, Murray ET, Hobbs FDR, Kakkar AK, Fitzmaurice DA. Patterns of antithrombotic therapy in relation to type of atrial fibrillation: insights from the UK cohort of the global GARFIELD registry. *Europace* 2014;**16**:iii14. https://doi.org/10.1093/europace/euu239.3

156. Faria R, Walker S, Whyte S, Dixon S, Palmer S, Sculpher M. How to invest in getting cost-effective technologies into practice? A framework for value of implementation analysis applied to novel oral anticoagulants. *Med Decis Making* 2016;**37**:148–61. https://doi.org/10.1177/0272989X16645577

157. Office for National Statistics. *ONS Consumer Price Inflation Index for Medical Services (DKC3) for 2016* (released 17 May 2016). URL: www.ons.gov.uk/economy/inflationandpriceindices/timeseries/dkc3 (accessed 20 May 2016).

158. Kansal AR, Sorensen SV, Gani R, Robinson P, Pan F, Plumb JM, Cowie MR. Cost-effectiveness of dabigatran etexilate for the prevention of stroke and systemic embolism in UK patients with atrial fibrillation. *Heart* 2012;**98**:573–8. http://dx.doi.org/10.1136/heartjnl-2011-300646

159. Bayer Plc. *Single Technology Appraisal (STA) of Rivaroxaban (Xarelto)*; 2011. URL: www.nice.org.uk/guidance/ta256/documents/atrial-fibrillation-stroke-prevention-rivaroxaban-buyer4

160. Wells G, Coyle D, Cameron C, Steiner S, Coyle K, Kelly S, *et al. Safety, Effectiveness, and Cost-effectiveness of New Oral Anticoagulants Compared with Warfarin in Preventing Stroke and Other Cardiovascular Events In Patients with Atrial Fibrillation*. Canada: Canadian Agency for Drugs and Technologies in Health (CADTH). Therapeutic Review 2012. URL: http://cadth.ca/media/pdf/NOAC_Therapeutic_Review_final_report.pdf (accessed 16 January 2017).

161. Aguilar MI, Hart R, Pearce LA. Oral anticoagulants versus antiplatelet therapy for preventing stroke in patients with non-valvular atrial fibrillation and no history of stroke or transient ischemic attacks. *Cochrane Database Syst Rev* 2007;**3**:CD006186. https://doi.org/10.1002/14651858.cd006186.pub2

162. Friberg L, Rosenqvist M, Lip GY. Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study. *Eur Heart J* 2012;**33**:1500–10. http://dx.doi.org/10.1093/eurheartj/ehr488

163. Andersen KK, Olsen TS. Reduced poststroke mortality in patients with stroke and atrial fibrillation treated with anticoagulants: results from a Danish quality-control registry of 22,179 patients with ischemic stroke. *Stroke* 2007;**38**:259–63. https://doi.org/10.1161/01.STR.0000254622.52483.03

164. Joint Formulary Committee. *British National Formulary*. 69th ed. London: BMJ Group and Pharmaceutical Press; 2015.

165. Department of Health. *NHS Reference Costs 2013–14*. London: Department of Health; 2014.

166. Luengo-Fernandez R, Yiin GS, Gray AM, Rothwell PM. Population-based study of acute- and long-term care costs after stroke in patients with AF. *Int J Stroke* 2013;**8**:308–14. http://dx.doi.org/10.1111/j.1747-4949.2012.00812.x

167. Scottish Centre for Telehealth and Telecare. *Atrial Fibrillation*. URL: www.sctt.scot.nhs.uk/programmes/health/cardiac/atrial-fibrillation/ (accessed 16 June 2016).

168. Health and Social Care Information Centre. *Numbers of Patients Registered at a GP Practice – April 2016*. URL: www.hscic.gov.uk/article/2021/Website-Search?productid=16843 (accessed 15 June 2016).

169. Wolff A, Long S, McComb J, Richley D, Mercer P. The gap between training and provision: a primary-care based ECG survey in North-East England. *Br J Cardiol* 2012;**19**:38–40. https://doi.org/10.5837/bjc.2012.008

170. Numed Healthcare. *12 Lead ECG Machines*. URL: www.numed.co.uk/products/category/12-lead-ecg (accessed 15 June 2016).

171. Bonanzamarket. URL: www.bonanzamarket.co.uk/listings/OLED-Pulse-Finger-Fingertip-Oximeter-Blood-SpO2-PR-Heart-Rate-Monitor/352640758?goog_pla=1&variation_id=125007605&gpid=18283950120&keyword=&goog_pla=1&pos=1o10&ad_type=pla&gclid=CKjuoLSmqM0CFQkq0wodnFMPBA (accessed 15 June 2016).

172. Omron. *Omron M6 Comfort Upper Arm Blood Pressure Monitor*. URL: http://omronwebshop.co.uk/blood-pressure-monitors/upper-arm-blood-pressure-monitors/new-m6-comfort.html (accessed 15 June 2016).

173. Wright J, Bibby J, Eastham J, Harrison S, McGeorge M, Patterson C, *et al.* Multifaceted implementation of stroke prevention guidelines in primary care: cluster-randomised evaluation of clinical and cost effectiveness. *Qual Saf Health Care* 2007;**16**:51–9. https://doi.org/10.1136/qshc.2006.019778

174. Caley M, Chohan P, Hooper J, Wright N. The impact of NHS Health Checks on the prevalence of disease in general practices: a controlled study. *Br J Gen Pract* 2014;**64**:e516–21. http://dx.doi.org/10.3399/bjgp14X681013

175. Abe R, Nishida T. [The criteria for the prediction of paroxysmal atrial fibrillation by time domain analysis of the P wave-triggered signal-averaged electrocardiogram.] *Nippon Rinsho* 1995;**53**:496–502.

176. Albilali A, Gulamhusein S, Butcher K, Khera S, Kivi P, Abassi M, *et al.* Occult atrial fibrillation detected by auto-triggered loop recorder screening in a community-based population (a result of the first 40 participants). *Int J Stroke* 2014;**9**:194.

177. Al-Fahoum AS, Qasaimeh AM. A practical reconstructed phase space approach for ECG arrhythmias classification. *J Med Eng Technol* 2013;**37**:401–8. http://dx.doi.org/10.3109/03091902.2013.819946

178. Alis C, del Rosario C, Buenaobra B, Mar Blanca C. Lifelink: 3G-based mobile telemedicine system. *Telemed J E Health* 2009;**15**:241–7. http://dx.doi.org/10.1089/tmj.2008.0098

179. Amir O, Barak-Shinar D, Wolff R, Amos Y, Paz H, Smart F, *et al.* Atrial fibrillation detection using a photoplethysmograph waveform. *Sleep* 2011;**34**:A234.

180. Andrikopoulos GK, Dilaveris PE, Richter DJ, Gialafos EJ, Synetos AG, Gialafos JE. Increased variance of P wave duration on the electrocardiogram distinguishes patients with idiopathic paroxysmal atrial fibrillation. *Pacing Clin Electrophysiol* 2000;**23**:1127–32. https://doi.org/10.1111/j.1540-8159.2000.tb00913.x

181. Anh D, Krishnan S, Bogun F. Accuracy of electrocardiogram interpretation by cardiologists in the setting of incorrect computer analysis. *J Electrocardiol* 2006;**39**:343–5. https://doi.org/10.1016/j.jelectrocard.2006.02.002

182. Aras D, Maden O, Ozdemir O, Aras S, Topaloglu S, Yetkin E, *et al.* Simple electrocardiographic markers for the prediction of paroxysmal atrial fibrillation in hyperthyroidism. *Int J Cardiol* 2005;**99**:59–64. https://doi.org/10.1016/j.ijcard.2003.11.040

183. Asl BM, Setarehdan SK, Mohebbi M. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artif Intell Med* 2008;**44**:51–64. http://dx.doi.org/10.1016/j.artmed.2008.04.007

184. Aytemir K, Aksoyek S, Ozer N, Aslamaci S, Oto A. Atrial fibrillation after coronary artery bypass surgery: P wave signal averaged ECG, clinical and angiographic variables in risk assessment. *Int J Cardiol* 1999;**69**:49–56. https://doi.org/10.1016/S0167-5273(99)00005-4

185. Aytemir K, Ozer N, Atalar E, Sade E, Aksöyek S, Ovünç K, *et al.* P wave dispersion on 12-lead electrocardiography in patients with paroxysmal atrial fibrillation. *Pacing Clin Electrophysiol* 2000;**23**:1109–12. https://doi.org/10.1111/j.1540-8159.2000.tb00910.x

186. Babaeizadeh S, Gregg RE, Helfenbein ED, Lindauer JM, Zhou SH. Improvements in atrial fibrillation detection for real-time monitoring. *J Electrocardiol* 2009;**42**:522–6. http://dx.doi.org/10.1016/j.jelectrocard.2009.06.006

187. Bae MH, Lee JH, Lee HS, Yang DH, Park HS, Cho Y, *et al.* Common errors in automatic computer ECG interpretation of atrial fibrillation. *Europace* 2011;**13**:i67.

188. Bae MH, Lee JH, Yang DH, Park HS, Cho Y, Chae SC, Jun JE. Erroneous computer electrocardiogram interpretation of atrial fibrillation and its clinical consequences. *Clin Cardiol* 2012;**35**:348–53. http://dx.doi.org/10.1002/clc.22000

189. Bansil S, Karim H. Detection of atrial fibrillation in patients with acute stroke. *J Stroke Cerebrovasc Dis* 2004;**13**:12–15. https://doi.org/10.1016/j.jstrokecerebrovasdis.2004.01.004

190. Barrett P, Komatireddy R, Topol EJ. Comparison of a novel 14-day adhesive continuous ECG monitoring patch to traditional 24-hour Holter monitoring in symptomatic arrhythmia patients. *Irish J Med Sci* 2013;**182**:S374.

191. 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. *Am J Med* 2014;**127**:95.e11–95.e17. https://doi.org/10.1016/j.amjmed.2013.10.003

192. Barthélémy JC, Féasson-Gérard S, Garnier P, Gaspoz JM, Da Costa A, Michel D, Roche F. Automatic cardiac event recorders reveal paroxysmal atrial fibrillation after unexplained strokes or transient ischemic attacks. *Ann Noninvasive Electrocardiol* 2003;**8**:194–9. https://doi.org/10.1046/j.1542-474X.2003.08305.x

193. Benavente L, Calleja S, Vega J, Larrosa D, Rico M, Para M, *et al.* Yield of Holter and echocardiography in the screening of TIA. *Cerebrovasc Dis* 2013;**35**:490.

194. Bernard P, Chaitman BR, Scholl JM, Chabot M. [Value of the automated analysis of the electrocardiogram by the Telemed program (V version).] *Arch Mal Coeur Vaiss* 1981;**74**:1155–62.

195. Boudaoud S, Rix H, Blanc JJ, Cornily JC, Meste O. Integrated shape averaging of the P-wave applied to AF risk detection. *Comput Cardiol* 2003;**30**:125–8. https://doi.org/10.1109/cic.2003.1291106

196. Bourdillon PJ, Kilpatrick D. Clinicians, the Mount Sinai program and the Veterans' Administration program evaluated against clinico-pathological data derived independently of the electrocardiogram. *Eur J Cardiol* 1978;**8**:395–412.

197. Boyle KO, Morra D, Dorian P, McCrorie A, Haddad P, Taylor L, *et al.* Atrial fibrillation screening using a handheld ECG device: results from the Heart and Stroke Foundation (HSF) 'be pulse aware' campaign. *Stroke* 2013;**44**:e184.

198. Budeus M, Hennersdorf M, Perings C, Strauer BE. Detection of atrial late potentials with P wave signal averaged electrocardiogram among patients with paroxysmal atrial fibrillation. *Z Kardiol* 2003;**92**:362–9. https://doi.org/10.1007/s00392-003-0921-8

199. Budeus M, Hennersdorf M, Wieneke H, Sack S, Erbel R, Perings C. P wave signal averaged ECG and chemoreflexsensitivity in paroxysmal atrial fibrillation. *Int J Cardiol* 2005;**100**:317–24. https://doi.org/10.1016/j.ijcard.2004.12.001

200. 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* 2012;**14**:191–6. http://dx.doi.org/10.1093/europace/eur304

201. Chao TF, Sung SH, Wang KL, Tsao HM, Lin YJ, Chang SL, *et al.* Atrial electromechanical interval can identify patients with paroxysmal atrial fibrillation and is associated with CHADS$_2$ score and peak velocity of left atrial appendage. *J Cardiovasc Electrophysiol* 2011;**22**:1325–30. http://dx.doi.org/10.1111/j.1540-8167.2011.02115.x

202. Chee K, Coles D, McGill D. Is P wave duration on 12-leads ECG useful to select ischemic stroke patients for 24h ambulatory Holter monitoring? *Heart Lung Circ* 2010;**19**:S220. https://doi.org/10.1016/j.hlc.2010.06.533

203. Choi S, Jiang Z. Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique. *Comput Biol Med* 2010;**40**:8–20. http://dx.doi.org/10.1016/j.compbiomed.2009.10.003

204. Clavier L, Boucher JM, Lepage R, Blanc JJ, Cornily JC. Automatic P-wave analysis of patients prone to atrial fibrillation. *Med Biol Eng Comput* 2002;**40**:63–71. https://doi.org/10.1007/BF02347697

205. Cole-Haskayne AL, Pushie A, Gairdner S, Coutts S, Demchuk A. The use of cardiac event monitoring in ambulatory care for the detection of atrial fibrillation. *Stroke* 2014;**45**:E281.

206. Coutts SB, Choi PM. Seven days of non-invasive cardiac monitoring early postischaemic stroke or TIA increases atrial fibrillation detection rate compared with current guideline-based practice. *Evid Based Med* 2014;**19**:152. http://dx.doi.org/10.1136/eb-2013-101607

207. Cubanski D, Cyganski D, Antman EM, Feldman CL. A neural network system for detection of atrial fibrillation in ambulatory electrocardiograms. *J Cardiovasc Electrophysiol* 1994;**5**:602–8. https://doi.org/10.1111/j.1540-8167.1994.tb01301.x

208. Czaplik M, Eilebrecht B, Ntouba A, Walter M, Schauerte P, Leonhardt S, Rossaint R. Clinical proof of practicability for an ECG device without any conductive contact. *Biomed Tech* 2010;**55**:291–300. http://dx.doi.org/10.1515/BMT.2010.035

209. Dash S, Chon KH, Lu S, Raeder EA. Automatic real time detection of atrial fibrillation. *Ann Biomed Eng* 2009;**37**:1701–9. http://dx.doi.org/10.1007/s10439-009-9740-z

210. Dash S, Raeder E, Merchant S, Chon KH. *A Statistical Approach for Accurate Detection of Atrial Fibrillation and Flutter*. 36th Annual Computers in Cardiology Conference; 2009. pp. 137–40.

211. Dash S, Raeder E, Merchant S, Chon KH. A statistical approach for accurate detection of atrial fibrillation and flutter. *J Electrocard* 2011;**44**:e7. https://doi.org/10.1016/j.jelectrocard.2010.12.024

212. Davidenko JM, Snyder LS. Causes of errors in the electrocardiographic diagnosis of atrial fibrillation by physicians. *J Electrocardiol* 2007;**40**:450–6. https://doi.org/10.1016/j.jelectrocard.2007.01.003

213. Davy JM, Laurendeau C, Donio V, Spiess N, Briand Y. The French screening campaign of atrial fibrillation in general practice: assessment of predictive criteria for atrial fibrillation. *Eur Heart J* 2014;**35**:868.

214. Deelawar S, Manegold JC, Kalyani M, Tschishow WN, Ridjab D, Buddecke J, *et al.* Diagnostic yield of extended-time external loop recording for the detection of paroxysmal atrial fibrillation in patients with cryptogenic stroke. *Europace* 2013;**15**:ii220.

215. Dilaveris PE, Gialafos EJ, Sideris SK, Theopistou AM, Andrikopoulos GK, Kyriakidis M, *et al.* Simple electrocardiographic markers for the prediction of paroxysmal idiopathic atrial fibrillation. *Am Heart J* 1998;**135**:733–8. https://doi.org/10.1016/S0002-8703(98)70030-4

216. Dilaveris PE, Gialafos EJ, Chrissos D, Andrikopoulos GK, Richter DJ, Lazaki E, Gialafos JE. Detection of hypertensive patients at risk for paroxysmal atrial fibrillation during sinus rhythm by computer-assisted P wave analysis. *J Hypertens* 1999;**17**:1463–70. https://doi.org/10.1097/00004872-199917100-00015

217. Dirschedl P, Lenz S, Löllgen H, Fahrenkrog U. [Validity of telephone ECG multichannel transmission.] *Z Kardiol* 1996;**85**:677–83.

218. Dobbels L, Thijs V. Prospective comparison of continuous cardiac monitoring and Holter on a stroke unit. *Eur J Neurol* 2014;**21**(Suppl. 1):116.

219. Dogan U, Dogan EA, Tekinalp M, Tokgoz OS, Aribas A, Akilli H, *et al.* P-wave dispersion for predicting paroxysmal atrial fibrillation in acute ischemic stroke. *Int J Med Sci* 2012;**9**:108–14. https://doi.org/10.7150/ijms.9.108

220. Doliwa PS, Frykman V, Rosenqvist M. Short-term ECG for out of hospital detection of silent atrial fibrillation episodes. *Scand Cardiovasc J* 2009;**43**:163–8. http://dx.doi.org/10.1080/14017430802593435

221. Doliwa Sobocinski P, Anggardh Rooth E, Frykman Kull V, von Arbin M, Wallen H, Rosenqvist M. Improved screening for silent atrial fibrillation after ischaemic stroke. *Europace* 2012;**14**:1112–16. https://doi.org/10.1093/europace/eur431

222. Dotan U, Dotan EA, Tekinalp M, Tokgoz OS, Aribas A, Akilli H, *et al.* Evaluation of potential predictive value of P-wave dispersion for predicting paroxysmal atrial fibrillation in patients with acute ischemic stroke. *Turk Kardiyoloji Dernegi Ars* 2011;**39**:196.

223. Douen AG, Pageau N, Medic S. Serial electrocardiographic assessments significantly improve detection of atrial fibrillation 2.6-fold in patients with acute stroke. *Stroke* 2008;**39**:480–2. http://dx.doi.org/10.1161/STROKEAHA.107.492595

224. Du X, Rao N, Qian M, Liu D, Li J, Feng W, *et al.* A novel method for real-time atrial fibrillation detection in electrocardiograms using multiple parameters. *Ann Noninvasive Electrocardiol* 2014;**19**:217–25. http://dx.doi.org/10.1111/anec.12111

225. Dukes JW, Whooley MA, Marcus GM. Cohort under-ascertainment of prevalent atrial fibrillation using the baseline 12-lead ECG: implications for atrial fibrillation 'prediction'. *Circulation* 2014;**130**:A18479.

226. Duverney D, Gaspoz JM, Pichot V, Roche F, Brion R, Antoniadis A, Barthélémy JC. High accuracy of automatic detection of atrial fibrillation using wavelet transform of heart rate intervals. *Pacing Clin Electrophysiol* 2002;**25**:457–62. https://doi.org/10.1046/j.1460-9592.2002.00457.x

227. Elgendi M, Mahalingam S, Jonkman M, De Boer F. *A Robust QRS Complex Detection Elgorithm Using Dynamic Thresholds*. International Symposium on Computer Science and its Applications 2008. URL: www.researchgate.net/publication/232644395_A_Robust_QRS_Complex_Detection_Algorithm_using_Dynamic_Thresholds (accessed 16 January 2017).

228. Elijovich L, Josephson SA, Fung GL, Smith W. 30-day event monitors in detecting atrial fibrillation in cryptogenic stroke. *Stroke* 2008;**39**:584.

229. Engdahl J, Andersson L, Mirskaya M, Rosenqvist M. Stepwise screening for atrial fibrillation in a 75-year old population: implications for stroke prevention. *Cerebrovasc Dis* 2012;**33**(Suppl. 2):830–1.

230. Feruglio GA, Feraco E, Maisano G. [The Caceres-USPHS (version HP-3) computer program. Evaluation of automatic analysis of 40,000 ECG.] *G Ital Cardiol* 1975;**5**:262–71.

231. Filos D, Chouvarda I, Dakos G, Vassilikos V, Maglaveras N. Beat to beat wavelet variability in atrial fibrillation. *Conf Proc IEEE Eng Med Biol Soc* 2011;**2011**:953–6. http://dx.doi.org/10.1109/IEMBS.2011.6090215

232. Finucane C, Frewen J, Nolan H, Murphy B, Kenny RA. Towards a smart app for increasing awareness and detection of atrial fibrillation in the community. Translating findings from the Irish longitudinal study on ageing into practice. *Ir J Med Sci* 2013;**182**:S214.

233. Foley PWX, Yung L, Barnes E. Comparison of wireless ambulatory cardiac monitoring with Holter monitoring. *Europace* 2011;**13**(Suppl. 3).

234. Franczuk PK, Grodzicki T. [Sensitivity and specificity of high resolution electrocardiography in the identification of patients with paroxysmal atrial fibrillation.] *Folia Cardiologica* 2004;**11**:547–53.

235. Fukunami M, Yamada T, Ohmori M, Kumagai K, Umemoto K, Sakai A, *et al.* Detection of patients at risk for paroxysmal atrial fibrillation during sinus rhythm by P wave-triggered signal-averaged electrocardiogram. *Circulation* 1991;**83**:162–9. https://doi.org/10.1161/01.CIR.83.1.162

236. Fukunami M. Signal-averaged ECG for diagnosing paroxysmal atrial fibrillation. *Primary Cardiol* 1992;**18**:30–5.

237. Gaillard N, Vilotijevic B, Deltour S, Hornych A, Leger A, Crozier S, *et al.* Detection of paroxysmal atrial fibrillation in patients after cerebral ischemic events with transtelephonic ECG monitoring. *Neurology* 2008;**70**:A156-A.

238. Ghrooda EM, Dobrowolski P, Basir G, Yaseen I, Khan N, Ahmad A, *et al.* Paroxysmal atrial fibrillation is common in patients with defined etiology for stroke: prolonged monitoring of cardiac rhythm for detection of atrial fibrillation after a cerebral ischemic event (PEAACE) study. *Stroke* 2014;**45**(Suppl. 1):A25.

239. Gialafos JE, Dilaveris PE, Gialafos EJ, Andrikopoulos GK, Richter DJ, Triposkiadis F, Kyriakidid MK. P wave dispersion: a valuable electrocardiographic marker for the prediction of paroxysmal lone atrial fibrillation. *Ann Noninvasive Electrocardiol* 1999;**4**:39–45. https://doi.org/10.1111/j.1542-474X.1999.tb00363.x

240. Gomis M, Millan M, De La Ossa NP, Dorado L, Guerrero C, Ricciardi AC, *et al.* Delayed ambulatory Holter for the detection of embolic arrhythmias after normal ECG continuous monitoring during 72 hours in acute ischemic stroke. *Cerebrovasc Dis* 2009;**27**:225.

241. Goricke B, Stahrenberg R, Weber-Kruger M, Seegers J, Edelmann F, Lahno R, *et al.* Atrial fibrillation: enhanced detection rates by prolonged Holter monitoring in stroke patients. *Cerebrovasc Dis* 2011;**31**:7.

242. Gradl S, Kugler P, Lohmuller C, Eskofier B. Real-time ECG monitoring and arrhythmia detection using Android-based mobile devices. *Conf Proc IEEE Eng Med Biol Soc* 2012;**2012**:2452–5. http://dx.doi.org/10.1109/EMBC.2012.6346460

243. Graja S, Boucher JM. *SVM classification of patients prone to atrial fibrillation*. 2005 IEEE International Workshop on Intelligent Signal Processing 2005. URL: www.researchgate.net/publication/4187535_SVM_Classification_of_patients_prone_to_atrial_fibrillation (accessed 16 January 2017).

