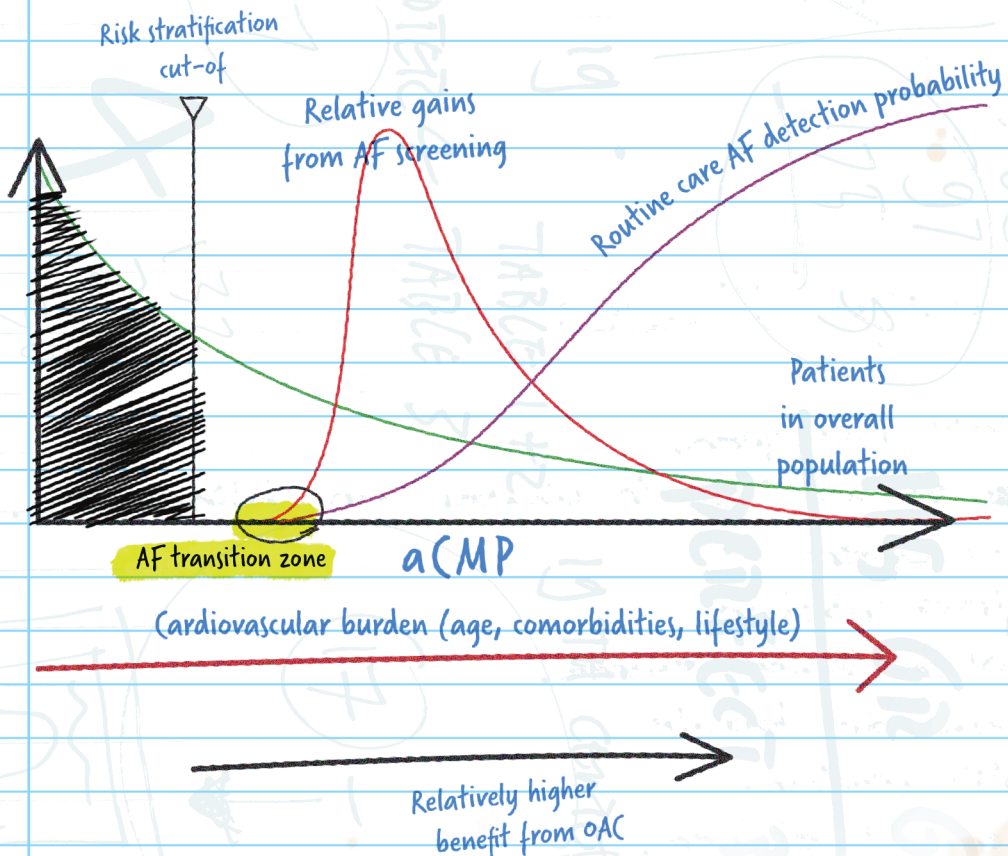
Early identification of atrial fibrillation in primary care and post-stroke patients

Himmelreich, J.C.L.

Publication date
2023
Document Version
Final published version



Early identification of atrial fibrillation in primary care and post-stroke patients

Jelle Himmelreich

Early identification of atrial fibrillation in primary care and post-stroke patients

Jelle Himmelreich

COLOFON
Cover design: ...
Layout and printing: Optima Grafische Communicatie (www.ogc.nl)
ISBN: 978-94-6361-877-9

Het hier beschreven onderzoek werd mede mogelijk gemaakt door steun van ZonMW (80-83910-98-13046), Stichting Stoffels-Hornstra (ond 74) en the European Research Council under the European Union's Horizon 2020 research and innovation programme (648 131).

Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged. Additional financial support for printing of this thesis was kindly provided by Amsterdam University Medical Centres, University of Amsterdam, Department of General Practice, and SBOH , employer of GP trainees.

Early identification of atrial fibrillation in primary care and post-stroke patients

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 14 september 2023, te 16.00 uur

door Jelle Caspar Lorenz Himmelreich
geboren te Heiloo

PROMOTIECOMMISSIE

Promotor:	prof. dr. H.C.P.M. van Weert	AMC-UvA
Copromotores:	dr. J. Coutinho	AMC-UvA
	dr. R.E. Harskamp	AMC-UvA
Overige leden:	prof. dr. Y.B.W.E.M. Roos	AMC-UvA
	dr. J.M. van Es	AMC-UvA
	dr. G.J. Geersing	UMC Utrecht
	prof. dr. P.J.E. Bindels	Erasmus Universiteit Rotterdam
	prof. dr. P.M.M. Bossuyt	AMC-UvA
	prof. dr. J.R. de Groot	AMC-UvA

Faculteit der Geneeskunde

TABLE OF CONTENTS

PART IV: GENERAL DISCUSSION AND SUMMARY

General introduction and outline of the thesis

GENERAL INTRODUCTION AND OUTLINE OF THE THESIS

Atrial fibrillation (AF) is a common arrhythmia, increasing in incidence with age.[1] The lifetime prevalence of AF in western countries is estimated to be around one in every five people with low cardiovascular morbidity, and up to almost two-fifths of people with highest cardiovascular risk.[2] The overall prevalence in the aging European population is expected to double from 8.8 million in 2010 to an estimated 17.9 million in 2060.[3]

On electrocardiogram (ECG), AF is typically distinguished by irregular R-R intervals with no discernible regular P waves.[4] These ECG features arise from the defective electric conduction pathways in the atria that are key to developing the arrhythmia. Pathophysiological factors contributing to the final common electrocardiological outcome that is AF are manifold, and include stretch-induced fibrosis, inflammation, cardiac ischemia and vascular remodelling.[4]

The burden caused by AF is significant and spans across patient-related, healthcare-related and societal factors.[4] Patients with AF can experience lower quality of life due to symptoms relating to AF episodes, especially when having experienced AF-related stroke.[5] Considerable healthcare costs are spent in treating AF symptoms and associated morbidity, as well as in preventing (further) complications from the arrhythmia.[6] And the loss of productivity experience by AF patients as well as the funding required to maintain the current standard of AF treatment and prevention weigh significantly on modern society.[7] Symptoms commonly reported by AF patients are shortness of breath, palpitations, chest discomfort and fatigue.[8] These symptoms can be mitigated with medication to improve heawrt rhythm and rate, or minimally invasive surgery of the area in the heart muscle that is the source of AF episodes ('ablation'). Moreover, such therapies are also effective in preventing later cardiac complications associated with AF such as heart failure. Therefore, in patients who present with aforementioned symptoms it is important to consider AF and to perform electrocardiographic (ECG) investigations.

AF and risk of ischemic stroke

Arguably the clinically most important reason to be aware of AF is the arrhythmia's association with an increased risk of ischemic stroke and transient ischemic attack (TIA) (Figure 1).[4,9] If not treated with adequate stroke prophylaxis through anticoagulation, this risk is thought to be up to five-fold higher than that of persons without AF.[9] It is further estimated that one in four cases of ischemic stroke or TIA are related to AF.[10-13] Although AF is thus not a sine qua non for the development of stroke or TIA,[14] it is a

common and easily identifiable ECG entity, with a well-established body of evidence on how to act in its presence to treat symptoms and to prevent complications.[4,15] Other prominent complications associated with AF include heart failure,[16] myocardial infarction,[16] chronic kidney disease exacerbation,[17] cognitive impairment,[18] falls,[19] sudden cardiac death,[20] and all-cause mortality[16] (Figure 1).[1]

Figure 1. *Common complications associated with atrial fibrillation.*

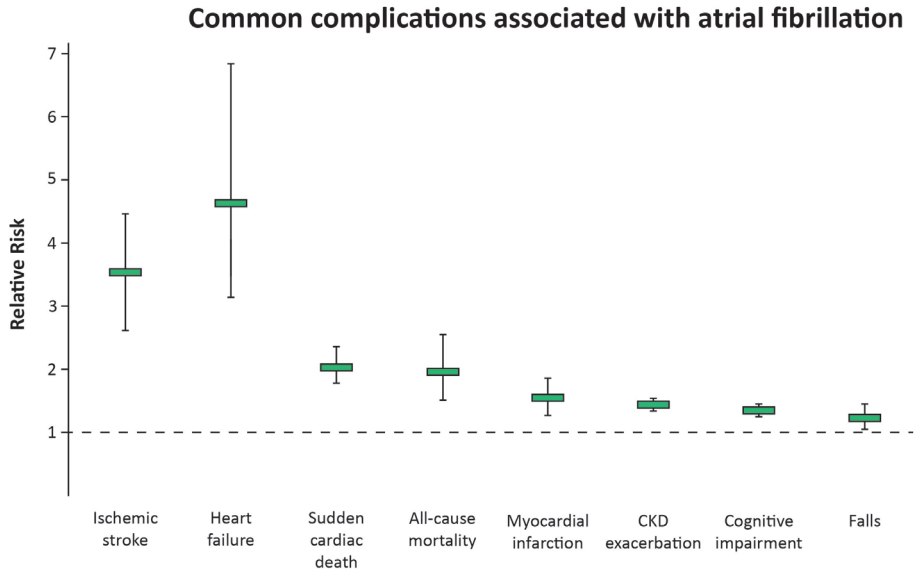


Diagram depicts relative risk in patients with versus without atrial fibrillation for ischemic stroke,[9] heart failure,[16] sudden cardiac death,[20] all-cause mortality,[16] myocardial infarction,[16] chronic kidney disease exacerbation,[17] cognitive impairment,[18] and falls.[19]
CKD, chronic kidney disease.

In assessing the risk of stroke in those with an established AF diagnosis, it is vital to introduce the CHA_2DS_2-VASc risk score.[21] Points are awarded in CHA_2DS_2-VASc for old age (1 point for age 65-74 years; 2 points for age ≥75 years), female sex (1 point), and a history of heart failure (1 point), hypertension (1 point), vascular disease (any of prior myocardial infarction, peripheral artery disease or aortic plaque; 1 point), diabetes (1 point) or any of ischemic stroke, TIA or systemic embolism (2 points).[21] The central place of CHA_2DS_2-VASc in clinical decision-making comes from its role in assessing whether an AF patient's cardiovascular risk profile is such that treatment with oral anticoagulation (OAC; vitamin K antagonists [VKA] or direct oral anticoagulation [DOAC]) to decrease the risk of ischemic events outweighs the increased risk of bleeding associated with OAC use.[4]

Other risk scores commonly turned to in the decision to initiate OAC are the HAS-BLED (Hypertension, Abnormal renal/liver function, Stroke, Bleeding history or predisposition, Labile international normalized ratio, Elderly (> 65 years), Drugs/alcohol concomitantly) and GARFIELD-AF (Global Anticoagulant Registry in the FIELD–Atrial Fibrillation) risk scores. HAS-BLED focuses on the bleeding risk associated with OAC initiation in AF patients. It can be used to assess the presence of modifiable risk factors that could be mitigated to reduce the risk of major bleeding when using OAC.[22] The GARFIELD-AF score comes with an online tool that allows one to enter a range of patient characteristics, resulting in a graphic display of how the shared decision to initiate (type of) stroke prophylaxis can affect the newly diagnosed AF patient's absolute risk of stroke, bleeding and mortality.[23] However, CHA_2DS_2-VASc currently remains most prominent in the choice for anticoagulation initiation.[4] In absence of contraindications for OAC treatment, initiation of OAC is indicated in AF patients with a CHA_2DS_2-VASc score ≥2 in men or ≥3 in women, and OAC initiation should be considered in those with CHA_2DS_2-VASc ≥1 in men or ≥2 in women.[4]

Challenges to early AF detection

There are two main challenges in diagnosing AF. First, the arrhythmia is often asymptomatic, commonly referred to as 'silent AF'. As a consequence, AF patients may be unaware of, and will therefore not seek medical care for, their arrhythmia.[24,25] In the meantime, asymptomatic AF patients are at similar risk of adverse outcomes compared to those who experience AF-related symptoms.[26]

A second challenge is that an AF patient's heart rhythm may not always be in AF, and instead can display normal sinus rhythm or other non-AF rhythms between 'episodes' of AF. Depending on the duration of AF episodes until remission into non-AF rhythm, guidelines distinguish 'paroxysmal AF' (AF episodes lasting up to 7 days), 'persistent AF' (AF episodes lasting 7 days up to 12 months, or AF episodes terminated by cardioversion <12 months' duration), 'long-standing persistent AF' (uninterrupted AF episodes over 12 months' duration), and 'permanent AF' (presence of AF that is accepted by both patient and physician without further efforts to restore or maintain sinus rhythm).[4] For ease of use, we will use the term 'paroxysmal AF' (pAF) throughout this thesis to refer to non-continuous AF patterns. In the presence of pAF, it is possible that a standard 10-second ECG, or even a continuous ECG recording lasting several days, could 'miss' AF episodes when by chance no episodes of pAF occurred within the observation window.

AF screening

Given these challenges, and in light of the importance of AF as a widely prevalent condition for which effective stroke prophylaxis is available, there has been increasing interest in early detection of AF through more systematic case-finding regimes.[27] The assumption here is that a proportion of AF-related strokes can be prevented by facilitating early OAC initiation in AF patients whose AF would otherwise have remained undetected, by reducing their time at unmitigated risk of stroke. In order to assess whether a more systematic approach to finding undetected AF cases is feasible, researchers have investigated the merits of screening for AF in different populations.[28] Here, screening should be understood as an approach to investigating the presence of a disease in persons who have not sought medical attention on their own initiative for the disorder being screened for (paraphrasing Wald, 2001).[29]