244. Grond M, Rosin L, Kirchhof P. Detection of undiagnosed silent AF in patients with acute stroke. *Cerebrovasc Dis* 2011;**31**:33.

245. 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**:3357–64. http://dx.doi.org/10.1161/STROKEAHA.113.001884

246. Guidera SA, Steinberg JS. The signal-averaged P wave duration: a rapid and noninvasive marker of risk of atrial fibrillation. *J Am Coll Cardiol* 1993;**21**:1645–51. https://doi.org/10.1016/0735-1097(93)90381-A

247. Güler I, Ubeyli ED. ECG beat classifier designed by combined neural network model. *Pattern Recogn* 2005;**38**:199–208. https://doi.org/10.1016/S0031-3203(04)00276-6

248. Gumbinger C, Hacke W, Ringleb P. Continuous monitoring vs. Holter ECG for detection of atrial fibrillation in patients with stroke. *Cerebrovasc Dis* 2010;**29**:167–8.

249. Gumbinger C, Krumsdorf U, Veltkamp R, Hacke W, Ringleb P. Continuous monitoring versus Holter ECG for detection of atrial fibrillation in patients with stroke. *Eur J Neurol* 2012;**19**:253–7. http://dx.doi.org/10.1111/j.1468-1331.2011.03519.x

250. Gunalp M, Atalar E, Coskun F, Yilmaz A, Aksoyek S, Aksu NM, Sivri B. Holter monitoring for 24 hours in patients with thromboembolic stroke and sinus rhythm diagnosed in the emergency department. *Adv Ther* 2006;**23**:854–60. https://doi.org/10.1007/BF02850206

251. Haberman ZC, Jahn RT, Bose R, Tun H, Shinbane JS, Doshi RN, *et al.* Wireless smart phone equipped ECG enables large scale screening in diverse populations. *Heart Rhythm* 2014;**11**:S312.

252. Haberman ZC, Jahn RT, Bose R, Tun H, Shinbane JS, Doshi RN, *et al.* Wireless smartphone ECG enables large-scale screening in diverse populations. *J Cardiovasc Electrophysiol* 2015;**26**:520–6. http://dx.doi.org/10.1111/jce.12634

253. Haeberlin A, Niederhauser T, Marisa T, Goette J, Jacomet M, Tanner H, *et al.* First validation of esophageal long-term electrocardiography as an alternative technique for long-term heart rhythm monitoring. *Eur Heart J* 2012;**33**:702.

254. Haeberlin A, Roten L, Schilling M, Scarcia F, Niederhauser T, Vogel R, *et al.* Software-based detection of atrial fibrillation in long-term ECGs. *Heart Rhythm* 2014;**11**:933–8. http://dx.doi.org/10.1016/j.hrthm.2014.03.014

255. Hakacova N, Trägårdh-Johansson E, Wagner GS, Maynard C, Pahlm O. Computer-based rhythm diagnosis and its possible influence on nonexpert electrocardiogram readers. *J Electrocardiol* 2012;**45**:18–22. http://dx.doi.org/10.1016/j.jelectrocard.2011.05.007

256. Hallioglu O, Aytemir K, Celiker A. The significance of P wave duration and P wave dispersion for risk assessment of atrial tachyarrhythmias in patients with corrected tetralogy of Fallot. *Ann Noninvasive Electrocardiol* 2004;**9**:339–44. https://doi.org/10.1111/j.1542-474X.2004.94569.x

257. Larkin Harrington J, Woon Chong J, Li J, Esa N, Pidikiti R, Maitas O, *et al.* The detection and differentiation of arrhythmias using a smartphone: a clinical study of patients with atrial fibrillation, premature atrial and premature ventricular contractions. *J Am Coll Cardiol* 2013;**61**:E362. https://doi.org/10.1016/S0735-1097(13)60362-9

258. Helfenbein E, Gregg R, Lindauer J, Zhou S. An automated algorithm for the detection of atrial fibrillation in the presence of paced rhythms. *Comput Cardiol* 2010;**37**:113–16.

259. Hendrikx T, Rosenqvist M, Sandstrom H, Hornsten R. Comparing efficacy of Holter ECG and handheld ECG in detecting relevant arrhythmias – the Primarytm study. *Cerebrovasc Dis* 2012;**33**:838–9.

260. Hendrikx T, Rosenqvist M, Wester P, Sandström H, Hörnsten R. Intermittent short ECG recording is more effective than 24-hour Holter ECG in detection of arrhythmias. *BMC Cardiovasc Disord* 2014;**14**:41. http://dx.doi.org/10.1186/1471-2261-14-41

261. Hickey B, Heneghan C, de Chazal P. Non-episode-dependent assessment of paroxysmal atrial fibrillation through measurement of RR interval dynamics and atrial premature contractions. *Ann Biomed Eng* 2004;**32**:677–87. https://doi.org/10.1023/B:ABME.0000030233.39769.a4

262. 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**:2134–6. http://dx.doi.org/10.1161/STROKEAHA.114.005405

263. Hiraki T, Ikeda H, Ohga M, Hamada T, Kubara I, Yoshida T, *et al.* Frequency- and time-domain analysis of P wave in patients with paroxysmal atrial fibrillation. *Pacing Clin Electrophysiol* 1998;**21**:56–64. https://doi.org/10.1111/j.1540-8159.1998.tb01061.x

264. Hoffmann BA, Bakucz P, Rostock T, Steven D, Drewitz I, Sultan A, *et al.* Ultra-fast and accurate surface ECG detection of very short episodes of atrial fibrillation using wavelet transformation. *Heart Rhythm* 2010;**7**:S135.

265. Hong-Wei L, Ying S, Min L, Pi-Ding L, Zheng Z. A probability density function method for detecting atrial fibrillation using R–R intervals. *Med Eng Phys* 2009;**31**:116–23. http://dx.doi.org/10.1016/j.medengphy.2008.04.013

266. Horstmann S, Rizos T, Güntner J, Hug A, Jenetzky E, Krumsdorf U, Veltkamp R. Does the STAF score help detect paroxysmal atrial fibrillation in acute stroke patients? *Eur J Neurol* 2013;**20**:147–52. http://dx.doi.org/10.1111/j.1468-1331.2012.03816.x

267. Hoshino T, Nagao T, Shiga T, Maruyama K, Toi S, Mizuno S, *et al.* Prolonged QTc interval predicts poststroke paroxysmal atrial fibrillation. *Stroke* 2015;**46**:71–6. http://dx.doi.org/10.1161/STROKEAHA.114.006612

268. Howlett PJ, Morritt J, Jabr R, Mahmoudi M, Fry CH, Leatham EW. Targeted screening for paroxysmal atrial fibrillation: prolonged monitoring significantly increases diagnostic yield. *Europace* 2014;**16**:iii26. https://doi.org/10.1093/europace/euu243.4

269. Howlett P, Morritt J, Greswell L, Findlay N, Mahmoudi M, Waheed A, *et al.* Symptom frequency is a poor predictor of onset of paroxysmal atrial fibrillation in a population presenting with palpitations. *Europace* 2014;**16**:iii31. https://doi.org/10.1093/europace/euu244.7

270. Hsieh JC, Tzeng WC, Yang YC, Shieh SM, Ieee. Detecting ECG characteristic points by novel hybrid wavelet transforms: an evaluation of clinical SCP-ECG database. *Comput Cardiol* 2005;**32**:751–4. https://doi.org/10.1109/cic.2005.1588213

271. Inoue K, Shirayama T, Shiraishi H, Matoba Y, Imai H, Inoue D, Nakagawa M. Clinical significance of the atrial fibrillation threshold in patients with paroxysmal atrial fibrillation. *Pacing Clin Electrophysiol* 2001;**24**:796–805. https://doi.org/10.1046/j.1460-9592.2001.00796.x

272. Jang KJ, Balakrishnan G, Syed Z, Verma N. Scalable customization of atrial fibrillation detection in cardiac monitoring devices: increasing detection accuracy through personalized monitoring in large patient populations. *Conf Proc IEEE Eng Med Biol Soc* 2011;**2011**:2184–7. http://dx.doi.org/10.1109/IEMBS.2011.6090411

273. Jeon T, Kim B, Jeon M, Lee BG. Implementation of a portable device for real-time ECG signal analysis. *Biomed Eng Online* 2014;**13**:160. http://dx.doi.org/10.1186/1475-925X-13-160

274. Jiang K, Huang C, Ye SM, Chen H. High accuracy in automatic detection of atrial fibrillation for Holter monitoring. *J Zhejiang Univ Sci B* 2012;**13**:751–6. http://dx.doi.org/10.1631/jzus.B1200107

275. Joseph G, Zywietz C, Murray A, Swiryn S. Computerized ECG diagnosis on a reduced lead set (limb leads). *Comput Cardiol* 1999;**26**:687–90. https://doi.org/10.1109/cic.1999.826064

276. Kallmünzer 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**:994–9. http://dx.doi.org/10.1161/STROKEAHA.111.642199

277. Kallmünzer 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**:598–603. http://dx.doi.org/10.1212/WNL.0000000000000690

278. Kandel A, Ching MI, Covey TJ, Shucard DW. Timing of mobile cardiac outpatient telemetry may increase diagnostic yield of atrial fibrillation in select patients with cryptogenic strokes. *Stroke* 2012;**43**:A2437.

279. Kar A, Ellis A, Guyler P, O'Brien A. Intensive cardiac monitoring after transient ischaemic attack identifies a significant number of previously unknown paroxysmal atrial fibrillation. *Int J Stroke* 2009;**4**:1.

280. Karapinar H, Kaya Z, Pala S, Karavelioglu Y, Dasli T, Sirma D, *et al.* Predictive value of atrial electromechanical coupling time for paroxysmal atrial fibrillation. *Eur Heart J* 2009;**30**:561–2.

281. Kessler DK, Kessler KM. Is ambulatory electrocardiography beneficial to detect atrial-fibrillation in patients with stroke? *Clin Res* 1993;**41**:A811.

282. Kikillus N, Hammer G, Wieland S, Bolz A. Algorithm for identifying patients with paroxysmal atrial fibrillation without appearance on the ECG. *Conf Proc IEEE Eng Med Biol Soc* 2007;**2007**:275–8. http://dx.doi.org/10.1109/IEMBS.2007.4352277

283. Kikillus N, Hammer G, Bolz A. [Identification of patients with atrial fibrillation using HRV parameters.] *Biomed Tech* 2008;**53**:8–15. http://dx.doi.org/10.1515/BMT.2008.002

284. Kim D, Seo Y, Youn CH. Detection of atrial fibrillation episodes using multiple heart rate variability features in different time periods. *Conf Proc IEEE Eng Med Biol Soc* 2008;**2008**:5482–5. http://dx.doi.org/10.1109/IEMBS.2008.4650455

285. Kim SH, Oh YS, Kim BK, Ha YW, Park JW, Kim TS, *et al.* Diagnostic usefulness of prolonged and real-time rhythm monitoring using mobile phone. *Heart Rhythm* 2013;**10**:S22–3.

286. Kinlay S, Leitch JW, Neil A, Chapman BL, Hardy DB, Fletcher PJ. Cardiac event recorders yield more diagnoses and are more cost-effective than 48-hour Holter monitoring in patients with palpitations. A controlled clinical trial. *Ann Intern Med* 1996;**124**:16–20. https://doi.org/10.7326/0003-4819-124-1_Part_1-199601010-00003

287. Klein MD, Key-Brothers I, Feldman CL, Can the vectorcardiographically derived EASI ECG be a suitable surrogate for the standard ECG in selected circumstances. *Comput Cardiol* 1997;**24**:721–4. https://doi.org/10.1109/cic.1997.648152

288. Köse S, Aytemir K, Sade E, Can I, Ozer N, Amasyali B, *et al.* Detection of patients with hypertrophic cardiomyopathy at risk for paroxysmal atrial fibrillation during sinus rhythm by P-wave dispersion. *Clin Cardiol* 2003;**26**:431–4. https://doi.org/10.1002/clc.4960260910

289. Koskinen R, Lehto M, Väänänen H, Rantonen J, Voipio-Pulkki LM, Mäkijärvi M, *et al.* Measurement and reproducibility of magnetocardiographic filtered atrial signal in patients with paroxysmal lone atrial fibrillation and in healthy subjects. *J Electrocardiol* 2005;**38**:330–6. https://doi.org/10.1016/j.jelectrocard.2005.03.012

290. Kostka PS, Tkacz EJ. Feature extraction and selection algorithms in biomedical data classifiers based on time-frequency and principle component analysis. In *11th Mediterranean Conference on Medical and Biological Engineering and Computing*, 2007. pp. 70–3. https://doi.org/10.1007/978-3-540-73044-6_19

291. Kostka PS, Tkacz EJ. Feature extraction based on time-frequency and independent component analysis for improvement of separation ability in atrial fibrillation detector. *Conf Proc IEEE Eng Med Biol Soc* 2008;**2008**:2960–3. http://dx.doi.org/10.1109/IEMBS.2008.4649824

292. Kostka PS, Tkacz EJ. Feature extraction for improving the support vector machine biomedical data classifier performance. In *International Special Topic Conference on Information Technology and Applications in Biomedicine*, 2008. pp. 434–7. https://doi.org/10.1109/itab.2008.4570638

293. Kostka P, Tkacz E. *Support Vector Machine with Rule Extraction for Medical Signal Classification*. 19th International EURASIP Conference Biosignal, 30 June–1 July 2008, Brno, Czech Republic, abstract no. 45.

294. Kostka PS, Tkacz EJ. Feature selection based on time-frequency analysis in SVM classifier with rules extraction stage. In Dössel O, Schlegel WC, editors. *Proceedings of the World Congress 2009 for Medical Physics and Biomedical Engineering, vol. 25/IV: Image Processing, Biosignal Processing, Modelling and Simulation, Biomechanics*. pp. 2261–4. https://doi.org/10.1007/978-3-642-03882-2_601

295. Kostka PS, Tkacz EJ. Feature extraction in time-frequency signal analysis by means of matched wavelets as a feature generator. *Conf Proc IEEE Eng Med Biol Soc* 2011;**2011**:4996–9. http://dx.doi.org/10.1109/IEMBS.2011.6091238

296. Krasteva V, Jekova I. QRS template matching for recognition of ventricular ectopic beats. *Ann Biomed Eng* 2007;**35**:2065–76. http://dx.doi.org/10.1007/s10439-007-9368-9

297. Lagido RB, Lobo J, Leite S, Sousa C, Ferreira L, Silva-Cardoso J, *et al.* Using the smartphone camera to monitor heart rate and rhythm in heart failure patients. *Proceedings of the 2014 IEEE-EMBS International Conference on Biomedical and Health Informatics*. pp. 556–9. https://doi.org/10.1109/BHI.2014.6864425

298. Lalani GG, Schricker AA, Clopton P, Krummen DE, Narayan SM. Frequency analysis of atrial action potential alternans: a sensitive clinical index of individual propensity to atrial fibrillation. *Circ Arrhythm Electrophysiol* 2013;**6**:859–67. http://dx.doi.org/10.1161/CIRCEP.113.000204

299. Langley P, di Bernardo D, Allen J, Bowers E, Smith FE, Vecchietti S, *et al.* Can paroxysmal atrial fibrillation be predicted? *Comput Cardiol* 2001;**28**:121–4. https://doi.org/10.1109/cic.2001.977606

300. Lau J, Lowres N, Neubeck L, Brieger DB, Sy RW, Galloway C, *et al.* Validation of an iphone ECG application suitable for community screening for silent atrial fibrillation: a novel way to prevent stroke. *Circulation* 2012;**126**:A16810.

301. Lavallee PC, Benchimol H, Labreuche J, Meseguer E, Mazighi M, Guidoux C, *et al.* Feasibility and sensibility of daily transtelephonic EKG monitoring to detect paroxysmal atrial fibrillation in a selected population of stroke patients. *Cerebrovasc Dis* 2013;**35**:429.

302. Lazzaro MA, Krishnan K, Prabhakaran S. The presence of supraventricular tachycardia may be a marker of paroxysmal atrial fibrillation in acute ischemic stroke and TIA patients. *Cerebrovasc Dis* 2010;**29**:155.

303. Lazzaro MA, Krishnan K, Prabhakaran S. Detection of atrial fibrillation with concurrent Holter monitoring and continuous cardiac telemetry following ischemic stroke and transient ischemic attack. *J Stroke Cerebrovasc Dis* 2012;**21**:89–93. http://dx.doi.org/10.1016/j.jstrokecerebrovasdis.2010.05.006

304. Lee J, McManus D, Chon K. Atrial fibrillation detection using time-varying coherence function and Shannon entropy. *Conf Proc IEEE Eng Med Biol Soc* 2011;**2011**:4685–8. http://dx.doi.org/10.1109/IEMBS.2011.6091160

305. Lee J, Nam Y, McManus DD, Chon KH. Time-varying coherence function for atrial fibrillation detection. *IEEE Trans Biomed Eng* 2013;**60**:2783–93. http://dx.doi.org/10.1109/TBME.2013.2264721

306. Lee J, Reyes BA, McManus DD, Maitas O, Mathias O, Chon KH. Atrial fibrillation detection using an iPhone 4S. *IEEE Trans Biomed Eng* 2013;**60**:203–6. http://dx.doi.org/10.1109/TBME.2012.2208112

307. Lelakowska-Pieła M, Pudło J, Engel A, Lelakowski J. [Analysis of P wave duration and dispersion in paroxysmal atrial fibrillation.] *Pol Merkur Lekarski* 2013;**35**:259–62.

308. Lepage R, Boucher JM, Blanc JJ, Cornilly JC, Ieee, Ieee, *et al.* ECG segmentation and P-wave feature extraction: application to patients prone to atrial fibrillation. *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vols 1–4: Building New Bridges at the Frontiers of Engineering and Medicine*; 2001. pp. 298–301.

309. Lever NA, Larsen PD, Dawes M, Wong A, Harding SA. Are our medical graduates in New Zealand safe and accurate in ECG interpretation? *N Z Med J* 2009;**122**:9–15.

310. Lewalter T, Remerie S, Yang A, Schrickel JW, Stoeckigt F, Andrie R, *et al.* Detection of atrial fibrillation: sensitivity of various ECG monitoring approaches. *Eur Heart J* 2007;**28**:537.

311. Lim HS, Lip GY. Asymptomatic atrial fibrillation on device interrogation. *J Cardiovasc Electrophysiol* 2008;**19**:891–3. http://dx.doi.org/10.1111/j.1540-8167.2008.01194.x

312. Linker DT, Dziubinski MJ, Badelt SW, Swerdlow CD. Accurate, unsupervised screening of Holter recordings for diagnosing atrial fibrillation. *Heart Rhythm* 2013;**10**:S167.

313. Liu CS, Tseng WK, Lee JK, Hsiao TC, Lin CW. The differential method of phase space matrix for AF/VF discrimination application. *Med Eng Phys* 2010;**32**:444–53. http://dx.doi.org/10.1016/j.medengphy.2010.04.001

314. Locati ET, Vecchi AM, Cattafi G, Sachero A, Lunati M. Role of new external loop recorders (ELR) with auto-trigger functions and high-memory capacity in a step-wise approach for the clinical diagnosis of infrequent arrhythmic events. *Europace* 2011;**13**:P675.

315. Logan B, Healey J. Robust detection of atrial fibrillation for a long term telemonitoring system. *Comput Cardiol* 2005;**32**:619–22. https://doi.org/10.1109/cic.2005.1588177

316. Lu H. [Detecting atrial fibrillation and normal sinus rhythm by R–R intervals.] *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi* 2010;**27**:183–7.

317. Madsen C, Henriksen IO, Tveskov C. Value of 48 hour telemetry of all acute stroke patients in detecting paroxysmal atrial fibrillation. *Cerebrovasc Dis* 2009;**27**:44.

318. Mahagne MH, Suissa L, Bertora D, Lachaud S. Score for targeting of atrial fibrillation (STAF): a new approach for detection of atrial fibrillation in ischemic stroke patients? *Cerebrovasc Dis* 2009;**27**:65.

319. Mahagne MH, Suissa L, Lachaud S. Continuous bedside ECG monitoring: an interesting method to increase atrial fibrillation diagnosing rate in stroke patients. *Cerebrovasc Dis* 2011;**31**:295.

320. Maier C, Bauch M, Dickhaus H. Screening and prediction of paroxysmal atrial fibrillation by analysis of heart rate variability parameters. *Comput Cardiol* 2001;**28**:129–32. https://doi.org/10.1109/cic.2001.977608

321. Mairesse GH, Scavee C, Claes N, Goethals P, Vijgen J, Blankoff I, *et al.* Similar heart rates between atrial fibrillation and sinus rhythm in asymptomatic populations justify large scale screening to improve the detection of atrial fibrillation. *Europace* 2013;**15**:ii224.

322. Maitas O, Lee J, Robotis D, Bourell P, Chon KH, McManus DD. Detection of atrial fibrillation using an iphone 4S. *Circulation* 2012;**126**:A9298.

323. Maitas O, Lee J, Robotis D, Chon KH, McManus DD. Detection of atrial fibrillation using a smartphone camera. *Circ Res* 2012;**111** A55.

324. Mäkijärvi M, Nenonen J, Toivonen L, Montonen J, Katila T, Siltanen P. Magnetocardiography: supraventricular arrhythmias and preexcitation syndromes. *Eur Heart J* 1993;**14**(Suppl. E):46–52. https://doi.org/10.1093/eurheartj/14.suppl_E.46

325. Makowska E, Szymot J, Soszyriska M, Kulakowski P. Evaluation of patients with palpitations: cardiac event recorder versus 48-hour Holter monitoring. *Ann Noninvasive Electrocardiol* 2000;**5**:315–21. https://doi.org/10.1111/j.1542-474X.2000.tb00068.x

326. Manina G, Agnelli G, Becattini C, Zingarini G, Paciaroni M. 96 hours ECG monitoring for patients with ischemic cryptogenic stroke or transient ischaemic attack. *Intern Emerg Med* 2014;**9**:65–7. http://dx.doi.org/10.1007/s11739-012-0755-3

327. Martínez A, Alcaraz R, Rieta JJ. Study on the P-wave feature time course as early predictors of paroxysmal atrial fibrillation. *Physiol Meas* 2012;**33**:1959–74. http://dx.doi.org/10.1088/0967-3334/33/12/1959

328. Martinez-Sanchez P, Cruz-Herranz A, Correas E, Fuentes B, Cazorla-Garcia R, Martinez-Martinez M, *et al.* Diagnosis of paroxysmal atrial fibrillation combining cardiac and Holter monitoring in patients with acute brain ischemia. *Eur J Neurol* 2011;**18**:115.

329. Martinez-Sanchez P, Callero EC, Herranz AC, Gimeno BF, Montes AM, De Los Angeles Mangas Guijarro M, *et al.* Detection of paroxysmal atrial fibrillation in patient with acute brain ischemia combining cardiac and Holter monitoring: prevalence and predictors. *Stroke* 2012;**43**:A3648.

330. Martis RJ, Acharya UR, Adeli H, Prasad H, Tan JH, Chua KC, *et al.* Computer aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation. *Biomed Signal Process Control* 2014;**13**:295–305. https://doi.org/10.1016/j.bspc.2014.04.001

331. Maslowsky F, Sarzi Braga S, Pedretti RFE, Tramarin R. Atrial depolarization dispersion on Holter monitoring is an useful marker of atrial fibrillation after cardiac surgery. *Eur J Prev Cardiol* 2012;**19**:S133.