One population in whom there is now consensus to screen for AF are patients with a recent ischemic stroke or TIA.[4] Not only is AF often involved as a cause for stroke, which would increase the chance of AF detection in these 'post-stroke' patients. Screening for AF after stroke is also clinically relevant, as an AF diagnosis would affect the choice of treatment to prevent a new stroke or TIA.[30] All patients with a history of ischemic stroke or TIA are to be treated with medication to prevent a new stroke episode. In those without AF, the choice for stroke prophylaxis is antiplatelet (AP) drugs, whereas in those with a concomitant AF diagnosis the physician should opt for OAC (VKA or DOAC).[4,31] This is because in patients with a history of AF and stroke, oral anticoagulation provides a much more effective reduction in risk of recurrent stroke than AP treatment, while only marginally increasing the risk of bleeding associated with oral anticoagulation.[4] In patients with recent stroke or TIA and without a history of AF, there is therefore consensus to perform at least a 10-second ECG at presentation, followed by continuous rhythm monitoring in those without AF on said ECG.[4,30,32,33] While there is ongoing debate on the recommended minimum of continuous monitoring, currently ranging from 24 hours (National Institute for Health and Care Excellence) to 72 hours (European Society of Cardiology, Netherlands Society of Neurology), the clinical value of screening for AF after stroke/TIA is thus well established.[4,30-33]

But would it not make more sense to screen for AF in community settings, *before* a stroke or TIA has even occurred? While this question seems intuitive, answering it has proven to be more difficult than one might expect. Several early studies on community AF screening resulted in a higher AF yield than through routine care. The British SAFE trial (Screening for AF in the Elderly), for instance, enrolled primary care patients aged ≥65 years who were free of AF at baseline. It found that opportunistic AF case finding and systematic AF screening resulted in a similarly high one-year AF yield (1.64%

and 1.62% , respectively) when compared to that of routine primary care (1.04%), with both intervention arms resulting in significantly higher AF yield than usual care.[34] Dutch research showed that screening for AF in a convenience sample of primary care patients aged 60 or older presenting for seasonal influenza vaccination resulted in 1.1% newly diagnosed AF of whom a majority had an indication for anticoagulation initiation.[35] And research performed in the United States of America saw that continuous screening for 14 days with wearable monitors in Medicare beneficiaries aged 75 years or aged over 55 (male) or 65 (female) years with one or more cardiovascular comorbidities resulted in a 4-fold increase in AF detection after 1 year.[36]

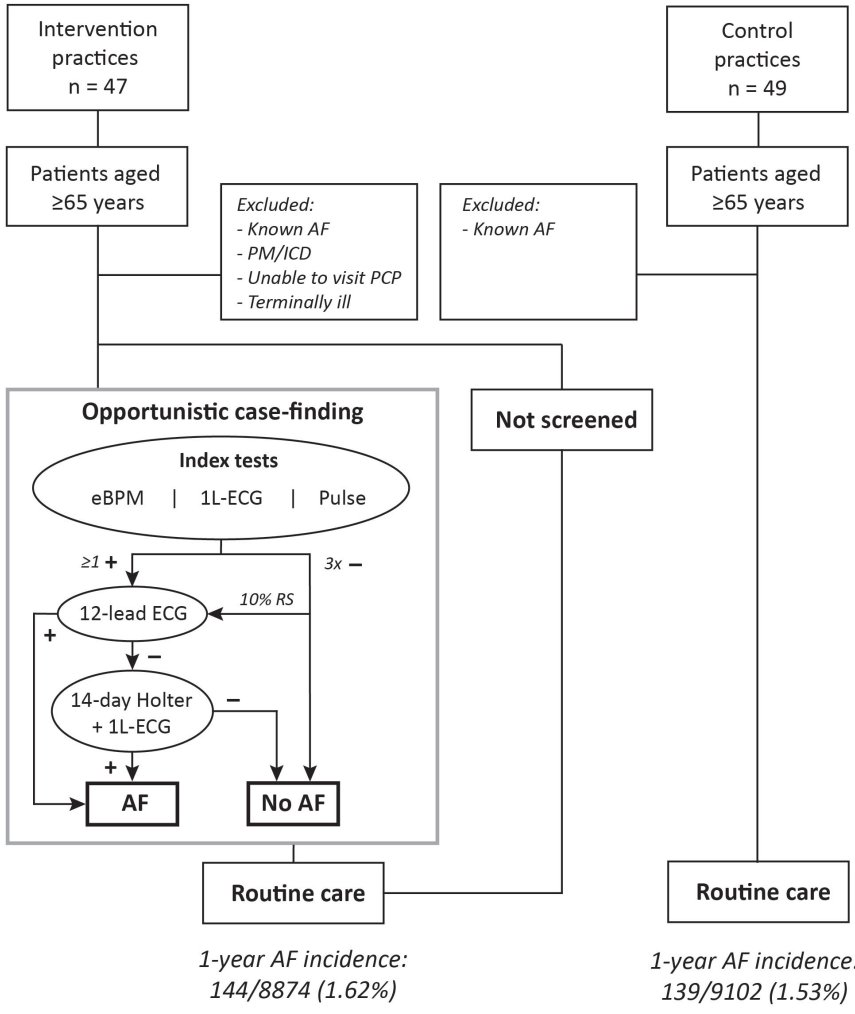
The D_2AF trial: neutral AF screening results in The Netherlands

However, validation of findings from the SAFE trial in Dutch primary care did not result in similar benefit from AF screening. The Detecting and Diagnosing AF study (D_2AF; conducted by our group in collaboration with the Department of Family Medicine, Care and Public Health Research Institute at Maastricht University, The Netherlands) was a large, open-label, cluster-randomized trial in 96 primary care practices in The Netherlands. Like SAFE, it included primary care patients aged ≥65 years without a history of AF. Intervention practices engaged in opportunistic case-finding: initiation of the AF screening protocol once patients visited their primary care practice on their own initiative, for any cause. This is thus different from systematic screening, in which patients are actively invited regardless of contact through routine care. Included intervention patients underwent three index tests: electronic sphygmomanometers with automated AF detection algorithm; MyDiagnostick single-lead ECG (1L-ECG) recording with automated AF detection algorithm; and pulse palpation for irregular rhythm. Those with ≥1 positive index test and a sample of 10% index-negative participants were invited for 12-lead ECG, followed by 'Holter' (continuous ECG) and twice-daily MyDiagnostick 1L-ECG for 14 days in case of AF-negative 12-lead ECG (Figure 2). The D_2AF trial did not result in higher one-year AF yield in the intervention practices compared to those who followed routine care (1.62% versus 1.53% incident AF; odds ratio: 1.06; 95% confidence interval: 0.84-1.35).[37]

Several potential reasons for the neutral results seen in D_2AF, as opposed to the significant benefit of AF screening seen in the SAFE trial, were identifiable. First, it was observed that the one-year AF incidence in D_2AF's routine care arm was similar to that of the intervention arm of its British counterpart, while baseline AF prevalence among potentially eligible 65-year-olds was higher prior to study start (10.1% in D_2AF vs 7.3% in SAFE).[34,37] While a Hawthorne effect among practices randomized to the routine care arm could not be excluded[38], the D_2AF results could thus indicate that Dutch routine care might already be relatively conducive to identifying new AF among

older patients. Second, it was seen that those who participated in D_2AF's intervention protocol were of lower cardiovascular burden than those in the intervention arm who were not screened. It could be that screening efforts aimed specifically at higher-risk patients would have resulted in higher AF yield in the intervention arm. Finally, those in D_2AF's intervention arm who completed the full intervention protocol by also undergoing 14-day 'Holter' (continuous ECG) were few, and again of lower *a priori* AF risk than those in the control arm.[37]

Figure 2. *The Detecting and Diagnosing Atrial Fibrillation (D_2AF) study protocol.*



1L-ECG, single-lead electrocardiogram; AF, atrial fibrillation; eBPM-AF, electronic automatic sphygmomanometer with atrial fibrillation detection algorithm; ECG, electrocardiogram; PCP, primary care practice; PM/ICD, pacemaker or implantable cardioverter-defibrillator; RS, random sample.
Adapted from Uittenbogaart et al., *BMJ 2020*.[37]

Towards higher yield of primary AF screening

The results of the D_2AF trial thus provided us with multiple options to potentially increase the yield of future primary AF screening efforts. The main opportunities seemed to be in patient selection for the screening intervention, rather than selecting for age ≥65 years alone. By boosting the *a priori* chance of detecting AF through better patient selection, efficiency of the screening program could be increased beyond the already high AF yield through Dutch routine care. Then, once high risk for AF is established, there also remained a need for novel screening methods with a lower threshold than the elaborate D_2AF protocol. The main questions that will be addressed in this thesis, will therefore be:

1) Whom to screen?

And if selected for screening:

2) How to screen?

The second question can be further subdivided into a number of subquestions. For instance, 'how long to screen?', 'how often to screen?', or 'with which device(s) to screen?'[27] As one thesis would not be sufficient to answer all these (sub)questions, we have focused primarily on the question whom to screen, with a foray into what device would be helpful in lowering the threshold for timely AF diagnosis.

Whom to screen: multivariable risk models for AF prediction

In assessing risk of AF in the context of whom to screen for the arrhythmia, cardiovascular risk factors are the intuitive starting point of one's investigation. Multivariable risk models combine the information contained in several clinical variables to assess a person's precise risk of an outcome.[39] If we could identify a multivariable risk model that has a higher predictive value for risk of AF than e.g. age alone, researchers could use such a risk model to more efficiently select patients eligible for AF screening.[40]

In clinical practice, this would have to be offset with the ease by which the variables within such a model would have to be gathered. In order to prevent patients from having to visit a clinic for additional investigations before screening selection can even commence – with all costs and effort involved – it is after all most practical if variables for a risk model could reliably be extracted from available, routinely collected healthcare data. This in turn means that different healthcare settings will have a different optimum in the amount and type of variables than could best be included in an efficient AF risk model.[39] For instance, whereas all patients presenting for a stroke in a Neurology ward will typically undergo brain imaging[31], such advanced investigations are not systematically performed in primary care. While brain imaging

features are increasingly recognised as predictors for AF[41], their value as predictors in a multivariable model as triage test for AF screening are thus specific to the context in which the model would be used.

In this thesis we therefore performed a systematic review and meta-analysis in which we assessed which multivariable models for predicting the risk of AF have been developed and/or validated specifically for primary care or community settings, and which of those would most likely be best for patient selection for community AF screening. We subsequently took the models that showed highest performance and validated these within a cohort of AF-free older patients included in a national routine primary care EHR database in the Netherlands. In this analysis we assessed the amount of missing data in the EHR database for the variables in our target multivariable models, and compared the performance of the models with that of age alone as a means of assessing whether multivariable models should be considered over age alone as a means of selecting patients for AF screening.

Whom to screen: premature atrial contractions on ECG signal

While many clinical risk factors for AF had previously been established as potential candidates in AF risk models,[42] one interesting candidate – an electrocardiographic entity just as AF – was up to recently quite poorly understood: premature atrial contractions (PACs).[43] All physicians have encountered these extrasystoles on ECG or Holter throughout their medical career, and up to recently PACs were considered a relatively benign finding in cardiac rhythm assessment.[43] However, while almost all older adults would show one or more PACs each day if they were to wear continuous ECG,[44] an increasing amount of evidence indicates that frequent PACs should be seen as electrocardiographic (bio)markers of pathophysiological changes to the atrium.[45,46]

The exact pathophysiological pathways that could explain the association between frequent PACs and AF or stroke are not clear, but their intricate relation is more and more understood. It is thought that a person's cardiovascular risk burden, e.g. high age or history of hypertension or coronary artery disease, increases the degree of pathophysiological changes to the atrium, a concept known as atrial cardiomyopathy.[14,47] This can be through e.g. altered autonomous innervation to, or chronic inflammation processes within, the atrial wall. This culminates in changes to the atrial tissue, most prominently fibrosis – a process known as atrial structural remodelling. This in turn translates into changes to electrophysiological conduction within the atrium, such as re-entry pathways, enhanced automaticity or early or delayed afterdepolarisations. These can be seen as PACs and other non-sustained atrial arrhythmias on ECG.[47]

Ultimately these atrial changes could lead to reoccurring bouts of AF (paroxysmal AF, or pAF), or even permanent AF, conditions with an established association with stroke and mortality.[4,9] How exactly the pathophysiological changes to the atrium leading to frequent PACS or (p)AF translate into risk of ischemic stroke and death is unclear. This could be as a direct mechanistic pathway through e.g. thrombus formation in the atrium, or indirectly with both atrial cardiomyopathy and stroke as pathophysiological consequences of prior cardiovascular burden.[47] For that, we will have to await the results of future investigations into the mechanistic link between atrial cardiomyopathy and stroke and death.

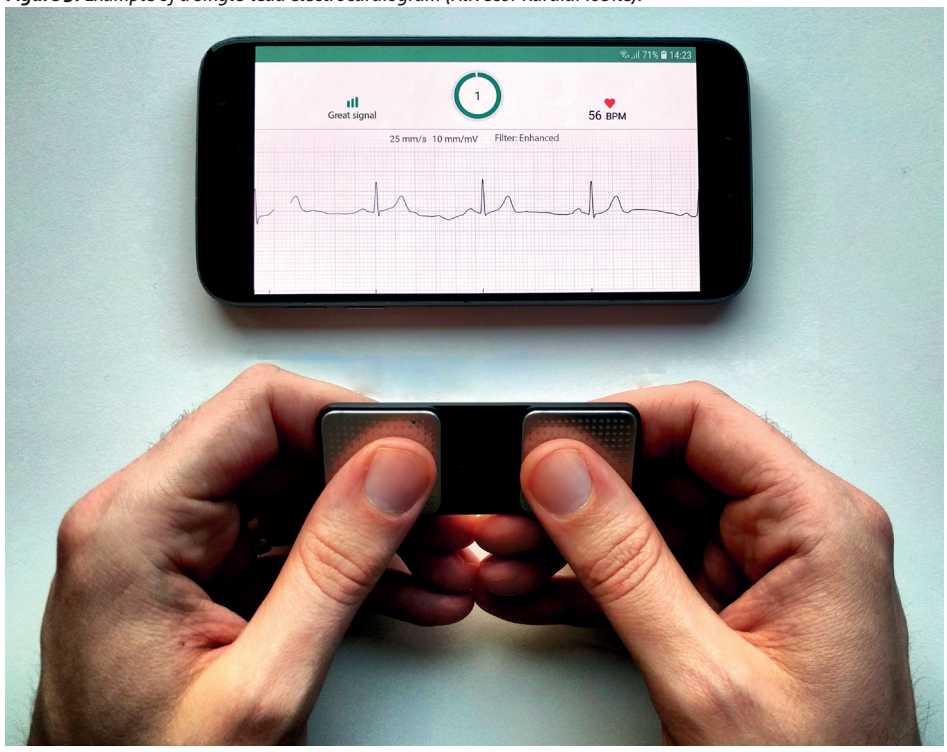
If sufficiently predictive, presence of frequent PACs on (continuous) ECG could be useful in assessing risk of AF for screening purposes, whether as singular variable or e.g. as addition to multivariable clinical models for AF prediction. Given the increasing evidence pointing towards frequent PACs as a predictor for adverse cardiovascular outcomes,[14] it is thus important to assess the relative risk associated with the presence of frequent PACs on ECG recordings, and what constitutes 'frequent' in this regard.