332. McCarthy L. ACP Journal Club: ambulatory ECG monitoring for 30 d increased AF detection more than 24 h of ECG monitoring after cryptogenic stroke. *Ann Intern Med* 2014;**161**:JC2. http://dx.doi.org/10.7326/0003-4819-161-10-201411180-02002

333. Miller DJ, Khan M, Schultz L, Simpson JR, Russman A, Panayiotis M. Outpatient cardiac telemetry detects a high rate of atrial fibrillation among patients with cryptogenic cerebral ischemia. *Stroke* 2012;**43**:A150.

334. Miller DJ. Randomised controlled trial: prolonged cardiac monitoring after cryptogenic stroke superior to 24 h ECG in detection of occult paroxysmal atrial fibrillation. *Evid Based Med* 2014;**19**:235. https://doi.org/10.1136/ebmed-2014-110074

335. Mohebbi M, Ghassemian H. Detection of atrial fibrillation episodes using SVM. *Conf Proc IEEE Eng Med Biol Soc* 2008;**2008**:177–80. http://dx.doi.org/10.1109/IEMBS.2008.4649119

336. Mohebbi M, Ghassemian H. Prediction of paroxysmal atrial fibrillation based on non-linear analysis and spectrum and bispectrum features of the heart rate variability signal. *Comput Methods Programs Biomed* 2012;**105**:40–9. http://dx.doi.org/10.1016/j.cmpb.2010.07.011

337. Montereggi A, Marconi P, Olivotto I, Castelli G, Dolara A, Luisi ML, Gheri RG. Signal-averaged P-wave duration and risk of paroxysmal atrial fibrillation in hyperthyroidism. *Am J Cardiol* 1996;**77**:266–9. https://doi.org/10.1016/S0002-9149(97)89391-5

338. Moreira JO, Moffa PJ, Uchida AH, Tobias NM, Grupi CJ, Luna Filho B, Tarasoutchi F. The signal-averaged electrocardiogram of atrial activation in patients with or without paroxysmal atrial fibrillation. *Arq Bras Cardiol* 2006;**87**:564–9.

339. Mueller A, Och W, Schweizer J. Automatic detection of atrial fibrillation and telemedical ECG transmission via a loop recorder. *Eur Heart J* 2006;**27**:139.

340. Müller 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. *J Telemed Telecare* 2009;**15**:391–6. http://dx.doi.org/10.1258/jtt.2009.090402

341. Murgatroyd FD, Xie BY, Copie X, Blankoff I, Camm AJ, Malik M. Identification of atrial fibrillation episodes in ambulatory electrocardiographic recordings: validation of a method for obtaining labeled R–R interval files. *Pacing Clin Electrophysiol* 1995;**18**:1315–20. https://doi.org/10.1111/j.1540-8159.1995.tb06972.x

342. Opolski G, Kraska T, Piatkowska-Janko E, Słomka K, Górecki A, Stanisławska J, *et al.* [P-wave EKG averaging technique – a new method of selecting patients with paroxysmal atrial fibrillation.] *Pol Tyg Lek* 1995;**50**:39–41.

343. Ozdemir O, Soylu M, Demir AD, Topaloglu S, Alyan O, Turhan H, *et al.* P-wave durations as a predictor for atrial fibrillation development in patients with hypertrophic cardiomyopathy. *Int J Cardiol* 2004;**94**:163–6. http://dx.doi.org/10.1016/j.ijcard.2003.01.001

344. Ozer N, Aytemir K, Atalar E, Sade E, Aksöyek S, Ovünç K, *et al.* P wave dispersion in hypertensive patients with paroxysmal atrial fibrillation. *Pacing Clin Electrophysiol* 2000;**23**:1859–62. https://doi.org/10.1111/j.1540-8159.2000.tb07038.x

345. Park J, Lee S, Jeon M. Atrial fibrillation detection by heart rate variability in Poincare plot. *Biomed Eng Online* 2009;**8**:38. http://dx.doi.org/10.1186/1475-925X-8-38

346. Pastor-Pérez FJ, Manzano-Fernández S, Goya-Esteban R, Pascual-Figal DA, Barquero-Pérez 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. *Am J Cardiol* 2010;**106**:677–81. http://dx.doi.org/10.1016/j.amjcard.2010.04.027

347. Peer C, Schreier G, Kastner P, Marko W, Messmer J, Rotman B, *et al.* easyG – electrocardiogram analysis system Graz. 4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine. *Conference Proceedings: New Solutions for New Challenges*; 2003 pp. 338–41.

348. Pellise A, Ustrell X, Milian M, Merce J, Vinas J, Mares R. Enhanced detection of occult atrial fibrillation by 24 hour Holter monitoring in selected patients with acute ischemic stroke. *Cerebrovasc Dis* 2011;**31**:296.

349. Poon K, Okin PM, Kligfield P. Diagnostic performance of a computer-based ECG rhythm algorithm. *J Electrocardiol* 2005;**38**:235–8. https://doi.org/10.1016/j.jelectrocard.2005.01.008

350. Portet F. P wave detector with PP rhythm tracking: evaluation in different arrhythmia contexts. *Physiol Meas* 2008;**29**:141–55. http://dx.doi.org/10.1088/0967-3334/29/1/010

351. Potpara TS, Lane DA. Diving to the foot of an iceberg: the SEARCH for undiagnosed atrial fibrillation. *Thromb Haemost* 2014;**112**:1–3. http://dx.doi.org/10.1160/TH14-05-0437

352. Pryor TA, Lindsay AE, England RW. Computer analysis of serial electrocardiograms. *Comput Biomed Res* 1972;**5**:709–14. https://doi.org/10.1016/0010-4809(72)90049-3

353. Pusalkar A, Venter M, Banerjee S, Kar A, Ames D. Inpatient continuous cardiac telemetry on a hyper-acute stroke unit increases the detection of paroxysmal atrial fibrillation after stroke and TIA. *Cerebrovasc Dis* 2012;**33**:839–40.

354. Quinto Villani G, Rosi A, Dieci G, Arruzzoli S, Gazzola U. [The high-resolution analysis of the P wave recorded via the esophagus: a new diagnostic approach in patients with paroxysmal atrial fibrillation.] *G Ital Cardiol* 1993;**23**:139–44.

355. Rabinstein AA. Prolonged cardiac monitoring for detection of paroxysmal atrial fibrillation after cerebral ischemia. *Stroke* 2014;**45**:1208–14. http://dx.doi.org/10.1161/STROKEAHA.113.003389

356. Reddy BR, Taha B, Swiryn S, Silberman R, Childers R. Prospective evaluation of a microprocessor-assisted cardiac rhythm algorithm: results from one clinical center. *J Electrocardiol* 1998;**30**:28–33. https://doi.org/10.1016/S0022-0736(98)80016-2

357. Reifart N, Weil HJ, Göhring S, Dietl J. [The reliability of a new 12-channel ECG with telephone transmission.] *Dtsch Med Wochenschr* 1997;**122**:1137–40. http://dx.doi.org/10.1055/s-2008-1047739

358. 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 Cardiol* 2012;**67**:525–31. http://dx.doi.org/10.2143/AC.67.5.2174126

359. Rhys GC, Mohammed F, Foster A. Screening for AF in patients 65 years and over attending annual flu vaccination clinics at Moorlands Medical Centre. *Int J Stroke* 2012;**7**:73.

360. Rincon F, Grassi PR, Khaled N, Atienza D, Sciuto D. Automated real-time atrial fibrillation detection on a wearable wireless sensor platform. *Conf Proc IEEE Eng Med Biol Soc* 2012;**2012**:2472–5. http://dx.doi.org/10.1109/EMBC.2012.6346465

361. Rizikou D, Margariti A, Labakis M, Vangelis S, Makrygiannis S, Labakis S, *et al.* Clinical suspicion of nursing personnel in the identification and prevention of atrial fibrillation in intensive care unit patients. *Intensive Care Med* 2013;**39**:S403.

362. Rizos T, Rasch C, Jenetzky E, Hametner C, Kathoefer S, Reinhardt R, *et al.* Detection of paroxysmal atrial fibrillation in acute stroke patients. *Cerebrovasc Dis* 2010;**30**:410–7. http://dx.doi.org/10.1159/000316885

363. Rizos T, Güntner J, Jenetzky E, Marquardt L, Reichardt C, Becker R, *et al.* Continuous stroke unit electrocardiographic monitoring versus 24-hour Holter electrocardiography for detection of paroxysmal atrial fibrillation after stroke. *Stroke* 2012;**43**:2689–94. https://doi.org/10.1161/STROKEAHA.112.654954

364. 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 Clin Electrophysiol* 2002;**25**:1587–93. https://doi.org/10.1046/j.1460-9592.2002.01587.x

365. Ros E, Mota S, Toro FJ, Díaz AF, Fernández FJ. Automatic paroxysmal atrial fibrillation based on not fibrillating ECGs. *Methods Inf Med* 2004;**43**:94–8.

366. Roten L, Schilling M, Häberlin 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;**98**:645–9. http://dx.doi.org/10.1136/heartjnl-2011-301455

367. Samol A, Masin M, Gellner R, Otte B, Pavenstädt HJ, Ringelstein EB, *et al.* Prevalence of unknown atrial fibrillation in patients with risk factors. *Europace* 2013;**15**:657–62. http://dx.doi.org/10.1093/europace/eus366

368. Sanak D, Kral M, Hutyra M, Bartkova A, Fedorco M, Zapletalova J, *et al.* Heart and brain long-term ECG-Holter monitoring in detection of paroxysmal atrial fibrillation in young cryptogenic ischemic stroke patients. *Int J Stroke* 2014;**9**:201–2.

369. Sato H, Takaki H, Shimizu W, Kamakura S, Sugimachi M. Magnetocardiography can identify abnormal atrial electrical activities during sinus rhythm in patients with paroxysmal atrial fibrillation. *Eur Heart J* 2009;**30**:559.

370. Sawant AC, Te Riele AS, Sandhu O, Sandhu L, Srivatsa S. Smartphone based applications accurately predict rhythm disturbances and diagnose silent atrial fibrillation. *Heart Rhythm* 2014;**11**:S230. https://doi.org/10.1016/s0735-1097(14)60299-0

371. 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**:e89328. http://dx.doi.org/10.1371/journal.pone.0089328

372. Schaer BA, Cron TA, Osswald S. Value of routine Holter screening for the detection of paroxysmal atrial fibrillation in patients with cerebral ischemic attacks. *J Am Coll Cardiol* 2003;**41**:531A. https://doi.org/10.1016/S0735-1097(03)82848-6

373. Schreier G, Kastner P, Marko W. An automatic ECG processing algorithm to identify patients prone to paroxysmal atrial fibrillation. *Comput Cardiol* 2001;**28**:133–5. https://doi.org/10.1109/cic.2001.977609

374. 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 Clin Electrophysiol* 1999;**22**:1082–4. https://doi.org/10.1111/j.1540-8159.1999.tb00574.x

375. Sezgin N. Nonlinear analysis of electrocardiography signals for atrial fibrillation. *Sci World J* 2013;**2013**:509784. http://dx.doi.org/10.1155/2013/509784

376. Shafqat S, Kelly PJ, Furie KL. Holter monitoring in the diagnosis of stroke mechanism. *Intern Med J* 2004;**34**:305–9. http://dx.doi.org/10.1111/j.1444-0903.2004.00589.x

377. Sheldon SH, Baturova MA, Platonov PG, Brady PA, Lin G, Rabinstein A, *et al.* Use of left atrial volume index to rule out paroxysmal atrial fibrillation on ambulatory electrocardiographic monitoring in patients with ischemic stroke. *Heart Rhythm* 2014;**11**:S392.

378. Shiyovich A, Wolak A, Yacobovich L, Grosbard A, Katz A. Accuracy of diagnosing atrial flutter and atrial fibrillation from a surface electrocardiogram by hospital physicians: analysis of data from internal medicine departments. *Am J Med Sci* 2010;**340**:271–5. http://dx.doi.org/10.1097/MAJ.0b013e3181e73fcf

379. Singh D, Xiang C, Tat HTC, Nan LQS, Yeo WT, Poh KK, *et al.* Validation study of a real-time wireless, cell phone-based single lead electrocardiographic monitoring system in dysrrhythmia. *Circulation* 2010;**122**:e385.

380. Sobocinski-Doliwa P, Rooth E, Frykman V, Von Arbin M, Wallen H, Rosenqvist M. Improved screening for silent atrial fibrillation after ischemic stroke. *Scand Cardiovasc J* 2010;**44**:10.

381. Stafford PJ, Robinson D, Vincent R. Optimal analysis of the signal averaged P wave in patients with paroxysmal atrial fibrillation. *Br Heart J* 1995;**74**:413–18. https://doi.org/10.1136/hrt.74.4.413

382. Stahrenberg R, Weber-Krüger 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**:2884–8. http://dx.doi.org/10.1161/STROKEAHA.110.591958

383. 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**:327–8. https://doi.org/10.1136/bmj.317.7154.327

384. Sugai TK, Yoshizawa M, Abe M, Shimizu K, Inagaki M, Sugimachi M, Sunagawa K. Preliminary study on the detection of cardiac arrhythmias based on multiple simultaneous electrograms. *Conf Proc IEEE Eng Med Biol Soc* 2009;**2009**:2498–501. http://dx.doi.org/10.1109/IEMBS.2009.5335157

385. Suissa L, Mahagne MH, Lachaud S. Score for the targeting of atrial fibrillation: a new approach to diagnosing paroxysmal atrial fibrillation. *Cerebrovasc Dis* 2011;**31**:442–7. http://dx.doi.org/10.1159/000323852

386. Suissa L, Bresch S, Lachaud S, Mahagne MH. What pertinent parameter(s) to rule out paroxysmal atrial fibrillation (pAF) in stroke patients? *Cerebrovasc Dis* 2012;**33**:76.

387. Suissa L, Bresch S, Lachaud S, Mahagne MH. Optimal timing and duration of continuous ECG monitoring for detecting atrial fibrillation in stroke patients. *Cerebrovasc Dis* 2012;**33**:77.

388. Suissa L, Lachaud S, Mahagne MH. Optimal timing and duration of continuous electrocardiographic monitoring for detecting atrial fibrillation in stroke patients. *J Stroke Cerebrovasc Dis* 2013;**22**:991–5. http://dx.doi.org/10.1016/j.jstrokecerebrovasdis.2012.01.015

389. Suissa L, Lachaud S, Mahagne MH. Continuous ECG monitoring for tracking down atrial fibrillation after stroke: Holter or automated analysis strategy? *Eur Neurol* 2014;**72**:7–12. http://dx.doi.org/10.1159/000358053

390. Sun R, Wang Y, Yang S, Fang Z. [Detecting cardiac arrhythmias based on phase space analysis.] *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi* 2008;**25**:934–7, 949.

391. Sutamnartpong P, Dharmasaroja PA, Ratanakorn D, Arunakul I. Atrial fibrillation and paroxysmal atrial fibrillation detection in patients with acute ischemic stroke. *J Stroke Cerebrovasc Dis* 2014;**23**:1138–41. http://dx.doi.org/10.1016/j.jstrokecerebrovasdis.2013.09.032

392. Tagawa M, Takeuchi S, Chinushi M, Saeki M, Taniguchi Y, Nakamura Y, *et al.* Evaluating patients with acute ischemic stroke with special reference to newly developed atrial fibrillation in cerebral embolism. *Pacing Clin Electrophysiol* 2007;**30**:1121–8. https://doi.org/10.1111/j.1540-8159.2007.00823.x

393. Tateno K, Glass L. Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and deltaRR intervals. *Med Biol Eng Comput* 2001;**39**:664–71. https://doi.org/10.1007/BF02345439

394. Temiz A, Gazi E, Güngör O, Altun B, Barutcu A, Bekler A, *et al.* Fragmented QRS and prediction of paroxysmal atrial fibrillation episodes. *Pak J Med Sci* 2014;**30**:862–7. https://doi.org/10.12669/pjms.304.5064

395. 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* 2014;**16**:1291–5. http://dx.doi.org/10.1093/europace/euu057

396. Torbey S, Akl SG, Redfearn DP. Automated detection of paroxysmal atrial fibrillation in sinus rhythm. *Can J Cardiol* 2013;**29**:S127. https://doi.org/10.1016/j.cjca.2013.07.176

397. 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. *Intern Med J* 2014;**44**:505–8. http://dx.doi.org/10.1111/imj.12410

398. Tung CE, Su D, Turakhia MP, Lansberg MG. Diagnostic yield of extended cardiac patch monitoring in patients with stroke or TIA. *Front Neurol* 2014;**5**:266. http://dx.doi.org/10.3389/fneur.2014.00266

399. Ustrell X, Pellise A, Milian M, Merce J, Vinas J, Mares R. Improved detection of unknown atrial fibrillation in acute stroke patients through prolonged serial daily ECG assessments. *Cerebrovasc Dis* 2011;**31**:297.

400. Velthuis BO, Bos J, Kraaier K, Stevenhagen J, van Opstal JM, van der Palen J, Scholten MF. Performance of an external transtelephonic loop recorder for automated detection of paroxysmal atrial fibrillation. *Ann Noninvasive Electrocardiol* 2013;**18**:564–70. http://dx.doi.org/10.1111/anec.12075

401. Veltkamp R, Rizos T, Guntner J, Jenetzky E, Reichardt C, Hepp T, *et al.* Continuous stroke unit ECG monitoring detects intermittent atrial fibrillation more sensitively than 24 h Holter ECG after acute stroke and TIA. *Cerebrovasc Dis* 2011;**31**:31.

402. Veltkamp R, Guentner J, Jenetzky E, Becker R, Kirchhof P, Reinhardt R, *et al.* Continuous stroke unit ECG monitoring is better than 24-h Holter ECG for detection of paroxysmal atrial fibrillation after stroke. *Stroke* 2012;**43**:A131.

403. Villani GQ, Piepoli M, Rosi A, Capucci A. P-wave dispersion index: a marker of patients with paroxysmal atrial fibrillation. *Int J Cardiol* 1996;**55**:169–75. https://doi.org/10.1016/0167-5273(96)02677-0

404. Villani M, Iannucci G, Iannuzzi S, Machella L, Baciarello G. New parameters able to better identify patients at risk of atrial fibrillation. *Heartweb* 1996;**2**:U10–16.

405. Vinther KH, Tveskov C, Rosen T, Auscher S, Egstrup K. Detection of atrial fibrillation in patients with ischemic stroke by continuous in-hospital cardiac telemetry. *Eur Heart J* 2014;**35**:814.

406. Vyas V, Duran J, Ansaripour A, Niedzielko M, Steel A, Bakhai A. Does a 12-lead ECG more reliably detect atrial fibrillation than a rhythm strip only ECG? *Value Health* 2014;**17**:A485–6. http://dx.doi.org/10.1016/j.jval.2014.08.1421

407. Wallmann D, Tüller D, Wustmann K, Meier P, Isenegger J, Arnold M, *et al.* Frequent atrial premature beats predict paroxysmal atrial fibrillation in stroke patients: an opportunity for a new diagnostic strategy. *Stroke* 2007;**38**:2292–4. https://doi.org/10.1161/STROKEAHA.107.485110

408. Wiesel J, Wiesel D, Suri R, Messineo FC. The use of a modified sphygmomanometer to detect atrial fibrillation in outpatients. *Pacing Clin Electrophysiol* 2004;**27**:639–43. http://dx.doi.org/10.1111/j.1540-8159.2004.00499.x

409. Wiesel J, Messineo FC, Wiesel D. Screening for atrial fibrillation at home using a modified sphygmomanometer. *Stroke* 2007;**38**:539.

410. Wiesel J, Fitzig L, Herschman Y, Messineo FC. Detection of atrial fibrillation using a modified microlife blood pressure monitor. *Am J Hypertens* 2009;**22**:848–52. http://dx.doi.org/10.1038/ajh.2009.98

411. Wiesel J, Abraham S, Messineo FC. Detecting atrial fibrillation at home with a novel blood pressure monitor with atrial fibrillation detection functionality. *J Am Coll Cardiol* 2010;**55**(10 Suppl.):A15.E144.

412. 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). *Am J Cardiol* 2013;**111**:1598–601. http://dx.doi.org/10.1016/j.amjcard.2013.01.331

413. Wiesel J, Arbesfeld B, Schechter D. Comparison of the Microlife blood pressure monitor with the Omron blood pressure monitor for detecting atrial fibrillation. *Am J Cardiol* 2014;**114**:1046–8. http://dx.doi.org/10.1016/j.amjcard.2014.07.016

414. Wiesel J, Arbesfeld BZ, Schechter D. Screening for atrial fibrillation using a blood pressure monitor: Omron vs microlife. *Cardiology* 2014;**128**:372.

415. 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. *J Electrocardiol* 2011;**44**:460–4. http://dx.doi.org/10.1016/j.jelectrocard.2011.01.011

416. Wyse DG. What is the best method of detecting atrial fibrillation in people aged 65 years and older? *Nat Clin Pract Cardiovasc Med* 2008;**5**:138–9. https://doi.org/10.1038/ncpcardio1063

417. Xu W, Tse HF, Chan FH, Fung PC, Lee KL, Lau CP. New Bayesian discriminator for detection of atrial tachyarrhythmias. *Circulation* 2002;**105**:1472–9. https://doi.org/10.1161/01.CIR.0000012349.14270.54

418. Yaghouby F, Ayatollahi A, Bahramali R, Yaghouby M, Alavi AH. Towards automatic detection of atrial fibrillation: a hybrid computational approach. *Comput Biol Med* 2010;**40**:919–30. http://dx.doi.org/10.1016/j.compbiomed.2010.10.004

419. Yiğit Z, Akdur H, Ersanli M, Okçün B, Güven O. The effect of exercise to P wave dispersion and its evaluation as a predictor of atrial fibrillation. *Ann Noninvasive Electrocardiol* 2003;**8**:308–12. https://doi.org/10.1046/j.1542-474X.2003.08408.x

420. Yung LTM, Brader S, Jordan D, Foley P. A comparison of mobile cardiac telemetry systems (MCT) with conventional cardiac monitoring for the investigation of common conditions suggestive of cardiac arrhythmias. *Europace* 2010;**12**:ii16.

421. Zenk BM, Bratton RL, Flipse TR, Page EE. Accuracy of detecting irregular cardiac rhythms via telemedicine. *J Telemed Telecare* 2004;**10**:55–8. http://dx.doi.org/10.1258/135763304322764211

422. Zhou X, Ding H, Ung B, Pickwell-MacPherson E, Zhang Y. Automatic online detection of atrial fibrillation based on symbolic dynamics and Shannon entropy. *Biomed Eng Online* 2014;**13**:18. http://dx.doi.org/10.1186/1475-925X-13-18.

Appendix 1 Search strategy for the diagnostic test accuracy review

The MEDLINE search strategy for the DTA review is provided below.