At the time of the start of this PhD program, these were all relatively unclear with only a few systematic reviews on the topic, and those that provided meta-analyses having considerable risks of bias in the way they accumulated the evidence available at that time.[48,49] For this thesis, we therefore performed a systematic review and meta-analysis into the association between PACs and AF, stroke and all-cause mortality with a rigorous protocol to assess the meaning of finding (frequent) PACs on two clinically distinct ECG platforms: standard (short) ECG and Holter (continuous) ECG. We subsequently tested the hypotheses generated in the systematic review in a prospective cohort of people with type-2 diabetes mellitus in a primary care setting and assessed the association between PACs, AF and stroke in this clinically important patient group.

Whom and how to screen: AI algorithms applied to ECG signal

Another emerging way in which early AF detection could be achieved from the use of ECG signal, is by the application of artificial intelligence (AI) algorithms on electrocardiographic or pulse signal data. Such AI algorithms can be trained to detect whether AF is currently manifest on e.g. 1L-ECG (Figure 3) or photoplethysmographic recordings, and have achieved good results for this purpose.[50-52] If sufficiently accurate and of high usability in the setting of intended use, a device that is immediately available upon AF suspicion and contains accurate AI for AF detection could lower the threshold for early AF detection compared the current standard, i.e. to have a patient come to a clinic for 12-lead ECG or for administering Holter.

Figure 3. *Example of a single-lead electrocardiogram (Alivecor KardiaMobile).*



Held lightly between the thumb and index finger of both hands, the device relays cardiac microvoltages analogous to lead I of the standard 12-lead electrocardiogram. The signal is immediately relayed to and displayed on a connected smartphone, where artificial intelligence-trained software analyses the 30-second recording for presence of atrial fibrillation.
Adapted from Himmelreich et al., *Annals of Family Medicine 2019.*[53]

Another potential use of AI is to assess whether a rhythm recording of non-AF ECG signal may still have the ECG 'fingerprint' of someone who would likely be diagnosed with pAF if he/she would be monitored for an extended duration.[54] The use of AI for this purpose is based on the atrial cardiomyopathy framework discussed previously in this chapter.[14] Here, a pAF patient currently in non-AF rhythm could still have traits of atrial cardiomyopathy (e.g. frequent PACs, irregular RR intervals short of AF or P-wave morphology variants) that could cross an AI-assessed 'threshold' for potential presence of underlying pAF.[55] If sufficiently accurate, such an AI algorithm could be used as a triage test for prolonged monitoring. Based on a short initial continuous monitoring period such an AI algorithm could tell us whom to screen for longer (those assessed as at high risk of underlying pAF), and for whom further rhythm monitoring would not be necessary (those at low pAF risk as per the AI algorithm). Ultimately, the AI algorithm could thus assist allocating costly and burdensome extended monitoring to only those at highest risk.

Given the dual potential use of ECG-processing AI algorithms in our quest for earlier AF detection, we performed two prospective studies for this thesis. First, we validated a hand-held ECG device for the detection of manifest AF in patients who underwent standard ECG in primary care against simultaneously performed 10-second ECG (reference standard). Second, we validated an AI algorithm that assesses risk of pAF during non-AF rhythm snippets of continuous ECG recording against the outcome of subsequent Holter monitoring (AF or no AF, reference standard) in prospectively enrolled primary care and post-stroke patients who underwent 14-day Holter recording.

AIM AND OUTLINE OF THIS THESIS

The aim of this thesis was to investigate means to better identify patients at high risk of AF in order to facilitate early AF detection in the primary care and post-stroke setting. To this end, this thesis includes the following chapters:

PART I: RISK MODELS USING CLINICAL PREDICTORS AND BIOMARKERS FOR ATRIAL FIBRILLATION IN THE COMMUNITY

Chapter 2 presents the results of a systematic review and meta-analysis on multivariable risk models developed and/or validated for AF risk prediction in community settings.

Chapter 3 validates a number of multivariable risk models featured in Chapter 2 against age alone for predicting 5-year risk of AF in a national routine primary care electronic health records database in the Netherlands.

PART II: PREMATURE ATRIAL CONTRACTIONS AS AN ELECTROCARDIOGRAPHIC RISK FACTOR FOR ATRIAL FIBRILLATION

Chapter 4 shows the results of a systematic review and meta-analysis on the association between PACs on standard ECG or frequent PACs on continuous ECG and the risk of AF, brain ischemia or all-cause mortality.

Chapter 5 describes the longitudinal association of PACs with AF and brain ischemia in a database of prospectively enrolled people with type 2 diabetes in Dutch primary care.

PART III: IDENTIFYING ATRIAL FIBRILLATION USING ARTIFICIAL INTELLIGENCE ALGORITHMS ON ELECTROCARDIOGRAPHIC RECORDINGS IN PRIMARY CARE AND POST-STROKE PATIENTS

Chapter 6 presents the results of a diagnostic accuracy study validating a hand-held, AI-enabled 1L-ECG device for detection of rhythm and conduction abnormalities including AF in prospectively enrolled primary care patients who underwent 12-lead ECG as per routine care.

Chapter 7 describes the details of enrolment, baseline characteristics and AF yield in a cohort of prospectively enrolled, consecutive patients presenting to an academic hospital in The Netherlands for TIA or ischemic stroke who underwent 14-day Holter for AF.

Chapter 8 presents a diagnostic accuracy study validating an AI algorithm that assesses the risk of underlying pAF during non-AF rhythm on Holter in the post-stroke/TIA patients featured in Chapter 7 as well as in patients from the intervention arm of a cluster-randomised, controlled AF screening trial in Dutch primary care, all of whom underwent 14-day Holter for AF.

PART IV: GENERAL DISCUSSION AND SUMMARY

Chapter 9 provides a general discussion on the findings of this thesis in relation to the current body of research, followed by an outline for potential future research that follow from the general discussion.