1. exp Atrial Fibrillation/ (35,633)
2. (atrial adj3 fibrillat$).ti,ab. (41,204)
3. ((auricular adj3 fibrillat$) or (supraventricul$ adj3 arrhythmi$)).ti,ab. (3058)
4. or/1-3 (51,726)
5. "Sensitivity and Specificity"/ (282,088)
6. ((diagnos$ or screening or triage) adj3 (accurate$ or accuracy)).ti,ab. (57,907)
7. ((sensitivity or specificity) adj6 (diagnos$ or screen$ or detect$ or PAF or AF)).ti,ab. (75,779)
8. ((diagnos$ or detect$) adj3 (rate or yield or PAF)).ti,ab. (27,930)
9. or/5-8 (392,368)
10. 4 and 9 (1684)
11. exp *Atrial Fibrillation/di [Diagnosis] (2201)
12. ((diagnos$ or underdiagnos$ or detect$ or identif$ or screen$ or predictive value$) adj6 (PAF or AF or atrial fibrillat$)).ti,ab. (4539)
13. *diagnosis/ or diagnosis, computer-assisted/ or diagnosis, differential/ or early diagnosis/ or predictive value of tests/ or ROC curve/ (584,709)
14. 11 or 12 or (13 and 4) (8483)
15. Electrocardiography, Ambulatory/ (9052)
16. (holter or single lead or 12-lead or event monitor$ or event record$ or loop record$ or ELR).ti,ab. (15894)
17. 15 and 16 (4003)
18. *Electrocardiography, Ambulatory/ (2842)
19. ((ECG or iECG or electrocardiogra$ or EKG) adj3 (single lead or serial or intermittent or bipolar or bi-polar or thumb or short-term or 12-lead or ambulatory or portable)).ti,ab. (9039)
20. ((ECG or iECG or electrocardiogra$ or EKG) adj6 (ELR or holter or event monitor$ or event record$ or loop record$)).ti,ab. (3589)
21. ((holter or cardiac event or R-test or 7-day) adj3 monitor$).ti,ab. (4951)
22. Pulse/ and (palpation/ or palpat$.ti,ab.) (217)
23. (pulse adj3 (finger-tip or palp$)).ti,ab. (404)
24. finger probe$.ti,ab. (87)
25. Watchbp.ti,ab. (18)
26. (exp Sphygmomanometers/ or Blood pressure Monitoring, ambulatory/) and (modified or atrial fibrillat$ or PAF or AF).ti,ab. (173)
27. ((modified or atrial fibrillat$ or PAF or AF) adj3 (BP monitor$ or blood pressure monitor$ or sphygmomanomet$)).ti,ab. (41)
28. photoplethysmograph$.ti,ab. (1628)
29. ((reveal or implantable) adj3 device$).ti,ab. (3458)
30. ((mobile or i-phone) adj3 app$).ti,ab. (1564)
31. or/17-30 (24,262)
32. 14 and 31 (725)
33. 10 or 32 (2181)
34. letter/ (858,649)
35. editorial/ (366,888)
36. news/ (165,913)
37. exp historical article/ (322,828)
38. Anecdotes as topic/ (4584)

39. comment/ (607,051)
40. case report/ (1,706,134)
41. (letter or comment$).ti. (100,072)
42. or/34-41 (3,406,548)
43. randomized controlled trial/ or Randomized Controlled Trials as Topic/ or random$.ti,ab. (883,882)
44. 42 not 43 (3,375,361)
45. animals/ not humans/ (3,881,514)
46. exp Animals, Laboratory/ (728,296)
47. exp Animal Experimentation/ (6461)
48. exp Models, Animal/ (422,793)
49. exp rodentia/ (2,679,540)
50. (rat or rats or mouse or mice).ti. (1,120,787)
51. or/44-50 (7,862,299)
52. 33 not 51 (1961)

Appendix 2 QUADAS-2 quality assessment

Ideal study

Population:	People who have not sought medical attention on account of symptoms associated with AF
Presentation:	Asymptomatic/not sought medical attention on account of symptoms associated with AF presenting to primary care or the community (for example community pharmacists). Individuals may be invited to screening regardless of medical history (this may be done on the basis of age, systematic screening); present to the GP for an unrelated issues (for example flu vaccination, opportunistic screening); or based on their medical history/the presence of risk factors that are associated with AF (targeted screening)
Prior tests:	No prior testing for AF
Index test:	Any non-invasive test that could be utilised in a primary care setting or the community
Purpose:	Screening test, to identify people with AF who have not sought medical attention on account of symptoms associated with AF
Target disorder:	AF
Reference standard:	12-lead ECG interpreted by a cardiologist

The 'ideal' study for AF screening tests

Low risk of bias

- A consecutive or random sample of people was enrolled

- A case-control design was avoided

- Inappropriate exclusions were avoided (for example the presence of a different condition that may cause arrhythmia for example atrial flutter, supraventricular tachycardia, ventricular tachycardia, ventricular fibrillation, heart block, tachy-brady syndrome)

- The index test was objective or was interpreted without knowledge of the reference standard.

- The reference standard was a gold standard diagnostic technique (12-lead ECG interpreted by a cardiologist)

- The reference standard was objective or was interpreted without the knowledge of the index test.

- All index tests and the reference standard were performed concurrently.

- If cut-offs were used these were pre-specified.

- All participants received all tests and the reference standard.

- There were no unclear/uninterpretable test results.

- All participants were included in the analysis

High applicability

- The population is asymptomatic/has not sought medical attention on account of symptoms associated with AF. The population has undergone no prior testing for AF.

- The population was recruited into the screening study based on:
  - No criteria or age ('ideal' for systematic screening)
  - Presentation to the GP/other setting for an issue unrelated to AF ('ideal' for opportunistic screening)
  - Medical history/presence of risk factors associated with AF ('ideal' for targeted screening)

- The test is performed and interpreted in primary care or the community.

- The reference standard is the gold standard diagnostic technique (12-lead ECG interpreted by a cardiologist)

QUADAS-2

Domain 1: Patient selection

Risk of Bias: Could the selection of patients have introduced bias?

Describe the method of patient selection

Signalling question 1:   Was a consecutive or random sample of patients enrolled?
                          Yes/No/Unclear

Yes   If a consecutive or random samples of patients was enrolled.

Signalling question 2:   Was a case control-design avoided?
                          Yes/No/Unclear

Yes   If the study did not use a two-gate entry procedure (i.e did not include any patients on the basis of having diagnosed AF)

Signalling question 3:   Did the study avoid inappropriate exclusions?

Yes/No/Unclear

Yes   If inappropriate exclusions were avoided (for example excluding based on the presence of a different condition that may cause arrhythmia for example atrial flutter, supraventricular tachycardia, ventricular tachycardia, ventricular fibrillation, heart block, tachy-brady syndrome)

Appropriate exclusions: diagnosed AF; patients with paced rhythms/pacemakers/defibrillators/other cardiac devices; severe medical condition preventing participation (e.g. severe dementia or terminal illness); age

Unclear   If exclusions are not detailed

Conclusion: **Could the selection of patients have introduced bias?**

High/Low/Unclear

(If the response to all the signalling questions is 'yes' the study can be considered at a low risk of bias; if the response to question 2 is 'no' (i.e. a case-control design was used) then it will be judged that the study is at high risk of bias)

Applicability: Are there concerns that the included patients and setting do not match the review question?

Describe included patients (prior testing, presentation, intended use of index test and setting)

Signalling question 1:   Was the population asymptomatic/had not sought medical attention on account of symptoms associated with AF?

Yes/No/Unclear

Signalling question 2:   Was the population recruited from primary care/the community?

Yes/No/Unclear

Signalling question 3:   Was inclusion into the study independent of the results of prior testing that could be used to detect AF?

Yes/No/Unclear

Type of screening programme question:

Was the population recruited based on:

- No criteria or age (systematic screening)

- Presentation to the GP/other setting for an issue unrelated to AF (opportunistic screening)

- Medical history/presence of risk factors associated with AF (targeted screening). NB GRASP-AF score includes items for congestive heart failure, hypertension, age≥75, diabetes mellitus, prior stroke/TIA/thromboembolism and vascular disease.

- Other

Signalling question 4:   Was the population representative of the population that would be expected to be tested by systematic screening, opportunistic screening or targeted screening?

Yes/No/Unclear

Conclusion: **Are there concerns that the included patients and setting do not match the review question?**

High/Low/Unclear

(If the response to all the signalling questions is 'yes' then concerns over applicability are low. If the population was not recruited from primary care/the community then applicability concerns are high)

Domain 2: Index test

Risk of Bias: Could the conduct or interpretation of the index test have introduced bias?

Describe the index test and how it was conducted and interpreted:

Signalling question 1:   Were the index test results interpreted without knowledge of the results of the reference standard?

Yes/No/Unclear

Yes   If the index test was always conducted and interpreted before the reference standard, or

If the index test was objective, or

If the interpreters of the index test were blinded to the results of the reference standard.

Signalling question 2:   If a threshold was used, was it pre-specified?

Yes/No/Unclear/NA

Yes   If threshold used were pre-specified, and were not defined post-hoc based on study data.

Signalling question 3: Did the person interpreting the index test have access to information or training that would not be available if the test was to be performed in the community/in primary care?

Conclusion: **Could the conduct or interpretation of the index test have introduced bias?**

Yes/No/Unclear

(If the response to all the signalling questions is 'yes' the study can be considered at a low risk of bias. If the threshold was not pre-specified then the risk of bias is high)

Applicability: Are there concerns that the index test, its conduct, or its interpretation differ from the review question?

Signalling question 1: Was the index test performed in primary care or the community?

Yes/No/Unclear

Signalling question 2: Was the index test interpreted in primary care, in the community, or using an automated method?

Yes/No/Unclear

Signalling question 3: Was the index test performed and interpreted without the person performing and interpreting the test having to undergo special training?

Signalling question 4: Were the same clinical data available when the test was interpreted as would be available when the test was used in practice?

Yes   If interpreters had accesses to the same clinical data as when the test would be interpreted in practice. NB studies that blinded interpreters to clinical data are still of high applicability because it may be that GP notes and medical records are not available in a screening setting.

Conclusion: **Are there concerns that the index test, its conduct, or its interpretation differ from the review question?**

High/Low/Unclear

(If the response to all the signalling questions is 'yes' then concerns over applicability are low. If the index test was interpreted by a cardiologist/someone in secondary care then the concerns about applicability are high.)

Domain 3: Reference standard

Risk of Bias: Could the reference standard, its conduct or its interpretation have introduced bias?

Describe the reference standard and how it was conducted and interpreted

Signalling question 1: Is the reference standard likely to correctly classify AF

Yes If 12-lead ECG interpreted by a cardiologist

Signalling question 2: Were the reference standard results interpreted without knowledge of the results of the index test?

Yes/No/Unclear

Yes If the reference standard was always conducted and interpreted before the reference standard or if the reference standard was objective or if the interpreters of the reference standard were blinded to the results of the index test.

Conclusion: Could the conduct or interpretation of the reference test have introduced bias?

High/Low/Unclear

(If the response to all the signalling questions is 'yes' the study can be considered at a low risk of bias.)

Applicability: Are there concerns that the target condition as defined by the reference standard does not match the condition?

Conclusion: Are there concerns that the target condition as defined by the reference standard does not match the condition?

High/Low/Unclear

(Low if 12-lead ECG interpreted by a cardiologist)

Domain 4: Flow and timing

Risk of Bias: Could the patient flow have introduced bias?

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram)

Describe the time interval and any interventions between index test(s) and reference standard

Signalling question 1: Were the index test and reference standard performed within 7 days of each other?

<div align="center">Yes/No/Unclear</div>

Signalling question 2: Did all patients receive the same reference standard?

Yes/No/Unclear

Yes   If all patients received the same reference standard

Signalling question 3: Were ≥80% of patients included in the analysis?

<div align="center">Yes/No/Unclear</div>

Yes   If <20% of participants were excluded due to missing/un-interpretable tests?

Conclusion: **Could the patient flow have introduced bias?**

<div align="center">**High/Low/Unclear**</div>

(If the response to all the signalling questions is 'yes' the study can be considered at a low risk of bias)

Appendix 3 Methods for meta-analysis of the diagnostic test accuracy studies

We assumed that the reference test is perfectly accurate, and so we used the HSROC model of Rutter and Gatsonis[53] to estimate the sensitivity and specificity of the index (experimental) test. We reparameterised the slope/asymmetry of their model to be multiplicative rather than exponential, thus easing interpretation, but the two parameterisations are equivalent. Each study i can include several groups j being tested by the same test. Each of these groups within studies are referred to as separate observations and are the unit of our analysis. All included studies reported the number of true positives, false positives, true negatives and false negatives for each of these observations.

For each separate index test, we used the notation in *Table 53* where r_{iD} is the number of 'true positives', $n_{iD} - r_{iD}$ is the number of 'false negatives', $s_{i\bar{D}}$ is the number of 'false positives' and $n_{i\bar{D}} - s_{i\bar{D}}$ is the number of 'true negatives' in observation j of study i of the index test.

We used the following likelihood for these observations:

$$r_{ijD} \sim Bin(\pi_{ij1}, n_{ijD}) \qquad \text{(sensitivity or 'true positives')} \tag{1}$$

$$s_{ij\bar{D}} \sim Bin(\pi_{ij2}, n_{ij\bar{D}}) \qquad \text{(1 – specificity or 'false positives')}, \tag{2}$$

for $i = 1, \ldots, N_1$ indicating the trials comparing the index to the reference test and observations $j = 1, \ldots, Nobs_i$ in trial i. Logistic link functions were used to model the probabilities:

$$logit(\pi_{ij1}) = \beta \times \theta_{ij} + \frac{1}{2}\alpha_{ij} \tag{3}$$

$$logit(\pi_{ij2}) = \theta_{ij} - \frac{1}{2}\alpha_{ij}, \tag{4}$$

with a random effect on the parameters to model between-observation (within-trial) variation:

$$\theta_{ij} \sim N(q_i, \omega_1^2) \tag{5}$$

$$\alpha_{ij} \sim N(a_i, \omega_2^2). \tag{6}$$

We used a hierarchical model for the parameters to capture between-study variation:

$$q_i \sim N(\Theta, \sigma_1^2) \tag{7}$$

$$a_i \sim N(\Lambda, \sigma_2^2), \tag{8}$$

TABLE 53 Notation for DTA study results for observation group j of study i

AF status	Positive test	Negative test	Total
AF	r_{ijD}	$n_{ijD} - r_{ijD}$	n_{ijD}
No AF	$s_{ij\bar{D}}$	$n_{ij\bar{D}} - s_{ij\bar{D}}$	$n_{ij\bar{D}}$

with priors on the hyper-parameters:

$$\beta \sim N(0, 10^4) \tag{9}$$

$$\theta \sim N(0, 10^4) \tag{10}$$

$$\Lambda \sim N(0, 10^4) \tag{11}$$

$$\omega_m \sim U(0, 5), l = 1, 2 \tag{12}$$

$$\sigma_m \sim U(0, 5), l = 1, 2, \tag{13}$$

where θ is the (absolute) threshold of the index test and increasing it increases both the number of true positives and the number of false positives, Λ is the (absolute) accuracy of the test and determines how good it is at distinguishing between true positives and false positives, and β is a slope parameter representing asymmetry around the central diagonal of the SROC curve. This slope may be dropped if too few data are available to fit the model or if model choice criteria (e.g. residual deviance, DIC) indicate that a symmetric SROC would provide a better fit.

The ω_m^2 and σ_m^2 are the between-observation and between-study variances respectively. We investigated a fixed-effects model, which assumes no such variation between tests:

$$logit(\pi_{ij1}) = \beta \times \Theta + \frac{1}{2}\Lambda \tag{14}$$

$$logit(\pi_{ij2}) = \Theta - \frac{1}{2}\Lambda. \tag{15}$$

Some test groups (observations) in the same study differed only in the interpreter used (e.g. two different nurses or a GP in one group and a cardiologist in another, interpreting the same single-lead ECG). These are considered as independent observations (e.g. two separate groups of individuals receiving an ECG interpreted by nurse). However, two analyses of included tests were conducted. One defined tests independently of interpreter (i.e. a single-lead ECG interpreted by a nurse and a GP were considered the same test) and the other defined tests by interpreter (e.g. a single-lead ECG interpreted by a nurse and a GP was considered to comprise two different types of test). A choice between these units of analyses was made on the basis of convergence, residual deviance and DIC.

Summary sensitivities and specificities were produced using the posterior means of the following functions of the hyper-parameters:

$$logit(sensitivity) = \beta \times \Theta + \frac{1}{2}\Lambda \tag{16}$$

$$logit(1 - specificity) = \Theta - \frac{1}{2}\Lambda. \tag{17}$$

The DOR is defined as:

$$DOR = \frac{sensitivity^2}{(1 - specificity)^2},$$ (18)

which reduces to:

$$DOR = e^{\Lambda}.$$ (19)

Subgroup analyses were mostly binary categories, in which the model was fitted for each group separately. The exception was prevalence of AF, which was explored using regression with covariates x_{ij} (AF prevalence), centred on mean $\bar{x}$ over all studies and observations, using the model:

$$\theta_{ij} \sim N(q_i + (x_{ij} - \bar{x})\gamma, \omega_1^2)$$ (20)

$$\alpha_{ij} \sim N(a_i + (x_{ij} - \bar{x})\lambda, \omega_2^2),$$ (21)

in which the index test-specific regression coefficients were given the following priors:

$$\gamma \sim N(0, 10^4)$$ (22)

$$\lambda \sim N(0, 10^4).$$ (23)

Alternative assumptions: independent and exchangeable slopes

In the above model, a common slope β was assumed for every test. We also considered the model in which the slope β_k for each test k is independent with priors:

$$\beta_k \sim N(0, 10^4).$$ (24)

We also considered the model in which the slopes are exchangeable and follow the distribution:

$$\beta_k \sim N(\beta, \sigma_\beta^2),$$ (25)

with priors:

$$\beta \sim N(0, 10^4)$$ (26)

$$\sigma_\beta \sim U(0, 5).$$ (27)

We chose between the four possible slope assumptions (no slope, common slope, independent slope, exchangeable slope) on the basis of convergence, residual deviance and DIC.

WinBUGS code

Random-effects hierarchical summary receiver operating characteristic model

```
# Inputs needed are
# tp[i,j,k], fp[i,j,k], pos[i,j,k], neg[i,j,k] i.e. summaries from
the 2x2 tables for replication i, study j, the kth test for study j
# You may want to re-order the indexes to make it easier for data
input to the arrays.
# ntests[j] a vector with number of tests per study
# nreps[j,k] the number of replcations of the kth test of study j
# t[j,k] indicates the test number for the kth test of study j.
# NTests is the total number of tests
# N1 is the number of studies with 2x2 data (ie your type 1 and type
3 studies)
# N2 is the number of studies with sensitivity/specificity data
# sens.mean[i,j,k] is the estimated sensitivity in ith rep of the
k'th test in study j, and sens.prec[i,j,k] is it's precision.

Model{
#Loop over studies reporting 2x2 tables. i=rep no., k=test no. in
study j
for (j in 1:N1){
     for (k in 1:ntests[j]){
          for (i in 1:nreps[j,k]){
               tp[i,j,k]~dbin(pi1[i,j,k],pos[i,j,k])
     #Likelihood
               fp[i,j,k]~dbin(pi2[i,j,k],neg[i,j,k])
               tphat[i,j,k]<-pi1[i,j,k]*pos[i,j,k]
     #Fitted values
               fphat[i,j,k]<-pi2[i,j,k]*neg[i,j,k]
               dev.tp[i,j,k]<-(2*(tp[i,j,k]*(log(tp[i,j,k])-
log(tphat[i,j,k]))      #Deviance contibution tp
          +  (pos[i,j,k]-tp[i,j,k])*(log(pos[i,j,k]-tp[i,j,k]) -
log(pos[i,j,k]-tphat[i,j,k]))))
               dev.fp[i,j,k]<-(2*(fp[i,j,k]*(log(fp[i,j,k])-
log(fphat[i,j,k]))      #Deviance contibution fp
          +  (neg[i,j,k]-fp[i,j,k])*(log(neg[i,j,k]-fp[i,j,k]) -
log(neg[i,j,k]-fphat[i,j,k]))))
          }
          dev[j,k]<-sum(dev.tp[1:nreps[j,k],j,k]) +
sum(dev.fp[1:nreps[j,k],j,k])
     }
     sumdev[j]<-sum(dev[j,1:ntests[j]])
}
```

```
totresdev<- sum(sumdev[])


#Loop over all studies
for (j in 1:(N1)){
     for (k in 1:ntests[j]){
          for (i in 1:nreps[j,k]){
               logit(pi1[i,j,k])<- beta[t[j,k]]*theta[i,j,k]
+alpha[i,j,k]/2        #Model for sensitivity
               logit(pi2[i,j,k])<- theta[i,j,k] - alpha[i,j,k]/2
                                #Model for (1-specificity)
               theta[i,j,k]~dnorm(q[j,k],p.reps[1])
     #Between rep variability
               alpha[i,j,k]~dnorm(a[j,k],p.reps[2])
          }
          q[j,k]~dnorm(Q[t[j,k]],p.stud[1])            #Between
study variability
          a[j,k]~dnorm(A[t[j,k]],p.stud[2])
     }
}


#PRIORS
for (k in 1:NTests){    #Loop over tests
     Q[k]~dnorm(0,.0001)
     A[k]~dnorm(0,.0001)
     log(beta.true[k])<-l.beta.true
     beta[k]<- include.slope[k]*beta.true[k] + (1-include.slope[k])
}
l.beta.true~dnorm(0,.0001)


for (m in 1:2){
     p.stud[m]<-pow(sd.stud[m],-2)
     p.reps[m]<-pow(sd.reps[m],-2)
     sd.stud[m]~dunif(0,5)
     sd.reps[m]~dunif(0,5)

     # Save the prior distributions
     sd.stud.prior[m]~dunif(0,5)
     sd.reps.prior[m]~dunif(0,5)
}
```

```
        # Posterior and posterior predictive distritbutions
        for(k in 1:NTests){
                theta.pred[k]~dnorm(Q[k],p.stud[1])
                alpha.pred[k]~dnorm(A[k],p.stud[2])
                logit(sens.pred[k])<-
beta[k]*theta.pred[k]+0.5*alpha.pred[k]
                logit(fpr.pred[k])<-theta.pred[k]-0.5*alpha.pred[k]
                spec.pred[k]<- 1 - fpr.pred[k]

                logit(TPR[k])<-(beta[k]*Q[k]+A[k]/2)
                logit(FPR[k])<-(Q[k]-A[k]/2)
                sensitivity[k]<-TPR[k]
                specificity[k]<-1-FPR[k]

                # Diagnostic odds ratio
                dor[k]<-(sensitivity[k]*specificity[k])/((1-
sensitivity[k])*(1-specificity[k]))
                #DOR[k] <- exp(A[k])
        }
}
```