Chapters 10 and 11 contain a summary of this thesis in English and in Dutch, respectively.

REFERENCES

1. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart Disease and Stroke Statistics-2022 Update: A Report From the American Heart Association. Circulation. 2022;145(8):e153-e639.
2. Staerk L, Wang BQ, Preis SR, Larson MG, Lubitz SA, Ellinor PT, et al. Lifetime risk of atrial fibrillation according to optimal, borderline, or elevated levels of risk factors: cohort study based on longitudinal data from the Framingham Heart Study. Bmj-Brit Med J. 2018;361.
3. 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(35):2746-51.
4. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). Eur Heart J. 2021;42(5):373-498.
5. Dorian P, Jung W, Newman D, Paquette M, Wood K, Ayers GM, et al. The impairment of health-related quality of life in patients with intermittent atrial fibrillation: implications for the assessment of investigational therapy. J Am Coll Cardiol. 2000;36(4):1303-9.
6. Kim MH, Johnston SS, Chu BC, Dalal MR, Schulman KL. Estimation of total incremental health care costs in patients with atrial fibrillation in the United States. Circ Cardiovasc Qual Outcomes. 2011;4(3):313-20.
7. Johnsen SP, Dalby LW, Tackstrom T, Olsen J, Fraschke A. Cost of illness of atrial fibrillation: a nationwide study of societal impact. BMC Health Serv Res. 2017;17(1):714.
8. De With RR, Rienstra M, Smit MD, Weijs B, Zwartkruis VW, Hobbelt AH, et al. Targeted therapy of underlying conditions improves quality of life in patients with persistent atrial fibrillation: results of the RACE 3 study. Europace. 2019;21(4):563-71.
9. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991;22(8):983-8.
10. Marini C, De Santis F, Sacco S, Russo T, Olivieri L, Totaro R, et al. Contribution of atrial fibrillation to incidence and outcome of ischemic stroke - Results from a population-based study. Stroke. 2005;36(6):1115-9.
11. Wolf PA, Dawber TR, Thomas HE, Jr., Kannel WB. Epidemiologic assessment of chronic atrial fibrillation and risk of stroke: the Framingham study. Neurology. 1978;28(10):973-7.
12. Hart RG, Diener H-C, Coutts SB, Easton JD, Granger CB, O'Donnell MJ, et al. Embolic strokes of undetermined source: the case for a new clinical construct. The Lancet Neurology. 2014;13(4):429-38.
13. Buchwald F, Norrving B, Petersson J. Atrial Fibrillation in Transient Ischemic Attack Versus Ischemic Stroke: A Swedish Stroke Register (Riksstroke) Study. Stroke. 2016;47(10):2456-61.
14. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model. Stroke. 2016;47(3):895-900.
15. Camm AJ, Fox KAA, Virdone S, Bassand JP, Fitzmaurice DA, Berchuck SI, et al. Comparative effectiveness of oral anticoagulants in everyday practice. Heart. 2021.

16. Ruddox V, Sandven I, Munkhaugen J, Skattebu J, Edvardsen T, Otterstad JE. Atrial fibrillation and the risk for myocardial infarction, all-cause mortality and heart failure: A systematic review and meta-analysis. Eur J Prev Cardiol. 2017;24(14):1555-66.
17. Chen TH, Chu YC, Ou SM, Tarng DC. Associations of atrial fibrillation with renal function decline in patients with chronic kidney disease. Heart. 2022;108(6):438-44.
18. Liu DS, Chen J, Jian WM, Zhang GR, Liu ZR. The association of atrial fibrillation and dementia incidence: a meta-analysis of prospective cohort studies. J Geriatr Cardiol. 2019;16(3):298-306.
19. O'Neal WT, Qureshi WT, Judd SE, Bowling CB, Howard VJ, Howard G, et al. Effect of Falls on Frequency of Atrial Fibrillation and Mortality Risk (from the REasons for Geographic And Racial Differences in Stroke Study). Am J Cardiol. 2015;116(8):1213-8.
20. Rattanawong P, Upala S, Riangwiwat T, Jaruvongvanich V, Sanguankeo A, Vutthikraivit W, et al. Atrial fibrillation is associated with sudden cardiac death: a systematic review and meta-analysis. J Interv Card Electrophysiol. 2018;51(2):91-104.
21. 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(2):263-72.
22. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJGM, Lip GYH. A Novel User-Friendly Score (HAS-BLED) To Assess 1-Year Risk of Major Bleeding in Patients With Atrial Fibrillation The Euro Head Survey. Chest. 2010;138(5):1093-100.
23. Fox KAA, Virdone S, Pieper KS, Bassand JP, Camm AJ, Fitzmaurice DA, et al. GARFIELD-AF risk score for mortality, stroke and bleeding within 2 years in patients with atrial fibrillation. Eur Heart J Qual Care Clin Outcomes. 2021.
24. 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(11):1587-93.
25. Boriani G, Laroche C, Diemberger I, Fantecchi E, Popescu MI, Rasmussen LH, et al. Asymptomatic atrial fibrillation: clinical correlates, management, and outcomes in the EORP-AF Pilot General Registry. Am J Med. 2015;128(5):509-18 e2.
26. Sgreccia D, Manicardi M, Malavasi VL, Vitolo M, Valenti AC, Proietti M, et al. Comparing Outcomes in Asymptomatic and Symptomatic Atrial Fibrillation: A Systematic Review and Meta-Analysis of 81,462 Patients. J Clin Med. 2021;10(17).
27. Freedman B, Camm J, Calkins H, Healey JS, Rosenqvist M, Wang JG, et al. Screening for Atrial Fibrillation A Report of the AF-SCREEN International Collaboration. Circulation. 2017;135(19):1851-+.
28. Af S, Collaborators A-E. Protocol for a Systematic Review and Individual Participant Data Meta-Analysis of Randomized Trials of Screening for Atrial Fibrillation to Prevent Stroke. Thromb Haemost. 2023;123(3):366-76.
29. Wald NJ. The definition of screening. J Med Screen. 2001;8(1):1-.
30. Rubiera M, Aires A, Antonenko K, Lemeret S, Nolte CH, Putaala J, et al. European Stroke Organisation (ESO) guideline on screening for subclinical atrial fibrillation after stroke or transient ischaemic attack of undetermined origin. European Stroke Journal. 2022;7(3):CVII–CXXXIX.