Fixed-effects model

```
# Fixed effects over both studies and reps
model.DTA.meta.analysis.noType2.fixed<-function(){
#Loop over studies reporting 2x2 tables. i=rep no., k=test no. in
study j
for (j in 1:N1){
     for (k in 1:ntests[j]){
          for (i in 1:nreps[j,k]){
               tp[i,j,k]~dbin(pi1[i,j,k],pos[i,j,k])
     #Likelihood
               fp[i,j,k]~dbin(pi2[i,j,k],neg[i,j,k])
               tphat[i,j,k]<-pi1[i,j,k]*pos[i,j,k]
     #Fitted values
               fphat[i,j,k]<-pi2[i,j,k]*neg[i,j,k]
               dev.tp[i,j,k]<-(2*(tp[i,j,k]*(log(tp[i,j,k])-
log(tphat[i,j,k]))     #Deviance contibution tp
          + (pos[i,j,k]-tp[i,j,k])*(log(pos[i,j,k]-tp[i,j,k]) -
log(pos[i,j,k]-tphat[i,j,k]))))
               dev.fp[i,j,k]<-(2*(fp[i,j,k]*(log(fp[i,j,k])-
log(fphat[i,j,k]))     #Deviance contibution fp
          + (neg[i,j,k]-fp[i,j,k])*(log(neg[i,j,k]-fp[i,j,k]) -
log(neg[i,j,k]-fphat[i,j,k]))))
          }
          dev[j,k]<-sum(dev.tp[1:nreps[j,k],j,k]) +
sum(dev.fp[1:nreps[j,k],j,k])
     }
     sumdev[j]<-sum(dev[j,1:ntests[j]])
}


totresdev<- sum(sumdev[])


#Loop over all studies
for (j in 1:(N1)){
     for (k in 1:ntests[j]){
          for (i in 1:nreps[j,k]){
               logit(pi1[i,j,k])<- beta[t[j,k]]*Q[t[j,k]] +
A[t[j,k]]/2     #Model for sensitivity
               logit(pi2[i,j,k])<- Q[t[j,k]] - A[t[j,k]]/2
                    #Model for (1-specificity)
          }


     }
}
```

```
#PRIORS
for (k in 1:NTests){     #Loop over tests
      Q[k]~dnorm(0,.0001)
      A[k]~dnorm(0,.0001)
      log(beta.true[k])<-l.beta.true
      beta[k]<- include.slope[k]*beta.true[k] + (1-include.slope[k])
}
l.beta.true~dnorm(0,.0001)


for (m in 1:2){
      p.stud[m]<-pow(sd.stud[m],-2)
      p.reps[m]<-pow(sd.reps[m],-2)
      sd.stud[m]~dunif(0,5)
      sd.reps[m]~dunif(0,5)


      # Save the prior distributions
      sd.stud.prior[m]~dunif(0,5)
      sd.reps.prior[m]~dunif(0,5)
}


      # Posterior and posterior predictive distritbutions
      for(k in 1:NTests){
            theta.pred[k]~dnorm(Q[k],p.stud[1])
            alpha.pred[k]~dnorm(A[k],p.stud[2])
            logit(sens.pred[k])<-
beta[k]*theta.pred[k]+0.5*alpha.pred[k]
            logit(fpr.pred[k])<-theta.pred[k]-0.5*alpha.pred[k]
            spec.pred[k]<- 1 - fpr.pred[k]

            logit(TPR[k])<-(beta[k]*Q[k]+A[k]/2)
            logit(FPR[k])<-(Q[k]-A[k]/2)
            sensitivity[k]<-TPR[k]
            specificity[k]<-1-FPR[k]

            # Diagnostic odds ratio
            dor[k]<-(sensitivity[k]*specificity[k])/((1-
sensitivity[k])*(1-specificity[k]))
            #DOR[k] <- exp(A[k])
      }
}
```

Meta-regression model

```
Model{
#Loop over studies reporting 2x2 tables. i=rep no., k=test no. in
study j
for (j in 1:N1){
     for (k in 1:ntests[j]){
          for (i in 1:nreps[j,k]){
               tp[i,j,k]~dbin(pi1[i,j,k],pos[i,j,k])
     #Likelihood
               fp[i,j,k]~dbin(pi2[i,j,k],neg[i,j,k])
               tphat[i,j,k]<-pi1[i,j,k]*pos[i,j,k]
     #Fitted values
               fphat[i,j,k]<-pi2[i,j,k]*neg[i,j,k]
               dev.tp[i,j,k]<-(2*(tp[i,j,k]*(log(tp[i,j,k])-
log(tphat[i,j,k]))     #Deviance contibution tp
          +  (pos[i,j,k]-tp[i,j,k])*(log(pos[i,j,k]-tp[i,j,k]) -
log(pos[i,j,k]-tphat[i,j,k]))))
               dev.fp[i,j,k]<-(2*(fp[i,j,k]*(log(fp[i,j,k])-
log(fphat[i,j,k]))     #Deviance contibution fp
          +  (neg[i,j,k]-fp[i,j,k])*(log(neg[i,j,k]-fp[i,j,k]) -
log(neg[i,j,k]-fphat[i,j,k]))))
          }
          dev[j,k]<-sum(dev.tp[1:nreps[j,k],j,k]) +
sum(dev.fp[1:nreps[j,k],j,k])
     }
     sumdev[j]<-sum(dev[j,1:ntests[j]])
}

totresdev<- sum(sumdev[])

#Loop over all studies
for (j in 1:(N1)){
     for (k in 1:ntests[j]){
          for (i in 1:nreps[j,k]){
               logit(pi1[i,j,k])<- beta[t[j,k]]*theta[i,j,k]
+alpha[i,j,k]/2          #Model for sensitivity
               logit(pi2[i,j,k])<- theta[i,j,k] - alpha[i,j,k]/2
                              #Model for (1-specificity)
               theta[i,j,k]~dnorm(qreg[i,j,k],p.reps[1])
     #Between study variability
               alpha[i,j,k]~dnorm(areg[i,j,k],p.reps[2])
```

```
                # Regression
                qreg[i,j,k]<-q[j,k]+(x[i,j,k]-
mean.x[t[j,k]])*gamma[t[j,k]]
                areg[i,j,k]<-a[j,k]+(x[i,j,k]-
mean.x[t[j,k]])*lambda[t[j,k]]
            }
            q[j,k]~dnorm(Q[t[j,k]],p.stud[1])                #Between
study variability
            a[j,k]~dnorm(A[t[j,k]],p.stud[2])
        }
}


#PRIORS
for (k in 1:NTests){    #Loop over tests
        Q[k]~dnorm(0,.0001)
        A[k]~dnorm(0,.0001)

        # Regression coefficients
        gamma[k]~dnorm(0,.0001)
        lambda[k]~dnorm(0,.0001)

        log(beta.true[k])<-l.beta.true
        beta[k]<- include.slope[k]*beta.true[k] + (1-include.slope[k])
}
l.beta.true~dnorm(0,.0001)

for (m in 1:2){
        p.stud[m]<-pow(sd.stud[m],-2)
        p.reps[m]<-pow(sd.reps[m],-2)
        sd.stud[m]~dunif(0,5)
        sd.reps[m]~dunif(0,5)

        # Save the prior distributions
        sd.stud.prior[m]~dunif(0,5)
        sd.reps.prior[m]~dunif(0,5)
}

        # Posterior and posterior predictive distritbutions
        for(k in 1:NTests){
                theta.pred[k]~dnorm(Q[k],p.stud[1])
                alpha.pred[k]~dnorm(A[k],p.stud[2])
```

```
                  logit(sens.pred[k])<-
beta[k]*theta.pred[k]+0.5*alpha.pred[k]
                  logit(fpr.pred[k])<-theta.pred[k]-0.5*alpha.pred[k]
                  spec.pred[k]<- 1 - fpr.pred[k]


                  logit(TPR[k])<-(beta[k]*Q[k]+A[k]/2)
                  logit(FPR[k])<-(Q[k]-A[k]/2)
                  sensitivity[k]<-TPR[k]
                  specificity[k]<-1-FPR[k]


                  # Diagnostic odds ratio
                  dor[k]<-(sensitivity[k]*specificity[k])/((1-
sensitivity[k])*(1-specificity[k]))
                  #DOR[k] <- exp(A[k])
            }
      }
```

Appendix 4 Search strategy for the systematic review of randomised controlled trials comparing screening strategies

The search strategy and MEDLINE search terms for the systematic review of RCTs comparing screening strategies are provided in *Table 54*.

Database: Ovid MEDLINE(R) < 1946 to present>

TABLE 54 Search protocol and MEDLINE search strategy for the systematic review of RCTs comparing screening strategies

Component	Description
Review area	Screening strategies for AF
Objectives	To investigate the effectiveness of systematic screening for AF
Populations/aspect	Adults of unknown AF disease status
Target condition	AF
Interventions	Systematic screening strategies – population, opportunistic, targeted
Comparisons/aspects covered by search	Usual care – no screening
Study designs	RCTs (including cluster and quasi-types), interrupted times series, controlled before-and-after studies
Exclusions	Animal studies
How the information was searched	Databases: MEDLINE and PreMEDLINE, EMBASE, CENTRAL – Cochrane Central Register of Controlled Trials
	Date parameters: April and May 2012; July–December 2015
	Note: because of the large number of conference abstracts added to EMBASE since the original search, these were excluded unless they met very specific AF/screening criteria
Search date	22 January 2016
Search terms	Atrial Fibrillation/ (38,865)
	(atrial fibrillation* or atrium fibrillation* or auricular fibrillation*).tw. (45,810)
	(af or a-fib).tw. (25,013)
	Atrial Flutter/ (5240)
	(atrial flutter* or auricular flutter*).tw. (4909)
	or/1-5 (67,777)
	Mass Screening/ (85,789)
	screen*.tw. (522,680)
	Diagnosis/ (16,823)
	"diagnostic techniques and procedures"/ (2609)
	(diagnos* or identif* or detect* or prevalence or incidence*).tw. (5,517,698)
	((systemat* or opportunist* or target* or population or mass) adj2 (test* or assess*)).tw. (20,842)
	Electrocardiography/ (171,668)

TABLE 54 Search protocol and MEDLINE search strategy for the systematic review of RCTs comparing screening strategies (*continued*)

Component	Description
	Electrocardiography, Ambulatory/ (9377)
	(electrocardiogram* or electrocardiograph* or ecg or ekg or holter or event monitor*).tw. (112,189)
	pulse/ (16,648)
	(pulse adj3 (test or tests)).tw. (867)
	or/7-17 (5,920,125)
	6 and 18 (31,631)
	randomized controlled trial.pt. (404,537)
	controlled clinical trial.pt. (90,002)
	randomi#ed.ab. (398,345)
	placebo.ab. (165,463)
	clinical trials as topic.sh. (174,355)
	randomly.ab. (240,675)
	trial.ti. (144,385)
	or/20-26 (1,012,739)
	19 and 27 (3572)
	exp animals/ not humans.sh. (4,175,239)
	28 not 29 (3504)
	intervention*1.ti. or (intervention*1 adj6 (clinician*1 or collaborat* or community or complex or DESIGN$ or GP*1 or educational or family GP* or family physician*1 or family practitioner*1 or financial or GP or general practice*1 or hospital*1 or impact*1 or improv$ or individuali?e*1 or individuali?ing or interdisciplin$ or multicomponent or multi-component or multidisciplin$ or multi-disciplin$ or multifacet$ or multi-facet$ or multimodal$ or multi-modal$ or personali?e*1 or personali?ing or pharmacies or pharmacist*1 or pharmacy or physician*1 or practitioner*1 or prescrib$ or prescription*1 or primary care or professional$ or provider*1 or regulatory or tailor$ or target$ or team$ or usual care)).ab. (191204)
	(pre-intervention*1 or preintervention*1 or post-intervention*1 or postintervention*1).tw. (12,843)
	(hospital$ or patient*1).hw. and (study or studies or care or health$ or practitioner*1 or provider*1 or physician*1 or nurse*1 or nursing or GP*1).ti,hw. (775,070)
	demonstration project*1.ti,ab. (2091)
	(pre-post or pre-test* or pretest$ or posttest$ or post-test* or (pre adj5 post)).ti,ab. (77,200)
	(pre-workshop or post-workshop or (before adj3 workshop) or (after adj3 workshop)).ti,ab. (723)
	(trial or study).ti. or ((study adj3 aim*1) or present study or our study).ab. (2,200,119)
	(before adj10 (after or during)).ti,ab. (389,176)
	(time points adj3 (over or multiple or three or four or five or six or seven or eight or nine or ten or eleven or twelve or month$ or hour*1 or day*1 or more than)).ab. (11,012)
	pilot.ti. (46,287)
	(multicentre or multicenter or multi-centre or multi-center).ti. (33,605)
	random$.ti,ab. or controlled.ti. (849,995)
	*experimental design/ or *pilot study/ or quasi experimental study/ (29,846)

TABLE 54 Search protocol and MEDLINE search strategy for the systematic review of RCTs comparing screening strategies (*continued*)

Component	Description
	(quasi-experiment$ or quasiexperiment$ or quasi-random$ or quasirandom$ or quasi control$ or quasicontrol$ or ((quasi$ or experimental) adj3 (method$ or study or trial or design$))).ti,ab. (113,933)
	(time series adj2 interrupt$).ti,ab. (1329)
	controlled before-after studies/ or interrupted time series analysis/ (200)
	or/31-46 (3,952,382)
	19 and 47 (9786)
	review.ti. (301,040)
	(animal$ not human$).sh,hw. (4,135,024)
	49 or 50 (4,428,155)
	48 not 51 (9136)
	30 or 52 (9895)
	(201204$ or 201205$ or 201507$ or 201508$ or 201509$ or 201510$ or 201511$ or 201512$ or 2016$).ed,ep,dc. (1,388,357)
	53 and 54 (1160)

Appendix 5 Studies excluded at full text

*T*able 55 presents the studies that were excluded from the DTA review at full-text stage. Although there was consensus between reviewers that these studies should be excluded, there were often multiple reasons for exclusion and consensus was not always obtained. One reviewer's reason for the exclusion of each study is presented in the table. Reasons for exclusion were:

- not a DTA study
- target condition not AF (this included studies from which DTA data for AF could not be extracted, for example when AF and atrial flutter were combined)
- information not per person (unit of analysis not the person, for example the unit of analysis was the reading or the segment of a reading)
- population – people with a diagnosis of AF who had had curative treatment such as ablation or cardioversion, people with pacemakers/paced rhythms, stroke inpatients and outpatients, cardiology inpatients and outpatients, anticoagulant outpatients, patients in intensive care and studies with < 40 participants
- index test – invasive, not possible in primary care or does not detect arrhythmia
- reference standard – when the reference standard was not a 12-lead ECG interpreted by a cardiologist.

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion

Study	Reason for exclusion
Abe 1995[175]	Index test
Albilali 2014[176]	Not a DTA study
Al-Fahoum 2013[177]	Information not per person
Alis 2009[178]	Population
Amir 2011[179]	Information not per person
Andrikopoulos 2000[180]	Index test
Anh 2006[181]	Not a DTA study
Aras 2005[182]	Index test
Asl 2008[183]	Information not per person
Aytemir 1999[184]	Not a DTA study
Aytemir 2000[185]	Index test
Babaeizadeh 2009[186]	Information not per person
Bae 2011[187]	Information not per person
Bae 2012[188]	Information not per person
Bansil 2004[189]	Population
Barrett 2013[190]	Not a DTA study
Barrett 2014[191]	Population
Barthéleémy 2003[192]	Population
Benavente 2013[193]	Not a DTA study
Bernard 1981[194]	Population
Boudaoud 2003[195]	Index test

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion (*continued*)

Study	Reason for exclusion
Bourdillon 1978[196]	Population
Boyle 2013[197]	Reference standard
Budeus 2003[198]	Index test
Budeus 2005[199]	Index test
Caldwell 2012[200]	Population
Chao 2011[201]	Index test
Chee 2010[202]	Not a DTA study
Choi 2010[203]	Information not per person
Clavier 2002[204]	Index test
Cole-Haskayne 2014[205]	Not a DTA study
Coutts 2014[206]	Not a DTA study
Cubanski 1994[207]	Information not per person
Czaplik 2010[208]	Target condition not AF
Dash 2009[209]	Information not per person
Dash 2009[210]	Population
Dash 2011[211]	Population
Davidenko 2007[212]	Information not per person
Davy 2014[213]	Reference standard
Deelawar 2013[214]	Population
Dilaveris 1998[215]	Index test
Dilaveris 1999[216]	Index test
Dirschedl 1996[217]	Population
Dobbels 2014[218]	Population
Dogan 2012[219]	Population
Doliwa 2009[220]	Population
Doliwa Sobocinski 2012[221]	Population
Dotan 2011[222]	Population
Douen 2008[223]	Population
Du 2014[224]	Information not per person
Dukes 2014[225]	Reference standard
Duverney 2002[226]	Index test
Elgendi 2008[227]	Information not per person
Elijovich 2008[228]	Not a DTA study
Engdahl 2012[229]	Not a DTA study
Engdahl 2013[22]	Not a DTA study
Feruglio 1975[230]	Target condition not AF
Filos 2011[231]	Not a DTA study
Finucane 2013[232]	Population

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion (*continued*)

Study	Reason for exclusion
Foley 2011[233]	Population
Franczuk 2004[234]	Index test
Fukunami 1991[235]	Index test
Fukunami 1992[236]	Not a DTA study
Gaillard 2008[237]	Population
Ghrooda 2014[238]	Not a DTA study
Gialafos 1999[239]	Index test
Gomis 2009[240]	Population
Goricke 2011[241]	Population
Gradl 2012[242]	Target condition not AF
Graja 2005[243]	Index test
Grond 2011[244]	Population
Grond 2013[245]	Population
Guidera 1993[246]	Population
Güler 2005[247]	Information not per person
Gumbinger 2010[248]	Population
Gumbinger 2012[249]	Population
Gunalp 2006[250]	Population
Haberman 2014[251]	Target condition not AF
Haberman 2015[252]	Target condition not AF
Haeberlin 2012[253]	Target condition not AF
Haeberlin 2014[254]	Information not per person
Hakacova 2012[255]	Target condition not AF
Hallioglu 2004[256]	Target condition not AF
Harrington 2013[257]	Target condition not AF
Helfenbein 2010[258]	Population
Hendrikx 2012[259]	Population
Hendrikx 2014[260]	Population
Hickey 2004[261]	Reference standard
Higgins 2014[262]	Not a DTA study
Hiraki 1998[263]	Index test
Hoffmann 2010[264]	Information not per person
Hong-Wei 2009[265]	Information not per person
Horstmann 2013[266]	Population
Hoshino 2015[267]	Population
Howlett 2014[268]	Population
Howlett 2014[269]	Not a DTA study

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion (*continued*)

Study	Reason for exclusion
Hsieh 2005[270]	Information not per person
Inoue 2001[271]	Population
Jang 2011[272]	Information not per person
Jeon 2014[273]	Information not per person
Jiang 2012[274]	Information not per person
Joseph 1999[275]	Target condition not AF
Kallmünzer 2012[276]	Population
Kallmünzer 2014[277]	Population
Kandel 2012[278]	Population
Kar 2009[279]	Population
Karapinar 2009[280]	Index test
Kessler 1993[281]	Not a DTA study
Kikillus 2007[282]	Information not per person
Kikillus 2008[283]	Information not per person
Kim 2008[284]	Information not per person
Kim 2013[285]	Not a DTA study
Kinlay 1996[286]	Target condition not AF
Klein 1997[287]	Population
Köse 2003[288]	Index test
Koskinen 2005[289]	Not a DTA study
Kostka 2007[290]	Population
Kostka 2008[291]	Population
Kostka 2008[292]	Population
Kostka 2008[293]	Population
Kostka 2010[294]	Population
Kostka 2011[295]	Information not per person
Krasteva 2007[296]	Target condition not AF
Krummen 2010[112]	Population
Lagido 2014[297]	Population
Lalani 2013[298]	Population
Langley 2001[299]	Index test
Lau 2012[300]	Population
Lavallee 2013[301]	Not a DTA study
Lazzaro 2010[302]	Population
Lazzaro 2012[303]	Population
Lee 2011[304]	Information not per person
Lee 2013[305]	Information not per person
Lee 2013[306]	Population

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion (*continued*)

Study	Reason for exclusion
Lelakowska-Piela 2013[307]	Not a DTA study
Lepage 2001[308]	Index test
Lever 2009[309]	Not a DTA study
Lewalter 2007[310]	Population
Lim 2008[311]	Not a DTA study
Linker 2013[312]	Population
Liu 2010[313]	Information not per person
Locati 2011[314]	Population
Logan 2005[315]	Information not per person
Lowres 2014[21]	Reference standard
Lu 2010[316]	Information not per person
Madsen 2009[317]	Population
Mahagne 2009[318]	Population
Mahagne 2011[319]	Population
Maier 2001[320]	Index test
Mairesse 2013[321]	Not a DTA study
Maitas 2012[322]	Information not per person
Maitas 2012[323]	Information not per person
Mäkijaärvi 1993[324]	Population
Makowska 2000[325]	Target condition not AF
Manina 2014[326]	Population
Martínez 2012[327]	Index test
Martinez-Sanchez 2011[328]	Population
Martinez-Sanchez 2012[329]	Population
Martis 2014[330]	Information not per person
Maslowsky 2012[331]	Index test
McCarthy 2014[332]	Not a DTA study
Miller 2012[333]	Population
Miller 2014[334]	Not a DTA study
Mohebbi 2008[335]	Information not per person
Mohebbi 2012[336]	Information not per person
Montereggi 1996[337]	Index test
Moreira 2006[338]	Index test
Morgan 2002[82]	Reference standard
Mueller 2006[339]	Information not per person
Müller 2009[340]	Reference standard
Murgatroyd 1995[341]	Information not per person

continued

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion (*continued*)

Study	Reason for exclusion
Opolski 1995[342]	Index test
Ozdemir 2004[343]	Index test
Ozer 2000[344]	Index test
Park 2009[345]	Information not per person
Pastor-Pérez 2010[346]	Population
Peer 2003[347]	Not a DTA study
Pellise 2011[348]	Not a DTA study
Poon 2005[349]	Information not per person
Portet 2008[350]	Target condition not AF
Potpara 2014[351]	Not a DTA study
Pryor 1972[352]	Target condition not AF
Pusalkar 2012[353]	Population
Quinto Villani 1993[354]	Index test
Rabinstein 2014[355]	Not a DTA study
Reddy 1998[356]	Information not per person
Reifart 1997[357]	Population
Renier 2012[358]	Target condition not AF
Rhys 2012[359]	Population
Rhys 2013[117]	Population
Rincon 2012[360]	Information not per person
Rizikou 2013[361]	Population
Rizos 2010[362]	Population
Rizos 2012[363]	Population
Roche 2002[364]	Population
Ros 2004[365]	Index test
Roten 2012[366]	Information not per person
Samol 2013[367]	Reference standard
Sanak 2014[368]	Population
Sato 2009[369]	Index test
Sawant 2014[370]	Population
Schaefer 2014[371]	Population
Schaer 2003[372]	Not a DTA study
Schreier 2001[373]	Index test
Schuchert 1999[374]	Population
Sezgin 2013[375]	Information not per person
Shafqat 2004[376]	Not a DTA study
Sheldon 2014[377]	Population
Shiyovich 2010[378]	Not a DTA study

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion (*continued*)

Study	Reason for exclusion
Singh 2010[379]	Population
Sobocinski-Doliwa 2010[380]	Population
Stafford 1995[381]	Index test
Stahrenberg 2010[382]	Population
Sudlow 1998[383]	Reference standard
Sugai 2009[384]	Target condition not AF
Suissa 2011[385]	Population
Suissa 2012[386]	Population
Suissa 2012[387]	Population
Suissa 2013[388]	Population
Suissa 2014[389]	Population
Sun 2008[390]	Information not per person
Sutamnartpong 2014[391]	Not a DTA study
Tagawa 2007[392]	Population
Tateno 2001[393]	Information not per person
Temiz 2014[394]	Population
[a]Tieleman 2014[395]	Population/reference standard
Torbey 2013[396]	Population
Tu 2014[397]	Population
Tung 2014[398]	Information not per person
Ustrell 2011[399]	Population
Velthuis 2013[400]	Information not per person
Veltkamp 2011[401]	Population
Veltkamp 2012[402]	Population
Villani 1996[403]	Index test
Villani 1996[404]	Index test
Vinther 2014[405]	Population
Vyas 2014[406]	Information not per person
Wallmann 2007[407]	Population
Wiesel 2004[408]	Information not per person
Wiesel 2007[409]	Population
Wiesel 2009[410]	Population
Wiesel 2010[411]	Reference standard
Wiesel 2013[412]	Reference standard
Wiesel 2014[413]	Population
Wiesel 2014[414]	Population
Winkler 2011[415]	Population

TABLE 55 Articles excluded at full text and one reviewer's reason for exclusion (*continued*)

Study	Reason for exclusion
Wyse 2008[416]	Not a DTA study
Xu 2002[417]	Information not per person
Yaghouby 2010[418]	Information not per person
Yiğit 2003[419]	Not a DTA study
Yung 2010[420]	Population
Zenk 2004[421]	Target condition not AF
Zhou 2014[422]	Information not per person

a This article represents two studies.

Appendix 6　Meta-analysis of diagnostic test accuracy studies

Summary of included data

There were 14 studies with a total of 50 test groups, and all either directly reported the number of true positives, false positives, true negatives and false negatives, or reported enough data for these numbers to be correctly calculated. Assuming that tests are the same no matter the interpreter, there are six groups of tests to be analysed (modified blood pressure monitor, single-lead ECG, two-stage screening strategy, 12-lead ECG, > 1- and < 12-lead ECG and pulse palpation). Assuming that tests are further split by interpreter (denoted 'Test/interpreter' category in *Tables 41–46* and *60–69*), there are 12 categories of test with single-lead ECG split into four categories (automatic/algorithm, cardiologist, GP and nurse) and 12-lead ECG split into three categories (automatic/algorithm, GP and nurse). The final two columns of *Table 2* summarise the number of test groups (arms/observations) and their test classifications and interpreter.

Model fit and model selection

We found that all three slope models (independent, common and exchangeable) would not converge for either categorisation of tests, and so we assumed all SROC curves to be symmetric. Alternative (mathematically equivalent) parameterisations of the slope did not help convergence.

In the no slope models, sensitivity and specificity converged for all tests except '12-lead ECG – GP' in the split by interpreter model. This group had only one data point. Models with and without splitting by interpreter both gave an adequate fit based on the posterior residual deviance, which is similar to the number of data points, 92 (*Table 56*). The DIC was very similar for both models, indicating that there is no evidence that DTA depends strongly on interpreter (see *Table 57*).

Heterogeneity across studies (σ_i in our model notation) and across observations (within studies) (ω_i in our notation) was assessed by comparing their respective standard deviations with the standard deviations of the model parameters Λ and Θ. The results of this assessment are presented in *Table 57* for the tests not split by interpreter analysis and *Table 58* for the split by interpreter analysis. We found evidence of heterogeneity across and within studies as the parameter standard deviations were comparable to the heterogeneity standard deviations.

TABLE 56 Model fit summaries comparing models for the index test defined with or without interpreter, assuming a random-effects model[a]

Summary statistic	Tests not split by interpreter	Tests split by interpreter
Posterior mean residual deviance	91.77	88.22
Effective number of parameters (pD)	57.64	61.22
DIC	149.40	149.45

a Models with a posterior mean residual deviance close to the number of data points ($n = 92$) indicate a good fit.

TABLE 57 Internal model parameters for the model heterogeneity estimate, tests not split by interpreter model[a]

Model parameter	Mean	SD	2.50%	97.50%	Rhat
$\Lambda[1]$	5.94	0.82	4.32	7.60	1.00
$\Lambda[2]$	6.14	0.52	5.21	7.27	1.00
$\Lambda[3]$	6.56	0.80	4.97	8.15	1.00
$\Lambda[4]$	41.70	21.37	10.12	85.39	1.06
$\Lambda[5]$	6.32	0.46	5.41	7.24	1.00
$\Lambda[6]$	7.12	0.97	5.38	9.26	1.00
$\Lambda[7]$	4.09	0.84	2.48	5.84	1.00
$\Theta[1]$	0.35	0.60	−0.83	1.56	1.00
$\Theta[2]$	0.24	0.36	−0.50	0.95	1.00
$\Theta[3]$	−0.27	0.56	−1.35	0.89	1.00
$\Theta[4]$	18.61	10.72	2.72	40.41	1.05
$\Theta[5]$	−0.54	0.33	−1.24	0.09	1.00
$\Theta[6]$	−1.69	0.66	−3.04	−0.42	1.00
$\Theta[7]$	0.60	0.63	−0.64	1.84	1.00
SD.stud[Θ] (σ_1)	0.51	0.28	0.05	1.11	1.04
SD.stud[Λ] (σ_2)	0.66	0.38	0.06	1.53	1.02
SD.obs[Θ] (ω_1)	0.66	0.13	0.43	0.95	1.00
SD.obs[Λ] (ω_2)	0.79	0.27	0.28	1.33	1.00

Rhat, Brooks-Rubin-Gelman convergence diagnostic reported by WinBUGS version 1.4.3; SD, standard deviation; SD.stud, across-study standard deviation; SD.obs, across-observation standard deviation.
a The across-study and across-observation (within-study) heterogeneity standard deviations should be compared with the individual parameter standard deviations.

TABLE 58 Internal model parameters for the model heterogeneity estimate, split interpreter model[a]

	Mean	SD	2.50%	97.50%	Rhat
$\Lambda[1]$	5.90	0.87	4.18	7.63	1.00
$\Lambda[2]$	6.02	0.85	4.41	7.77	1.00
$\Lambda[3]$	6.52	0.85	4.81	8.23	1.00
$\Lambda[4]$	6.11	0.82	4.65	7.93	1.00
$\Lambda[5]$	51.33	31.51	10.79	125.20	2.21
$\Lambda[6]$	6.32	0.53	5.28	7.40	1.00
$\Lambda[7]$	6.48	1.09	4.26	8.65	1.00
$\Lambda[8]$	4.14	0.90	2.41	5.98	1.00
$\Lambda[9]$	5.88	1.25	3.44	8.43	1.00
$\Lambda[10]$	8.38	2.13	4.52	12.88	1.00
$\Lambda[11]$	6.30	1.69	3.27	10.03	1.00

TABLE 58 Internal model parameters for the model heterogeneity estimate, split interpreter model[a] (*continued*)

	Mean	SD	2.50%	97.50%	Rhat
$\Lambda[12]$	55.53	23.13	16.70	101.60	1.29
$\Lambda[13]$	73.80	45.49	20.51	171.70	1.61
$\Theta[1]$	0.35	0.63	−0.89	1.62	1.00
$\Theta[2]$	0.65	0.58	−0.48	1.82	1.00
$\Theta[3]$	−0.26	0.60	−1.42	0.97	1.00
$\Theta[4]$	0.33	0.56	−0.81	1.49	1.00
$\Theta[5]$	23.41	15.78	3.09	60.36	2.23
$\Theta[6]$	−0.85	0.38	−1.63	−0.13	1.00
$\Theta[7]$	−1.38	0.81	−3.00	0.21	1.00
$\Theta[8]$	0.62	0.64	−0.66	1.91	1.00
$\Theta[9]$	0.12	0.85	−1.56	1.84	1.00
$\Theta[10]$	−0.50	1.30	−3.09	2.11	1.00
$\Theta[11]$	1.19	1.04	−0.74	3.37	1.00
$\Theta[12]$	23.14	11.57	3.57	45.48	1.31
$\Theta[13]$	2.38	19.05	−28.17	51.36	1.22
SD.stud[Θ] (σ_1)	0.59	0.32	0.06	1.29	1.00
SD.stud[Λ] (σ_2)	0.76	0.44	0.04	1.74	1.00
SD.obs[Θ] (ω_1)	0.61	0.14	0.36	0.92	1.00
SD.obs[Λ] (ω_2)	0.78	0.28	0.21	1.33	1.02

SD, standard deviation.
a　The across-study and across-observation (within-study) heterogeneity standard deviations should be compared with the individual parameter standard deviations.

A further investigation was to apply the fixed-effects model described in *Appendix 3*, which assumes fixed parameters across studies and across observations within studies. The model fit statistics are presented in *Table 59* and should be compared with those presented in *Table 56*. Although the fixed-effects models have fewer effective parameters (pD), the residual deviance and DIC clearly favour the random-effects models. This confirms our assessment that heterogeneity is present across studies and observations.

TABLE 59 Model fit summaries comparing models for the index test defined with or without interpreter, assuming a fixed-effects model[a]

Summary statistic	Tests not split by interpreter	Tests split by interpreter
Posterior mean residual deviance	672.44	549.35
Effective number of parameters (pD)	13.01	22.54
DIC	685.44	571.88

a　Models with a posterior mean residual deviance close to the number of data points ($n = 92$) indicate a good fit, so fixed-effects models did not fit well.

Appendix 7 Results tables for diagnostic test accuracy subgroup and sensitivity analyses

TABLE 60 Proportion of observations (arms from each study) on each test that are included in each subgroup analysis or at low risk of each category of bias or applicability concerns

Test/interpreter	Cohort study	Systematic or targeted screening	Primary care	RoB patient selection	RoB index test	RoB reference test	Applicability patient selection	Applicability index test	Applicability reference test
Modified blood pressure monitor	1.00	1.00	0.33	0.67	1.00	1.00	0.33	1.00	1.00
Single-lead ECG – automatic/algorithm	0.33	0.67	0.67	0.00	1.00	1.00	0.67	1.00	1.00
Single-lead ECG – nurse	0.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00
Single-lead ECG – GP	0.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00
Single-lead ECG – cardiologist	1.00	0.89	0.89	0.11	0.56	1.00	0.89	0.00	1.00
Two-stage screening strategy	0.14	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00
Photoplethysmography	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00
12-lead ECG – automatic/algorithm	0.89	0.00	0.11	0.11	0.67	0.22	0.11	1.00	0.67
12-lead ECG – nurse	0.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00
12-lead ECG – GP	0.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00
> 1- and < 12-lead ECG – automatic/algorithm	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
> 1- and < 12-lead ECG – cardiologist	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00
Pulse palpation – nurse	0.33	0.67	1.00	0.00	1.00	1.00	1.00	1.00	1.00

RoB, risk of bias.

TABLE 61 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for cohort study subgroup analysis

Test/interpreter	Number of observations restricted to cohort studies	Number of observations excluding cohort studies	DOR restricted to cohort studies (95% CrI)	DOR excluding cohort studies (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	3	0	2.5 (2.08 to 2.68)	– (– to –)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/ algorithm	1	2	2.07 (1.26 to 2.64)	2.47 (1.34 to 2.72)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	0	3	– (– to –)	2.38 (1.08 to 2.72)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	0	1	– (– to –)	2.55 (1.28 to 2.72)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	9	0	2.53 (2.2 to 2.68)	– (– to –)	2.53 (2.23 to 2.67)
Two-stage screening strategy	1	6	2.47 (1.77 to 2.7)	2.54 (1.28 to 2.72)	2.62 (2.43 to 2.7)
Photoplethysmography	1	0	2.38 (1.61 to 2.69)	– (– to –)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/ algorithm	8	1	2.65 (2.57 to 2.7)	2.71 (2.69 to 2.72)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	0	2	– (– to –)	2.21 (1.03 to 2.72)	2.33 (1.62 to 2.67)
12-lead ECG – GP	0	1	– (– to –)	2.55 (1.27 to 2.72)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	2	0	2.67 (2.52 to 2.72)	– (– to –)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	2	0	2.72 (2.72 to 2.72)	– (– to –)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	1	2	2.18 (1.33 to 2.66)	2.1 (1.02 to 2.72)	2.21 (1.64 to 2.58)

TABLE 62 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for systematic or targeted screening studies subgroup analysis

Test/interpreter	Number of observations restricted to systematic or targeted screening studies	Number of observations excluding systematic or targeted screening studies	DOR restricted to systematic or targeted screening studies (95% CrI)	DOR excluding systematic or targeted screening studies (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	3	0	2.46 (1.7 to 2.7)	– (– to –)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/ algorithm	2	1	2.3 (1.42 to 2.69)	2.56 (1.81 to 2.71)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	0	2.47 (1.47 to 2.71)	– (– to –)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	0	2.61 (1.89 to 2.72)	– (– to –)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	8	1	2.42 (1.39 to 2.71)	2.52 (1.68 to 2.71)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	0	2.6 (2.13 to 2.71)	– (– to –)	2.62 (2.43 to 2.7)
Photoplethysmography	0	1	– (– to –)	2.36 (1.36 to 2.7)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/ algorithm	0	9	– (– to –)	2.67 (2.58 to 2.71)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	0	2.29 (1.21 to 2.71)	– (– to –)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	0	2.61 (1.89 to 2.72)	– (– to –)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	0	2	– (– to –)	2.65 (2.31 to 2.72)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	0	2	– (– to –)	2.72 (2.7 to 2.72)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	2	1	2.15 (1.14 to 2.7)	2.16 (1.19 to 2.69)	2.21 (1.64 to 2.58)

TABLE 63 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for primary care studies subgroup analysis

Test/interpreter	Number of observations restricted to primary care studies	Number of observations excluding primary care studies	DOR restricted to primary care studies (95% CrI)	DOR excluding primary care studies (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	1	2	2.31 (1.2 to 2.71)	2.46 (1.4 to 2.71)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	2	1	2.29 (1.38 to 2.69)	2.53 (1.53 to 2.72)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	0	2.46 (1.4 to 2.71)	– (– to –)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	0	2.61 (1.84 to 2.72)	– (– to –)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	8	1	2.41 (1.33 to 2.71)	2.5 (1.46 to 2.72)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	0	2.59 (2.08 to 2.71)	– (– to –)	2.62 (2.43 to 2.7)
Photoplethysmography	0	1	– (– to –)	2.33 (1.21 to 2.71)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	1	8	2.64 (2.1 to 2.72)	2.66 (2.51 to 2.71)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	0	2.28 (1.18 to 2.71)	– (– to –)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	0	2.61 (1.84 to 2.72)	– (– to –)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	0	2	– (– to –)	2.63 (2.07 to 2.72)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	0	2	– (– to –)	2.72 (2.72 to 2.72)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	3	0	2.17 (1.28 to 2.67)	– (– to –)	2.21 (1.64 to 2.58)

TABLE 64 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for risk of bias in relation to patient selection sensitivity analysis

Test/interpreter	Number of observations restricted to low RoB patient selection	Number of observations excluding low RoB patient selection	DOR restricted to low RoB patient selection (95% CrI)	DOR excluding low RoB patient selection (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	2	1	2.29 (1.01 to 2.72)	2.35 (1.47 to 2.7)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	0	3	– (– to –)	2.45 (2 to 2.67)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	0	3	– (– to –)	2.49 (1.82 to 2.7)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	0	1	– (– to –)	2.63 (2.18 to 2.72)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	1	8	2.31 (1.01 to 2.72)	2.45 (1.77 to 2.69)	2.53 (2.23 to 2.67)
Two-stage screening strategy	0	7	– (– to –)	2.61 (2.34 to 2.7)	2.62 (2.43 to 2.7)
Photoplethysmography	1	0	2.16 (1 to 2.72)	– (– to –)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	1	8	2.46 (1.03 to 2.72)	2.66 (2.57 to 2.7)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	0	2	– (– to –)	2.31 (1.46 to 2.69)	2.33 (1.62 to 2.67)
12-lead ECG – GP	0	1	– (– to –)	2.64 (2.21 to 2.72)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	2	0	2.48 (1.06 to 2.72)	– (– to –)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardio	0	2	– (– to –)	2.72 (2.71 to 2.72)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	0	3	– (– to –)	2.2 (1.52 to 2.62)	2.21 (1.64 to 2.58)

RoB, risk of bias.

TABLE 65 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for RoB in relation to index test sensitivity analysis

Test/interpreter	Number of observations restricted to low RoB index test	Number of observations excluding low RoB index test	DOR restricted to low RoB index test (95% CrI)	DOR excluding low RoB index test (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	3	0	2.51 (2.19 to 2.66)	– (– to –)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	3	0	2.46 (2.14 to 2.64)	– (– to –)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	0	2.52 (2.07 to 2.69)	– (– to –)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	0	2.66 (2.38 to 2.72)	– (– to –)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	5	4	2.51 (2.21 to 2.66)	2.3 (1.02 to 2.72)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	0	2.62 (2.45 to 2.7)	– (– to –)	2.62 (2.43 to 2.7)
Photoplethysmography	0	1	– (– to –)	2.19 (1.01 to 2.72)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	6	3	2.69 (2.64 to 2.71)	2.25 (1.01 to 2.72)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	0	2.34 (1.71 to 2.66)	– (– to –)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	0	2.65 (2.35 to 2.72)	– (– to –)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	2	0	2.68 (2.57 to 2.71)	– (– to –)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	2	0	2.72 (2.72 to 2.72)	– (– to –)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	3	0	2.21 (1.7 to 2.56)	– (– to –)	2.21 (1.64 to 2.58)

RoB, risk of bias.

TABLE 66 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for RoB in relation to reference test selection sensitivity analysis

Test/interpreter	Number of observations restricted to low RoB reference test	Number of observations excluding low RoB reference test	DOR restricted to low RoB reference test (95% CrI)	DOR excluding low RoB reference test (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	3	0	2.5 (2.1 to 2.68)	– (– to –)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	3	0	2.46 (2.06 to 2.66)	– (– to –)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	0	2.51 (1.97 to 2.7)	– (– to –)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	0	2.64 (2.26 to 2.72)	– (– to –)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	9	0	2.53 (2.22 to 2.68)	– (– to –)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	0	2.63 (2.43 to 2.7)	– (– to –)	2.62 (2.43 to 2.7)
Photoplethysmography	1	0	2.38 (1.59 to 2.69)	– (– to –)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	2	7	2.69 (2.62 to 2.72)	2.63 (2.43 to 2.71)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	0	2.33 (1.58 to 2.68)	– (– to –)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	0	2.64 (2.26 to 2.72)	– (– to –)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	0	2	– (– to –)	2.65 (2.21 to 2.72)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	2	0	2.72 (2.72 to 2.72)	– (– to –)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	3	0	2.21 (1.59 to 2.6)	– (– to –)	2.21 (1.64 to 2.58)

RoB, risk of bias.

TABLE 67 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for studies with a low level of concern about patient selection applicability sensitivity analysis

Test/interpreter	Number of observations restricted to low applicability of patient selection concerns	Number of observations excluding low applicability of patient selection concerns	DOR restricted to low applicability of patient selection concerns (95% CrI)	DOR excluding low applicability of patient selection concerns (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	1	2	2.31 (1.2 to 2.71)	2.46 (1.4 to 2.71)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	2	1	2.29 (1.38 to 2.69)	2.53 (1.53 to 2.72)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	0	2.46 (1.4 to 2.71)	– (– to –)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	0	2.61 (1.84 to 2.72)	– (– to –)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	8	1	2.41 (1.33 to 2.71)	2.5 (1.46 to 2.72)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	0	2.59 (2.08 to 2.71)	– (– to –)	2.62 (2.43 to 2.7)
Photoplethysmography	0	1	– (– to –)	2.33 (1.21 to 2.71)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	1	8	2.64 (2.1 to 2.72)	2.66 (2.51 to 2.71)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	0	2.28 (1.18 to 2.71)	– (– to –)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	0	2.61 (1.84 to 2.72)	– (– to –)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	0	2	– (– to –)	2.63 (2.07 to 2.72)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	0	2	– (– to –)	2.72 (2.72 to 2.72)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	3	0	2.17 (1.28 to 2.67)	– (– to –)	2.21 (1.64 to 2.58)

TABLE 68 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for studies with a low level of concern about index test applicability sensitivity analysis

Test/interpreter	Number of observations restricted to low applicability of index test concerns	Number of observations excluding low applicability of index test concerns	DOR restricted to low applicability of index test concerns (95% CrI)	DOR excluding low applicability of index test concerns (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	3	0	2.5 (2.11 to 2.67)	– (– to –)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	3	0	2.47 (2.12 to 2.65)	– (– to –)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	0	2.48 (1.86 to 2.7)	– (– to –)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	0	2.64 (2.31 to 2.72)	– (– to –)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	0	9	– (– to –)	2.43 (1.25 to 2.72)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	0	2.6 (2.35 to 2.7)	– (– to –)	2.62 (2.43 to 2.7)
Photoplethysmography	1	0	2.4 (1.7 to 2.68)	– (– to –)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	9	0	2.67 (2.62 to 2.7)	– (– to –)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	0	2.33 (1.58 to 2.68)	– (– to –)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	0	2.65 (2.33 to 2.72)	– (– to –)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	2	0	2.67 (2.52 to 2.72)	– (– to –)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardio	0	2	– (– to –)	2.72 (2.72 to 2.72)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	3	0	2.2 (1.62 to 2.58)	– (– to –)	2.21 (1.64 to 2.58)

TABLE 69 Diagnostic test accuracy summaries for different categories of test and interpreter combinations for studies with a low level of concern about reference test applicability sensitivity analysis

Test/interpreter	Number of observations restricted to low applicability of reference test concerns	Number of observations excluding low applicability of reference test concerns	DOR restricted to low applicability of reference test concerns (95% CrI)	DOR excluding low applicability of reference test concerns (95% CrI)	DOR all (95% CrI)
Modified blood pressure monitor	3	0	2.51 (2.2 to 2.67)	– (– to –)	2.5 (2.14 to 2.67)
Single-lead ECG – automatic/algorithm	3	0	2.47 (2.13 to 2.65)	– (– to –)	2.46 (2.1 to 2.65)
Single-lead ECG – nurse	3	0	2.52 (2.09 to 2.69)	– (– to –)	2.52 (2.01 to 2.7)
Single-lead ECG – GP	1	0	2.65 (2.37 to 2.72)	– (– to –)	2.65 (2.31 to 2.72)
Single-lead ECG – cardiologist	9	0	2.53 (2.28 to 2.67)	– (– to –)	2.53 (2.23 to 2.67)
Two-stage screening strategy	7	0	2.63 (2.47 to 2.7)	– (– to –)	2.62 (2.43 to 2.7)
Photoplethysmography	1	0	2.41 (1.74 to 2.68)	– (– to –)	2.39 (1.68 to 2.69)
12-lead ECG – automatic/algorithm	6	3	2.69 (2.64 to 2.71)	2.25 (1.01 to 2.72)	2.67 (2.61 to 2.7)
12-lead ECG – nurse	2	0	2.34 (1.7 to 2.66)	– (– to –)	2.33 (1.62 to 2.67)
12-lead ECG – GP	1	0	2.65 (2.37 to 2.72)	– (– to –)	2.65 (2.32 to 2.72)
> 1- and < 12-lead ECG – automatic/algorithm	2	0	2.68 (2.58 to 2.71)	– (– to –)	2.68 (2.55 to 2.71)
> 1- and < 12-lead ECG – cardiologist	2	0	2.72 (2.72 to 2.72)	– (– to –)	2.72 (2.72 to 2.72)
Pulse palpation – nurse	3	0	2.21 (1.7 to 2.56)	– (– to –)	2.21 (1.64 to 2.58)

Appendix 8 Consolidated Health Economic Evaluation Reporting Standards checklist

TABLE 70 Consolidated Health Economic Evaluation Reporting Standards checklist for economic evaluations identified in the review

CHEERS checklist item		Aronsson 2015[100]	[a]Aronsson 2015[99]	[b]Gordon 2012[118]	Moran 2015[105]	Hobbs 2005[19]	[c]Lord 2013[116]	Lowres 2014[21]	Maeda 2004[96]	[b]Rhys 2013[117]
1	Title identified as economic evaluation	Y	N	N	Y	Y	N	Y	Y	N
2	Structured abstract	Y	N	N	Y	Y	N	N	Y	N
3	Background and objectives	Y	Y	N	Y	Y	Y	Y	Y	N
4	Population characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y
5	Setting and location	Y	N	Y	Y	Y	Y	Y	Y	Y
6	Study perspective	Y	N	N	Y	N	Y	Y	Y	N
7	Comparators	N	Y	N	Y	Y	Y	Y	Y	N
8	Time horizon	Y	N	N	Y	Y	Y	Y	Y	N
9	Discount rate	Y	N	N	Y	Y	Y	Y	Y	N
10	Description of health outcomes and relevance	Y	Y	Y	Y	Y	Y	Y	N	Y
11	Measurement of effectiveness	Y	N	Y	Y	Y	Y	Y	Y	Y
12	Measurement of preference-based outcomes	NA	NA	NA	NA	NA	NA	NA	NA	NA
13	Unit costs and methods reported	N	N	N	Y	Y	Y	Y	Y	Y
14	Currency, date and conversion	Y	N	N	Y	Y	Y	Y	Y	N
15	Choice of model described	Y	Y	N	Y	Y	Y	N	Y	N
16	Model assumptions	Y	N	N	Y	Y	Y	Y	Y	N
17	Analytical methods described	Y	Y	N	Y	Y	Y	N	Y	N
18	Study parameters reported in full	N	N	N	Y	Y	Y	Y	Y	N
19	Incremental costs and outcomes reported	Y	Y	N	Y	Y	Y	Y	Y	N
20	Sensitivity analysis undertaken	Y	N	N	Y	Y	NA	N	Y	N

continued

TABLE 70 Consolidated Health Economic Evaluation Reporting Standards checklist for economic evaluations identified in the review (*continued*)

CHEERS checklist item		Aronsson 2015[100]	[a]Aronsson 2015[99]	[b]Gordon 2012[118]	Moran 2015[105]	Hobbs 2005[19]	[c]Lord 2013[116]	Lowres 2014[21]	Maeda 2004[96]	[b]Rhys 2013[117]
21	Heterogeneity explored	N	N	N	N	N	NA	N	Y	N
22	Findings, limitations and generalisability	Y	N	N	Y	Y	NA	Y	Y	N
23	Funding source	Y	N	N	Y	Y	Y	Y	N	Y
24	Conflicts of interest	Y	N	Y	Y	Y	Y	Y	N	Y
	Score	19/23	7/23	5/23	22/23	21/23	18/23	18/23	20/23	7/23

N, no; NA, not applicable; Y, yes.
a Based on abstract.
b Insufficient information on economic evaluation – secondary to the objectives of the paper; included in the discussion section.
c Guideline economic evaluation developed for NICE using AF screening as a case study.