31. Nederlandse Vereniging voor Neurologie. Richtlijn Hersenifarct en Hersenbloeding. 2017. https://richtlijnendatabase.nl/richtlijn/herseninfarct_en_hersenbloeding/startpagina_herseninfarct_-bloeding.html. p. 1-1151.
32. National Institute for Health and Care Excellence. Atrial fibrillation: diagnosis and management (NICE guideline NG196). 2021. https://www.nice.org.uk/guidance/ng196. p. 1-43.
33. Kleindorfer DO, Towfighi A, Chaturvedi S, Cockroft KM, Gutierrez J, Lombardi-Hill D, et al. 2021 Guideline for the Prevention of Stroke in Patients With Stroke and Transient Ischemic Attack: A Guideline From the American Heart Association/American Stroke Association. Stroke. 2021;52(7):e364-e467.
34. 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(7616):383.
35. 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(10):1514-20.
36. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. JAMA. 2018;320(2):146-55.
37. Uittenbogaart SB, Verbiest-van Gurp N, Lucassen WAM, Winkens B, Nielen M, Erkens PMG, et al. Opportunistic screening versus usual care for detection of atrial fibrillation in primary care: cluster randomised controlled trial. BMJ. 2020;370:m3208.
38. McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne Effect: a randomised, controlled trial. BMC Med Res Methodol. 2007;7.
39. Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 2012;98(9):683-90.
40. Poorthuis MHF, Jones NR, Sherliker P, Clack R, de Borst GJ, Clarke R, et al. Utility of risk prediction models to detect atrial fibrillation in screened participants. Eur J Prev Cardiol. 2020.
41. Kim JG, Boo K, Kang CH, Kim HJ, Choi JC. Impact of Neuroimaging Patterns for the Detection of Atrial Fibrillation by Implantable Loop Recorders in Patients With Embolic Stroke of Undetermined Source. Front Neurol. 2022;13:905998.
42. Allan V, Honarbakhsh S, Casas JP, Wallace J, Hunter R, Schilling R, et al. Are cardiovascular risk factors also associated with the incidence of atrial fibrillation? A systematic review and field synopsis of 23 factors in 32 population-based cohorts of 20 million participants. Thromb Haemost. 2017;117(5):837-50.
43. Marcus GM, Dewland TA. Premature Atrial Contractions: A Wolf in Sheep's Clothing? J Am Coll Cardiol. 2015;66(3):242-4.
44. Conen D, Adam M, Roche F, Barthelemy JC, Felber Dietrich D, Imboden M, et al. Premature atrial contractions in the general population: frequency and risk factors. Circulation. 2012;126(19):2302-8.
45. Binici Z, Intzilakis T, Nielsen OW, Kober L, Sajadieh A. Excessive supraventricular ectopic activity and increased risk of atrial fibrillation and stroke. Circulation. 2010;121(17):1904-11.

46. Gaita F, Castagno D. Do supraventricular premature beats identify patients at high risk for atrial fibrillation? J Cardiovasc Med. 2017;18:e117-e20.
47. Guichard JB, Guasch E, Roche F, Da Costa A, Mont L. Premature atrial contractions: A predictor of atrial fibrillation and a relevant marker of atrial cardiomyopathy. Front Physiol. 2022;13.
48. Huang BT, Huang FY, Peng Y, Liao YB, Chen F, Xia TL, et al. Relation of premature atrial complexes with stroke and death: Systematic review and meta-analysis. Clin Cardiol. 2017;40:962-9.
49. Sejr MH, Riahi S, Larsen TB, Nielsen JC, Nielsen PB. Premature atrial complexes in an ischemic stroke population and risk of recurrent stroke: a systematic review. Expert Rev Cardiovasc Ther. 2017;15(6):447-55.
50. 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(6):1167-76.
51. 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(9):1291-5.
52. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, et al. Passive Detection of Atrial Fibrillation Using a Commercially Available Smartwatch. JAMA Cardiol. 2018;3(5):409-16.
53. Himmelreich JCL, Karregat EPM, Lucassen WAM, van Weert H, de Groot JR, Handoko ML, et al. Diagnostic Accuracy of a Smartphone-Operated, Single-Lead Electrocardiography Device for Detection of Rhythm and Conduction Abnormalities in Primary Care. Ann Fam Med. 2019;17(5):403-11.
54. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet. 2019.
55. Duning T, Kirchhof P, Wersching H, Hepp T, Reinhardt R. Extended Electrocardiographic Poincare Analysis (EPA) for Better Identification of Patients with Paroxysmal Atrial Fibrillation. J Clin Exp Cardiolog. 2011;02(02).

2

Prediction models for atrial fibrillation applicable in the community: a systematic review and meta-analysis

Jelle C.L. Himmelreich, Lieke Veelers, Wim A.M. Lucassen, Renate B. Schnabel, Michiel Rienstra, Henk C.P.M. van Weert, Ralf E. Harskamp

EP Europace, Volume 22, Issue 5,

May 2020, Pages 684–694,

https://doi.org/10.1093/europace/euaa005

ABSTRACT

Aims: Atrial fibrillation (AF) is a common arrhythmia associated with an increased stroke risk. The use of multivariable prediction models could result in more efficient primary AF screening by selecting at-risk individuals. We aimed to determine which model may be best suitable for increasing efficiency of future primary AF screening efforts.

Methods and results: We performed a systematic review on multivariable models derived, validated, and/or augmented for AF prediction in community cohorts using Pubmed, Embase, and CINAHL (Cumulative Index to Nursing and Allied Health Literature) through 1 August 2019. We performed meta-analysis of model discrimination with the summary C-statistic as the primary expression of associations using a random effects model. In case of high heterogeneity, we calculated a 95% prediction interval. We used the CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies) checklist for risk of bias assessment. We included 27 studies with a total of 2 978 659 unique participants among 20 cohorts with mean age ranging from 42 to 76 years. We identified 21 risk models used for incident AF risk in community cohorts. Three models showed significant summary discrimination despite high heterogeneity: CHARGE-AF (Cohorts for Heart and Aging Research in Genomic Epidemiology) [summary C-statistic 0.71; 95% confidence interval (95% CI) 0.66–0.76], FHS-AF (Framingham Heart Study risk score for AF) (summary C-statistic 0.70; 95% CI 0.64–0.76), and CHA_2DS_2-VASc (summary C-statistic 0.69; 95% CI 0.64–0.74). Of these, CHARGE-AF and FHS-AF had originally been derived for AF incidence prediction. Only CHARGE-AF, which comprises easily obtainable measurements and medical history elements, showed significant summary discrimination among cohorts that had applied a uniform (5-year) risk prediction window.

Conclusion: CHARGE-AF appeared most suitable for primary screening purposes in terms of performance and applicability in older community cohorts of predominantly European descent.

WHAT'S NEW

- This is the first systematic review and meta-analysis designed to capture and evaluate a broad range of prognostic models used for incident atrial fibrillation (AF) risk prediction, and the first to focus specifically on models that are applicable in and have been derived, validated and/or augmented in community cohorts.
- This work was open to any model used for incident AF prediction in the community, which also enabled inclusion of models that had not been developed for incident AF but that may have merits for this aim. We hereby identified 21 models used for incident AF prediction in community cohorts.
- This work suggests that the CHARGE-AF model is likely most robust for incident AF prediction in terms of performance as well as applicability in the community.