Appendix 9 Search strategy for the review of the natural history of atrial fibrillation screening

The search strategy for the review of the natural history of AF screening for use in the economic modelling is provided below.

1. *Atrial Fibrillation/ep (1706)
2. *atrial fibrillation/ or Atrial fibrillation/ep (31,285)
3. (atrial fibrillation* or atrium fibrillation* or auricular fibrillation*).tw. (45,795)
4. (af or a-fib).tw. (25,007)
5. Atrial Flutter/ (5240)
6. (atrial flutter* or auricular flutter*).tw. (4909)
7. or/2-6 (63,948)
8. Mass Screening/ (85,782)
9. disease progression/ (114,587)
10. incidence/ or prevalence/ (391,325)
11. ((AF or atrial fibrillation) adj6 (prevalence or incidence or screen*)).tw. (3318)
12. or/8-10 (574,026)
13. or/8-11 (576,168)
14. 1 and 12 (963)
15. 7 and 11 (3318)
16. 14 or 15 (3826)
17. letter/ (898,830)
18. editorial/ (391,744)
19. news/ (173,688)
20. exp historical article/ (363,693)
21. Anecdotes as topic/ (4684)
22. comment/ (647,935)
23. case report/ (1,760,067)
24. (letter or comment$).ti. (107,532)
25. animals/ not humans/ (4,142,787)
26. exp Animals, Laboratory/ (753,015)
27. exp Animal Experimentation/ (7739)
28. exp Models, Animal/ (448,998)
29. exp rodentia/ (2,813,026)
30. (rat or rats or mouse or mice).ti. (1,171,895)
31. or/17-30 (8,328,614)
32. 16 not 31 (3524)
33. epidemiologic studies/ (6952)
34. exp case control studies/ (748,592)
35. exp cohort studies/ (1,484,230)
36. cross-sectional studies/ (204,195)
37. (case control or cohort analys$).ti,ab. (92,391)
38. (cohort adj (study or studies)).ti,ab. (107,645)
39. ((follow up or prospective or retrospective or observational) adj (study or studies)).ti,ab. (318,257)
40. (Longitudinal or cross sectional).ti,ab. (351,536)
41. or/33-40 (2106,518)
42. 32 and 41 (1757)
43. (child* or paediatric or pediatric or neonat* or newborn* or infant*).ti. (978,749)

44. 42 not 43 (1754)
45. limit 44 to english language (1641)
46. limit 45 to yr="2000 –Current" (1469)
47. limit 45 to yr="2005-current" (1257)

Appendix 10　Cost-effectiveness planes: base-case economic evaluation

*F*igures 29–34 show the cost-effectiveness planes for each age group in the base-case economic evaluation. Although there is uncertainty in the incremental costs and QALYs, systematic population screening has both higher costs and higher QALYs than systematic opportunistic screening in all cases.
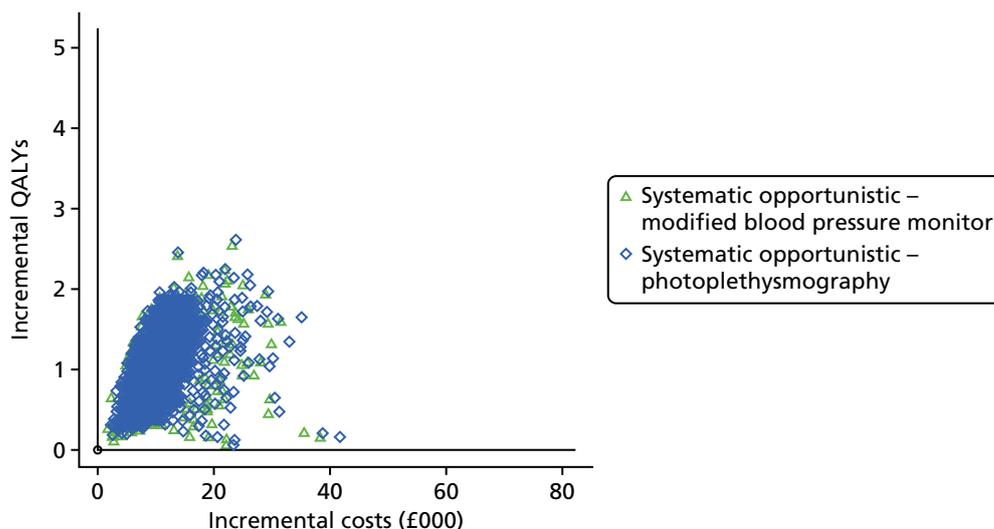
FIGURE 29 Cost-effectiveness plane for a single screen at age 55 years. Screening strategies that have a < 5% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.
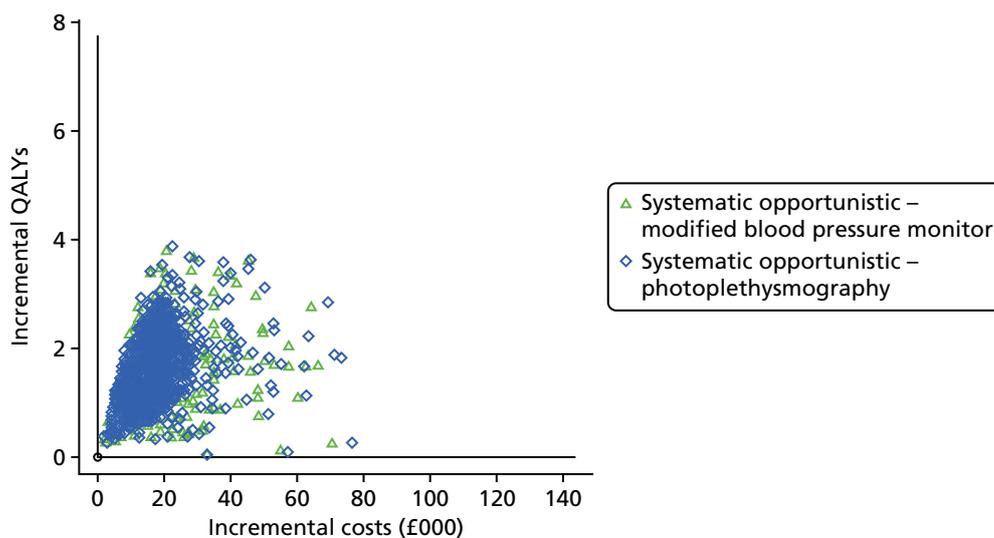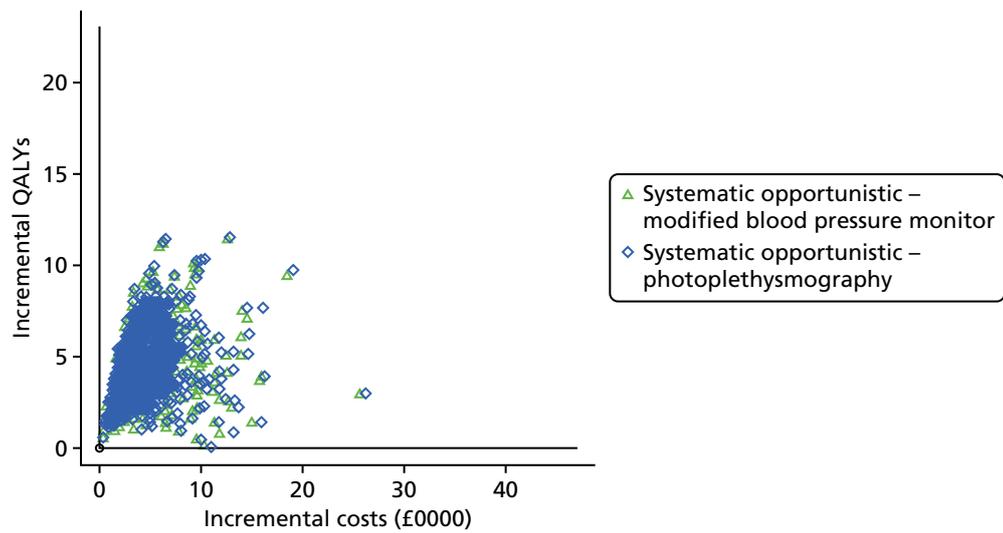

FIGURE 30 Cost-effectiveness plane for a single screen at age 60 years. Screening strategies that have a < 5% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.
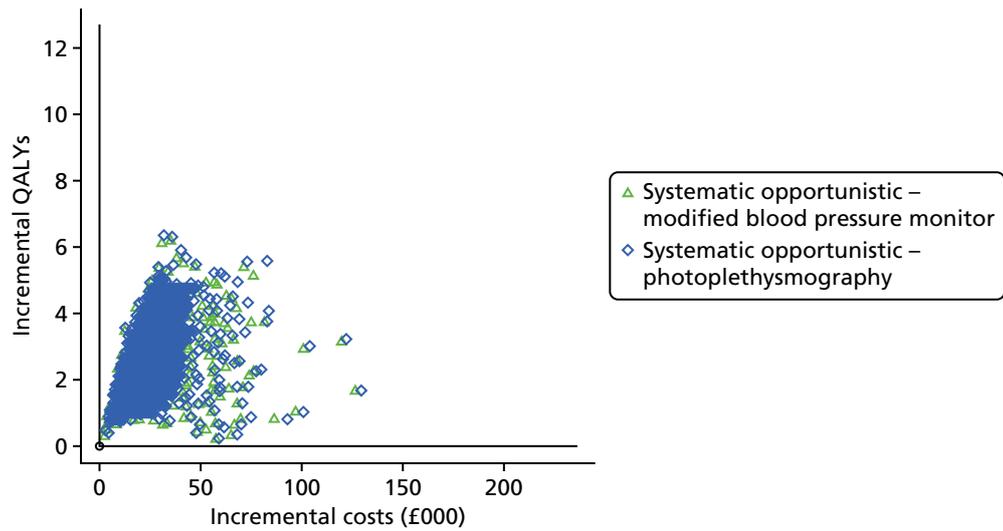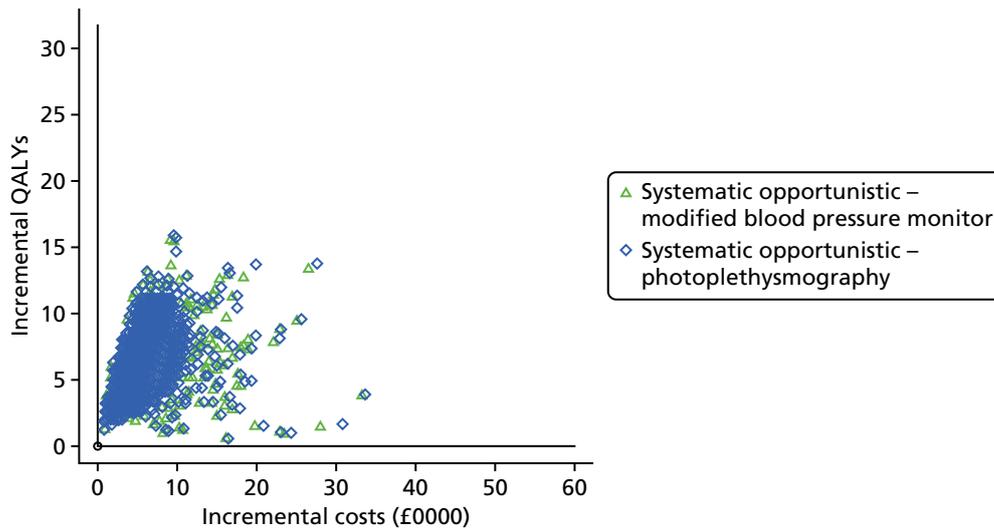
FIGURE 31 Cost-effectiveness plane for a single screen at age 65 years. Screening strategies that have a < 5% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

FIGURE 32 Cost-effectiveness plane for a single screen at age 70 years. Screening strategies that have a < 5% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.
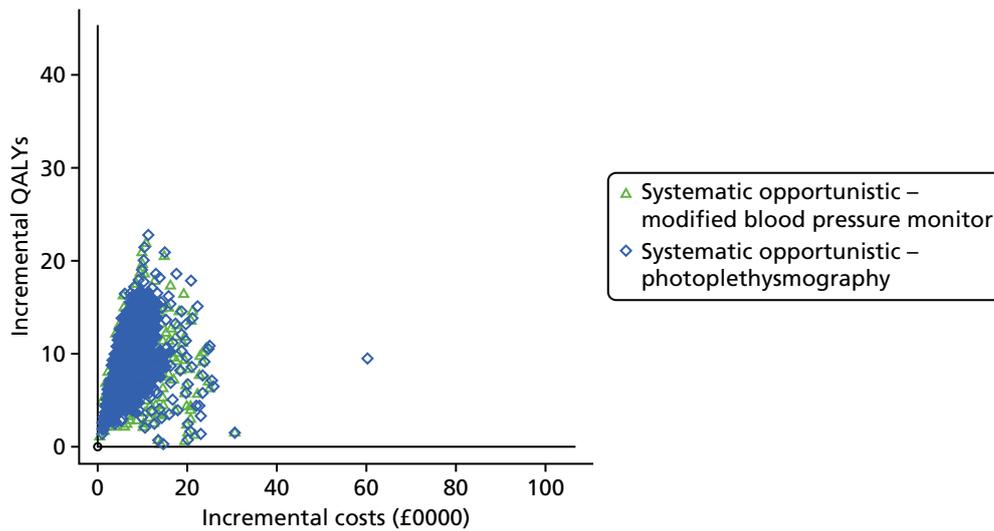
FIGURE 33 Cost-effectiveness plane for a single screen at age 75 years. Screening strategies that have a < 5% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

FIGURE 34 Cost-effectiveness plane for a single screen at age 80 years. Screening strategies that have a < 5% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Appendix 11 Results of the economic evaluation sensitivity analyses

Screening test

Using a modified blood pressure monitor as the screening test

The results for the across-age analysis are presented in *Table 71* and the CEACs are presented in *Figure 35*. We see almost no change in the estimated INBs for screening at each age. The conclusion that screening at age 80 years had the highest probability of being cost-effective and the highest expected INB was also unchanged.

TABLE 71 Sensitivity analysis using opportunistic screening with a modified blood pressure monitor as the screening test: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages

| Screening age (years)	Population size of age cohort for given ages	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	14,305 (6411 to 30,346)	2147 (1074 to 3652)	28,628 (10,672 to 51,349)	29,849 (11,152 to 53,445)
75	379,913.4	15,275 (6672 to 32,993)	2088 (1052 to 3505)	26,484 (7590 to 47,568)	27,603 (8010 to 49,625)
70	466,605.4	12,631 (5738 to 27,323)	1766 (893 to 3009)	22,697 (7509 to 41,679)	23,593 (7570 to 43,372)
65	633,658.8	10,930 (5247 to 21,845)	1423 (718 to 2433)	17,532 (5904 to 32,195)	18,137 (5959 to 33,476)
60	620,292.8	6818 (3316 to 14,226)	813 (405 to 1389)	9449 (1763 to 18,066)	9700 (1721 to 18,876)
55	675,855.0	5037 (2623 to 9870)	574 (281 to 992)	6443 (1598 to 12,496)	6556 (1240 to 12,932)
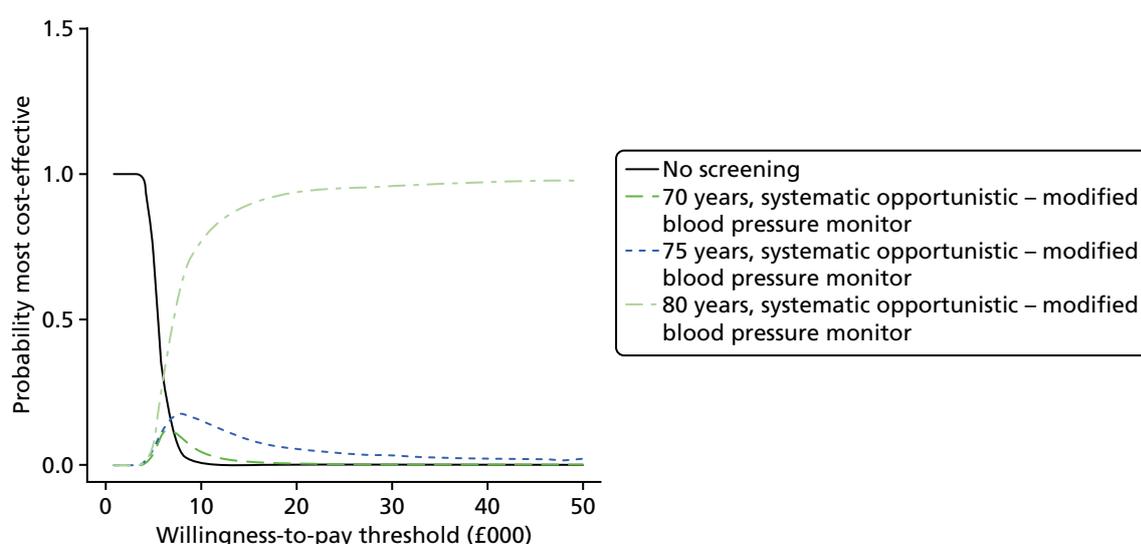
FIGURE 35 Cost-effectiveness acceptability curves comparing different ages for a single one-off screen in a given age cohort, with a modified blood pressure monitor as the screening test. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Using pulse palpation interpreted by a nurse as the screening test

The results for the across-age analysis are presented in *Table 72* and the CEACs are presented in *Figure 36*. Although the estimated INBs for screening are slightly lower than in the base case at each age, the same patterns with age are observed. The conclusion that screening at age 80 years had the highest probability of being cost-effective and the highest expected INB was unchanged.

TABLE 72 Sensitivity analysis using opportunistic screening with pulse palpation by a nurse as the screening test: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages

| Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	13,942 (6204 to 29,584)	2054 (1014 to 3527)	27,129 (9749 to 49,368)	29,849 (11,152 to 53,445)
75	379,913.4	14,996 (6654 to 32,434)	1998 (986 to 3424)	24,962 (6565 to 45,866)	27,603 (8010 to 49,625)
70	466,605.4	12,564 (5817 to 27,064)	1690 (833 to 2896)	21,245 (6475 to 39,661)	23,593 (7570 to 43,372)
65	633,658.8	11,110 (5449 to 21,973)	1362 (677 to 2343)	16,132 (4730 to 30,512)	18,137 (5959 to 33,476)
60	620,292.8	7154 (3577 to 14,185)	777 (378 to 1359)	8393 (990 to 16,949)	9700 (1721 to 18,876)
55	675,855.0	5491 (2928 to 10,085)	550 (267 to 971)	5507 (743 to 11,482)	6556 (1240 to 12,932)
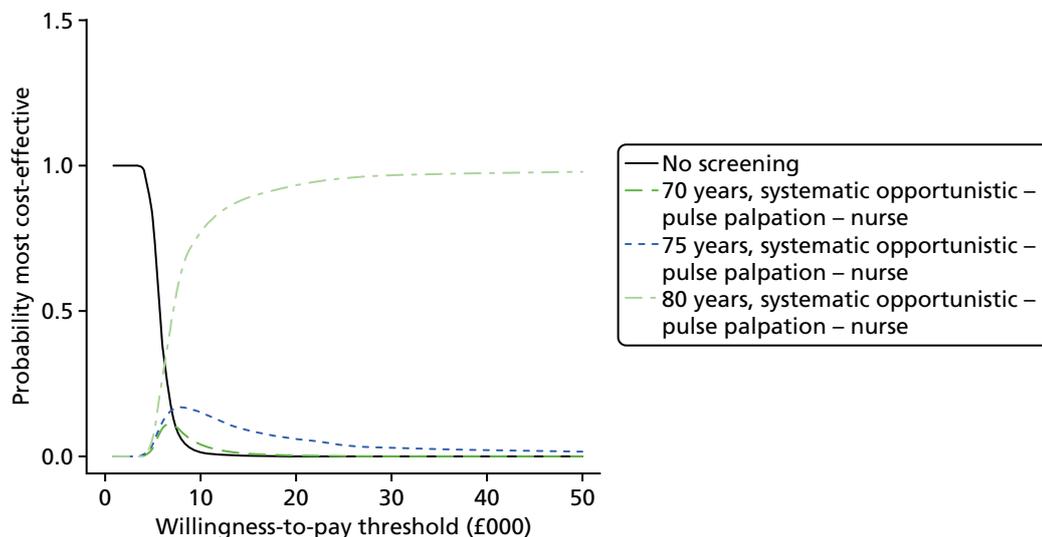
FIGURE 36 Cost-effectiveness acceptability curves comparing different ages for a single one-off screen in a given age cohort, with pulse palpation by a nurse as the screening test. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Uptake of an electrocardiogram in those with a positive screening result

In a sensitivity analysis, we assumed that only 72.5% of those with a positive screening test result agree to have an ECG, compared with 100% in the base case. The results of this sensitivity analysis are presented in *Table 73*; the CEACs are presented in *Figure 37*.

TABLE 73 Sensitivity analysis assuming a lower uptake of an ECG in screen-positive individuals: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages, all under opportunistic screening with photoplethysmography

| Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age-cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	11,045 (5032 to 23,014)	1628 (821 to 2740)	21,513 (7832 to 38,566)	29,849 (11,152 to 53,445)
75	379,913.4	11,868 (5347 to 25,258)	1583 (798 to 2679)	19,795 (5509 to 35,883)	27,603 (8010 to 49,625)
70	466,605.4	9929 (4664 to 21,201)	1338 (676 to 2269)	16,832 (5310 to 31,042)	23,593 (7570 to 43,372)
65	633,658.8	8772 (4380 to 17,146)	1078 (548 to 1836)	12,792 (3862 to 23,892)	18,137 (5959 to 33,476)
60	620,292.8	5637 (2838 to 11,265)	615 (307 to 1054)	6663 (769 to 13,272)	9700 (1721 to 18,876)
55	675,855.0	4319 (2340 to 8004)	435 (215 to 743)	4382 (420 to 9056)	6556 (1240 to 12,932)
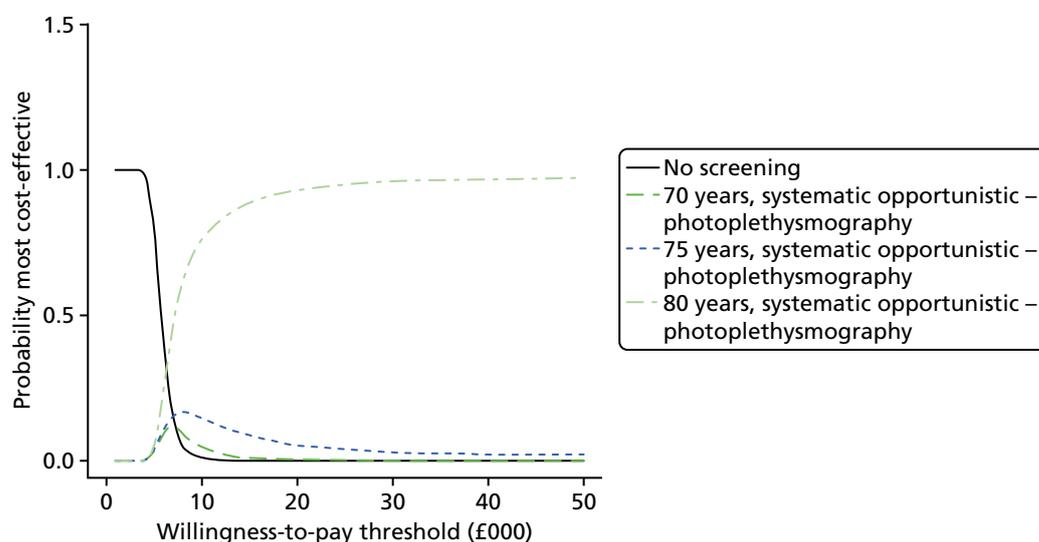
FIGURE 37 Sensitivity analysis assuming a lower uptake of ECGs in screen-positive individuals: CEACs comparing different ages for a single one-off screen in a given age cohort. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Meta-analysis of previous history of stroke in UK screen-detected atrial fibrillation

In sensitivity analysis, we used the results of a random-effects meta-analysis for the proportion of screen-detected AF patients with a previous history of stroke (0.112, 95% CrI 0.072 to 0.170); the value used in the base case, based on the SAFE study,[19] was 0.074 (95% CI 0.032 to 0.116. The results of this sensitivity analysis are presented in *Table 74* and the CEACs are presented in *Figure 38*.

TABLE 74 Sensitivity analysis using the meta-analysis estimate of the previous history of stroke rather than the SAFE study estimate: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages, all under opportunistic screening with photoplethysmography

| Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age-cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	13,868 (6273 to 27,412)	2189 (1045 to 3967)	29,920 (11,869 to 57,838)	29,849 (11,152 to 53,445)
75	379,913.4	15,263 (6493 to 29,999)	2225 (975 to 3970)	29,233 (9653 to 57,245)	27,603 (8010 to 49,625)
70	466,605.4	12,786 (5796 to 26,126)	1809 (817 to 3374)	23,394 (7475 to 46,322)	23,593 (7570 to 43,372)
65	633,658.8	11,153 (5057 to 22,089)	1477 (673 to 2768)	18,393 (5024 to 37,705)	18,137 (5959 to 33,476)
60	620,292.8	6974 (3507 to 13,692)	843 (377 to 1537)	9890 (2617 to 20,142)	9700 (1721 to 18,876)
55	675,855.0	5283 (2612 to 10,228)	570 (258 to 1082)	6114 (784 to 13,820)	6556 (1240 to 12,932)
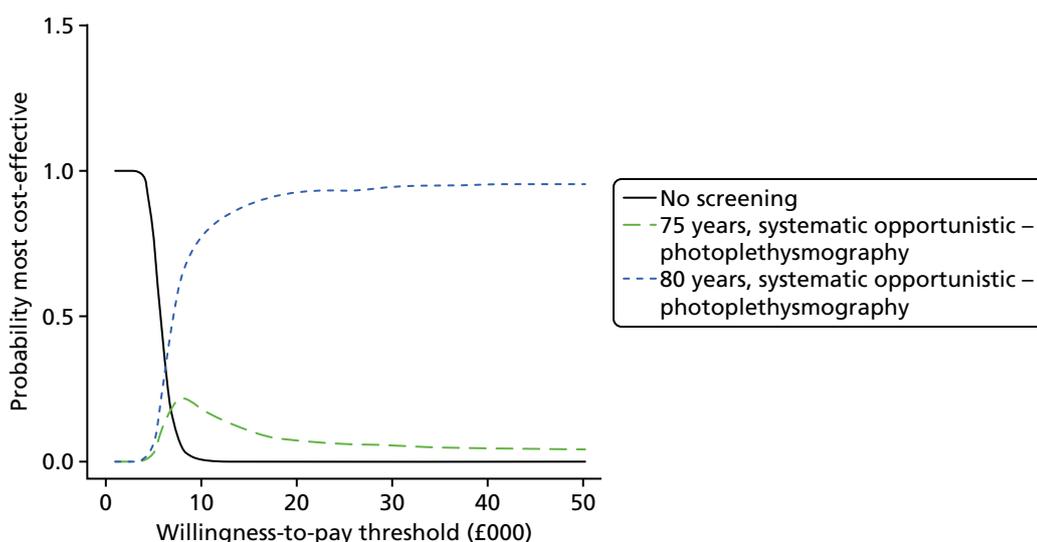
FIGURE 38 Sensitivity analysis using the meta-analysis estimate of the previous history of stroke rather than the SAFE study estimate: CEACs comparing different ages for a single one-off screen in a given age cohort. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Proportion of anticoagulant patients receiving directly acting oral anticoagulants instead of warfarin

We ran sensitivity analyses on the proportion of OAC patients prescribed DOACs instead of warfarin, assuming values of 50% and 100%. The results are presented in *Tables 75* and *76*, respectively, and the CEACs are presented in *Figures 39* and *40*, respectively.

TABLE 75 Sensitivity analysis assuming that 50% of anticoagulant patients take apixaban instead of warfarin: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages, all under opportunistic screening with photoplethysmography

Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age-cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	14,395 (6394 to 30,655)	2244 (1141 to 3792)	30,478 (11,425 to 54,297)	29,849 (11,152 to 53,445)
75	379,913.4	15,450 (6678 to 34,129)	2185 (1100 to 3692)	28,255 (8431 to 50,552)	27,603 (8010 to 49,625)
70	466,605.4	12,833 (5795 to 27,805)	1846 (926 to 3093)	24,085 (7887 to 43,922)	23,593 (7570 to 43,372)
65	633,658.8	11,212 (5366 to 22,474)	1489 (755 to 2523)	18,562 (6199 to 33,984)	18,137 (5959 to 33,476)
60	620,292.8	7065 (3396 to 14,528)	848 (420 to 1468)	9897 (1902 to 19,199)	9700 (1721 to 18,876)
55	675,855.0	5305 (2735 to 10,305)	601 (296 to 1039)	6710 (1294 to 13,424)	6556 (1240 to 12,932)

TABLE 76 Sensitivity analysis assuming that 100% of anticoagulant patients take apixaban instead of warfarin: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages, all under opportunistic screening with photoplethysmography

| Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age-cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	15,646 (7131 to 32,450)	2243 (1137 to 3767)	29,209 (10,679 to 52,239)	29,849 (11,152 to 53,445)
75	379,913.4	16,718 (7478 to 35,834)	2184 (1112 to 3703)	26,955 (7600 to 49,211)	27,603 (8010 to 49,625)
70	466,605.4	13,824 (6389 to 29,241)	1844 (929 to 3125)	23,051 (7304 to 42,443)	23,593 (7570 to 43,372)
65	633,658.8	12,020 (5888 to 23,666)	1488 (754 to 2544)	17,737 (5853 to 32,598)	18,137 (5959 to 33,476)
60	620,292.8	7530 (3678 to 15,152)	850 (419 to 1454)	9462 (1534 to 18,435)	9700 (1721 to 18,876)
55	675,855.0	5600 (2939 to 10,565)	599 (296 to 1028)	6381 (1115 to 12,688)	6556 (1240 to 12,932)
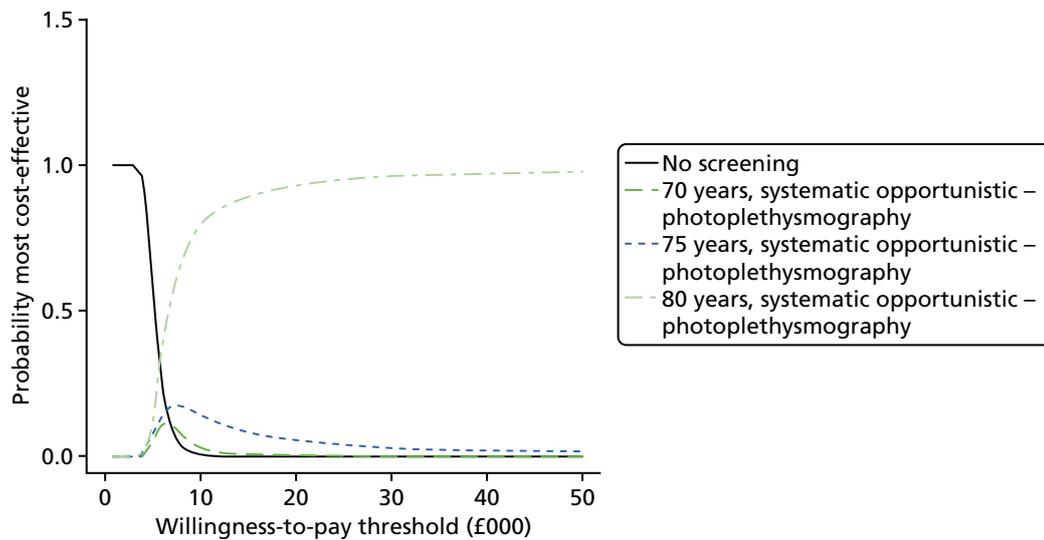
FIGURE 39 Sensitivity analysis assuming that 50% of anticoagulant patients take apixaban instead of warfarin: CEACs comparing different ages for a single one-off screen in a given age cohort. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.
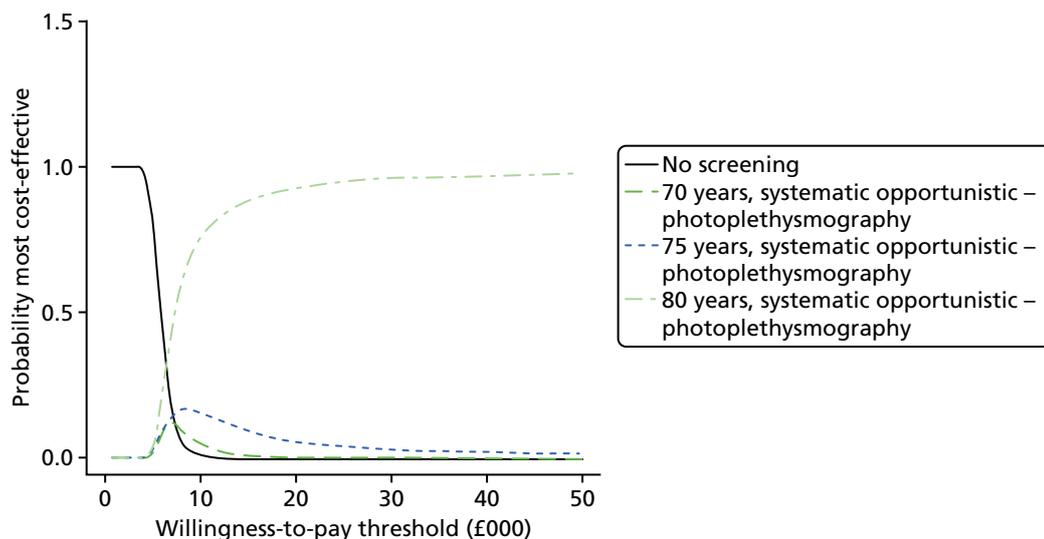

FIGURE 40 Sensitivity analysis assuming that 100% of anticoagulant patients take apixaban instead of warfarin: CEACs comparing different ages for a single one-off screen in a given age cohort. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Uptake of oral anticoagulants

In sensitivity analyses we assumed (1) a reduced OAC uptake of 50% and (2) that all 87% of patients who are not contraindicated or who prefer not to take OACs receive OACs. The results of the sensitivity analysis assuming an OAC uptake rate of 50% are presented in *Table 77*, with the CEACs in *Figure 41*, and the results of the sensitivity analysis assuming an OAC uptake rate of 87% are presented in *Table 78*, with the CEACs in *Figure 42*. Overall, the INB was lower if the OAC uptake rate was lower, but this did not alter our conclusions.

TABLE 77 Sensitivity analysis assuming an OAC uptake rate of 50%. Age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages, all under opportunistic screening with photoplethysmography

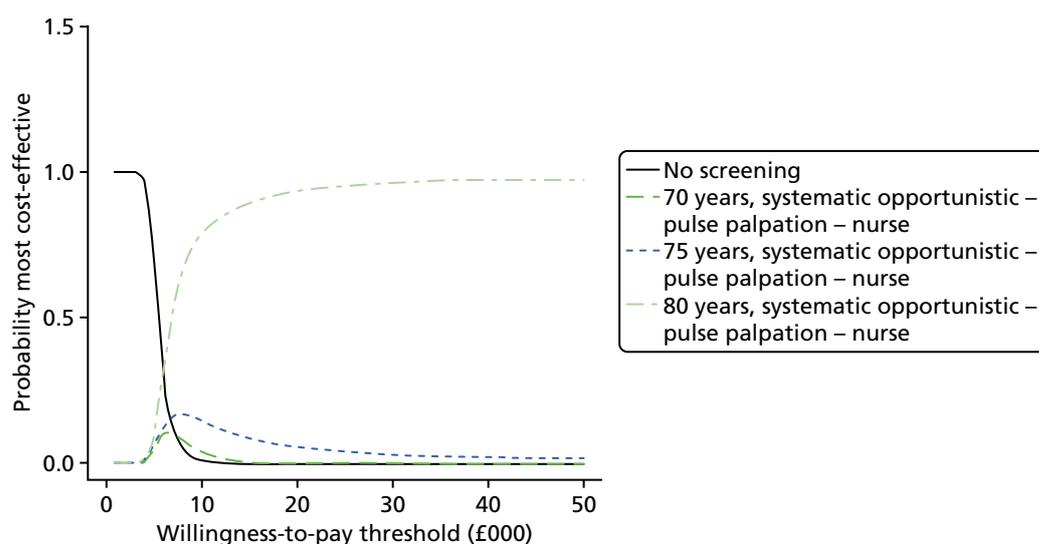| Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age-cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	13,297 (5836 to 28,334)	2044 (1005 to 3506)	27,574 (9839 to 50,323)	29,849 (11,152 to 53,445
75	379,913.4	14,337 (6170 to 31,485)	1989 (966 to 3392)	25,437 (6796 to 46,770)	27,603 (8010 to 49,625)
70	466,605.4	12,048 (5421 to 26,654)	1682 (822 to 2896)	21,601 (6125 to 40,455)	23,593 (7570 to 43,372)
65	633,658.8	10,688 (5141 to 21,546)	1356 (667 to 2344)	16,429 (4713 to 30,622)	18,137 (5959 to 33,476)
60	620,292.8	6925 (3428 to 13,980)	773 (377 to 1334)	8538 (1102 to 17,059)	9700 (1721 to 18,876)
55	675,855.0	5327 (2805 to 9950)	546 (257 to 960)	5584 (660 to 11,700)	6556 (1240 to 12,932)

FIGURE 41 Sensitivity analysis assuming an OAC uptake rate of 50%: CEACs comparing different ages for a single one-off screen in a given age cohort. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

TABLE 78 Sensitivity analysis assuming an OAC uptake rate of 87%: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages, all under opportunistic screening with photoplethysmography

| Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age-cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	15,401 (6987 to 31,998)	2251 (1141 to 3821)	29,628 (10,982 to 53,262)	29,849 (11,152 to 53,445)
75	379,913.4	16,457 (7317 to 35,160)	2190 (1110 to 3703)	27,342 (7846 to 49,353)	27,603 (8010 to 49,625)
70	466,605.4	13,635 (6311 to 29,174)	1852 (927 to 3138)	23,400 (7499 to 43,012)	23,593 (7570 to 43,372)
65	633,658.8	11,852 (5779 to 23,143)	1492 (758 to 2528)	17,987 (5720 to 33,227)	18,137 (5959 to 33,476)
60	620,292.8	7438 (3623 to 15,098)	852 (425 to 1460)	9601 (1607 to 18,690)	9700 (1721 to 18,876)
55	675,855.0	5554 (2866 to 10,523)	603 (297 to 1040)	6510 (1190 to 12,996)	6556 (1240 to 12,932)
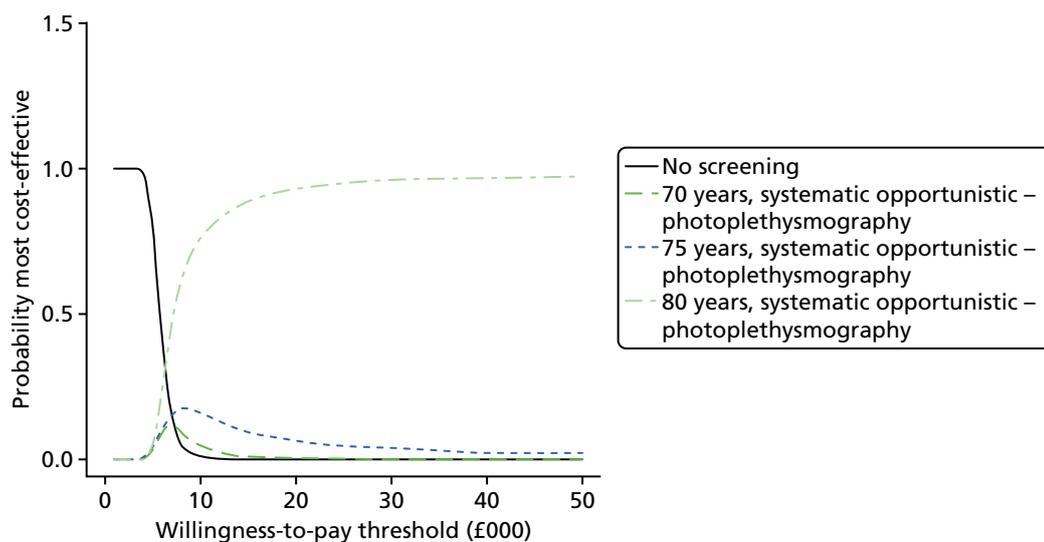
FIGURE 42 Sensitivity analysis assuming an OAC uptake rate of 87%: CEACs comparing different ages for a single one-off screen in a given age cohort. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

Hazard ratios for stroke and mortality risk for routine-detected atrial fibrillation compared with screen-detected atrial fibrillation

In sensitivity analysis, we set the HR for symptomatic AF (representing non-screen-detected AF) compared with asymptomatic AF (representing screen-detected AF) equal to 1.31 (95% CI 1.04 to 1.65), based on the unadjusted analysis reported by Flaker et al.,[35] rather than 1, as assumed in the base case. The results of the sensitivity analysis are shown in Table 79 and the CEACs are shown in Figure 43. Our conclusions were unchanged from the base case.

TABLE 79 Sensitivity analysis using a lower hazard of stroke and mortality in asymptomatic AF compared with symptomatic AF: age cohort population mean incremental costs, incremental QALYs and net benefit for willingness to pay per QALY thresholds of £20,000 for different screening ages, all under opportunistic screening with photoplethysmography

| Screening age (years)	Population size of age cohort	Age cohort population incremental costs (£000) (95% CrI)	Age cohort population incremental QALYs (95% CrI)	Age cohort population INB at £20,000 (£000) (95% CrI)	Age cohort population INB at £20,000 (£000): base case (95% CrI)
80	279,437.0	13,814 (5898 to 28,017)	2138 (983 to 3945)	28,948 (10391 to 56,452)	29,849 (11,152 to 53,445)
75	379,913.4	15,153 (6557 to 31,669)	2238 (997 to 4227)	29,598 (8116 to 58,651)	27,603 (8010 to 49,625)
70	466,605.4	12,623 (5544 to 26,710)	1806 (810 to 3304)	23,498 (8579 to 45,520)	23,593 (7570 to 43,372)
65	633,658.8	10,918 (4890 to 22,992)	1470 (679 to 2680)	18,474 (4819 to 36,787)	18,137 (5959 to 33,476)
60	620,292.8	7021 (3342 to 13,994)	825 (377 to 1561)	9472 (1389 to 20,005)	9700 (1721 to 18,876)
55	675,855.0	5234 (2597 to 10,563)	579 (249 to 1079)	6337 (871 to 13,594)	6556 (1240 to 12,932)
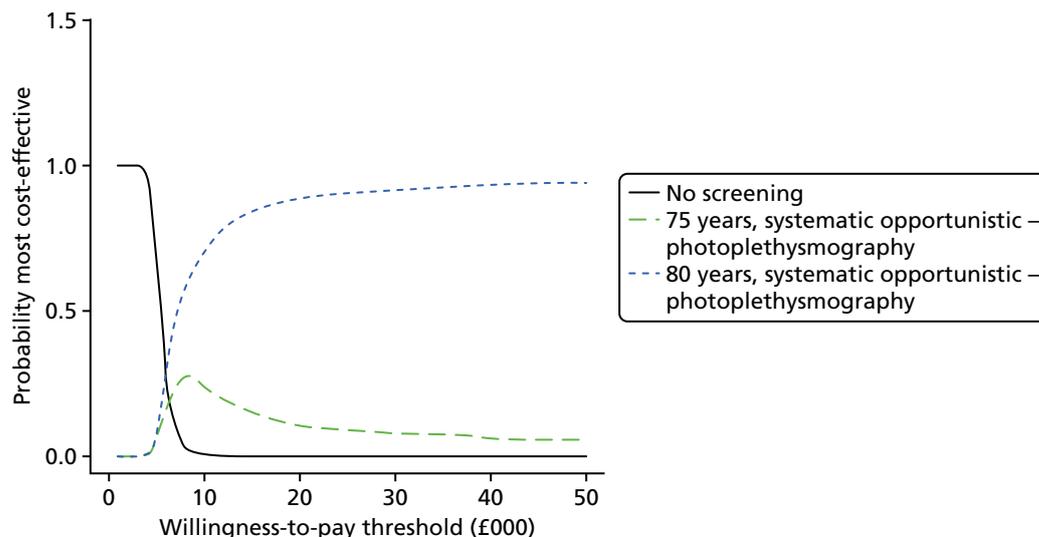
FIGURE 43 Sensitivity analysis using a lower hazard of stroke and mortality in asymptomatic AF compared with symptomatic AF: CEACs comparing different ages for a single one-off screen in a given age cohort. Screening strategies that have a < 10% probability of being cost-effective at any willingness-to-pay threshold are not shown for clarity.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library