INTRODUCTION

Atrial fibrillation (AF) is a common cardiac arrhythmia affecting over 33 million people worldwide.[1] Its incidence increases with age, with a lifetime risk of over 30%.[2] Due to ageing populations, the number of AF cases in Europe is expected to double to >17 million by 2060.[3] Atrial fibrillation is associated with a five-fold increased risk of ischaemic stroke, which can be largely prevented by antithrombotic prophylaxis in at-risk patients.[4,5]

Screening for AF in the community has been proposed as an approach to optimize early AF detection and to prevent AF-associated sequelae.[6] Prior research has shown that AF screening is cost-effective when selecting patients at older age, with thresholds for screening eligibility varying from 65 to 75 years.[7–9] The screening regimes in these primary care studies often involved single-point screening, while multiple-point screening could result in a higher yield of new AF cases.[9,10] Multiple-point or prolonged rhythm monitoring schemes are, however, likely to be more costly for society and more burdensome to patients.[7–9] Multivariable prediction models for incident AF could contribute to AF screening by determining a risk category for each patient.[11] The more intensive regimes could be assigned to those with highest risk, while those in lower-risk strata could be assigned to less stringent follow-up, or none at all. It remains, however, insufficiently clear from consensus documents whether other parameters beyond age could be used to differentiate between degrees of AF risk within the community.[5,12]

We therefore set out to perform a systematic review and meta-analysis with two aims. First, we wished to provide an overview of AF risk models that are applicable to and have been validated in community cohorts. Such models should consist of variables that can be quickly assessed and/or are commonly available from patient records and should not require advanced diagnostic testing. Second, by synthesizing the discriminatory abilities of each included risk model, we aimed to determine which of these may be best suitable for increasing efficiency of future primary AF screening efforts.

METHODS

We reported this systematic review and meta-analysis in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[13]

Data searches

We searched Pubmed, Embase, and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases from inception through 1 August 2019. We used the keywords 'AF or atrial flutter (AFl)' and 'risk model' as well as related terms. We filtered for studies conducted on humans and written in English. The full search is shown in Supplementary material, Table S1. We checked the reference list of included studies for additional relevant references.

Study selection

To be eligible for inclusion, studies had: (i) to be original studies in adults (≥18 years of age); (ii) to derive, validate, and/or augment a tool for predicting risk of incident AF/AFl based on multivariable analysis; (iii) to include only patients without a diagnosis of AF/AFl at baseline; and (iv) to incorporate into their risk prediction tool only variables that are applicable and/or commonly available in primary care settings. We included studies with AFl as co-outcome, since AF and AFl have similar clinical relevance.[5] In light of inclusion criterion iv, we included only studies that used medical history, physical examination, simple laboratory findings, or electrocardiogram (ECG) parameters as variables in the prediction model. We excluded studies that required advanced diagnostic testing [e.g. echocardiography, genetic markers, or specialized (laboratory) tests] for their simple (non-augmented) model. We only included studies written in English. We included studies that diagnosed AF/AFl through medical records, hospitalizations, death certificates, and/or ECG during follow-up examinations. We excluded studies that selected patients for a common disease or risk factor, as such studies would not be generalizable to the community. Moreover, we excluded studies with a mean follow-up duration under 3 months since with shorter follow-up durations there would be an increasing risk of measuring prevalent AF missed at baseline recording, rather than actual incident AF. We uploaded references to a systematic review web application (Covidence, Veritas Health Innovation Ltd, Melbourne, Australia). Three investigators (J.C.L.H., L.V., and R.E.H.) assessed studies for eligibility by screening studies on title and abstract, followed by a full-text screening. Disagreements were resolved by panel discussion (J.C.L.H., L.V., R.E.H., and W.A.M.L.).

Data extraction and quality assessment

Two investigators (J.C.L.H. and L.V.) extracted data from the included studies regarding study methods, population characteristics, risk prediction model(s), and model performance. For the latter, we extracted the C-statistic and corresponding 95% confidence interval (95% CI) for discrimination, and the P-value of a goodness-of-fit test and the ratio of observed and expected AF/AFl cases (O:E ratio) for calibration. When authors did not report an O:E ratio we derived the O:E ratio by analysing calibra-

tion plots.[14] When authors performed augmentation of pre-existing models by adding variables with an aim to enhance predictive value of models, we retrieved the net reclassification improvement (NRI) index of the augmented model compared with the original 'simple' model, as well as the augmented model's performance in terms of discrimination and calibration. We included augmentation data only when the augmentation variables were applicable to primary care settings as outlined previously.

Two investigators (J.C.L.H. and L.V.) used the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist to assess the risk of bias and the applicability for our research aims.[15] Our interpretation of each CHARMS domain can be found in the Supplementary Methods. We assessed risk of bias at the cohort level for each of the included studies. We scored each domain as either low, unclear, or high risk of bias. We defined overall risk of bias as: low, when all domains of a cohort within one study were scored as low risk of bias; unclear, when one or more domains of a cohort within one study were scored as unclear risk of bias; and high, when one or more domains of a cohort within one study were scored as high risk of bias. We resolved disagreements by discussion.

Data synthesis and statistical analysis

We reported continuous variables as means ± standard deviations, and categorical variables as percentages. We evaluated statistical significance in all analyses at the 0.05 level. In individual studies, we assessed the C-statistic of a model, where a 95% CI containing 0.5 indicated insufficient discrimination. Calibration of a model was deemed sufficient when authors reported a P-value of a goodness-of-fit test >0.05 and/or an O:E ratio ranging between 0.95 and 1.05. In assessing augmentation, we defined significant improvement as a positive NRI index with a reported 95% CI that did not contain 0. When a study reported on multiple cohorts, and presented separate data for each cohort, we assessed model performance separately for each cohort within that study.

We performed meta-analysis to assess overall discrimination of included models. The primary expression of associations in meta-analysis was the summary C-statistic and corresponding 95% CI using a random effects inverse variance model with restricted maximum likelihood estimation and Hartung–Knapp corrections.[14] We conducted the meta-analyses in R using the meta and metafor packages (R Foundation for Statistical Computing, version 3.5.1). We performed meta-analysis only when C-statistic data for a prediction model were available for ≥3 cohorts.[16] When studies presented a C-statistic without 95% CI, we calculated the 95% CI using methods described previously.[14] In each meta-analysis, we calculated the mean as the summary effect measure,

its 95% CI, and the I2 statistic as an expression of the heterogeneity between studies.[17] When heterogeneity in meta-analysis of C-statistics was high (I2 > 30%[18]), we derived a 95% prediction interval (95% PI) using methods described previously.[19] We assessed overall discrimination of models by the summary C-statistic. When the 95% CI (or, in case of high heterogeneity, the 95% PI) of the summary C-statistic included 0.5, we concluded that there was insufficient evidence that the prediction model has significant discriminatory ability for incident AF in such populations as included in the meta-analysis.

We assessed eligibility for inclusion into meta-analysis at the cohort level. Cohorts with low or unclear overall risk of bias were eligible for inclusion into meta-analysis. When studies reported C-statistic data based on the aggregation of multiple cohorts, and one of these cohorts was assessed as having high overall risk of bias, we did not include the aggregate C-statistic data into meta-analysis. When multiple studies reported C-statistic data on the same cohort, we included the first published study into the primary analysis. In the primary meta-analysis of each model, we included cohorts with any follow-up duration.

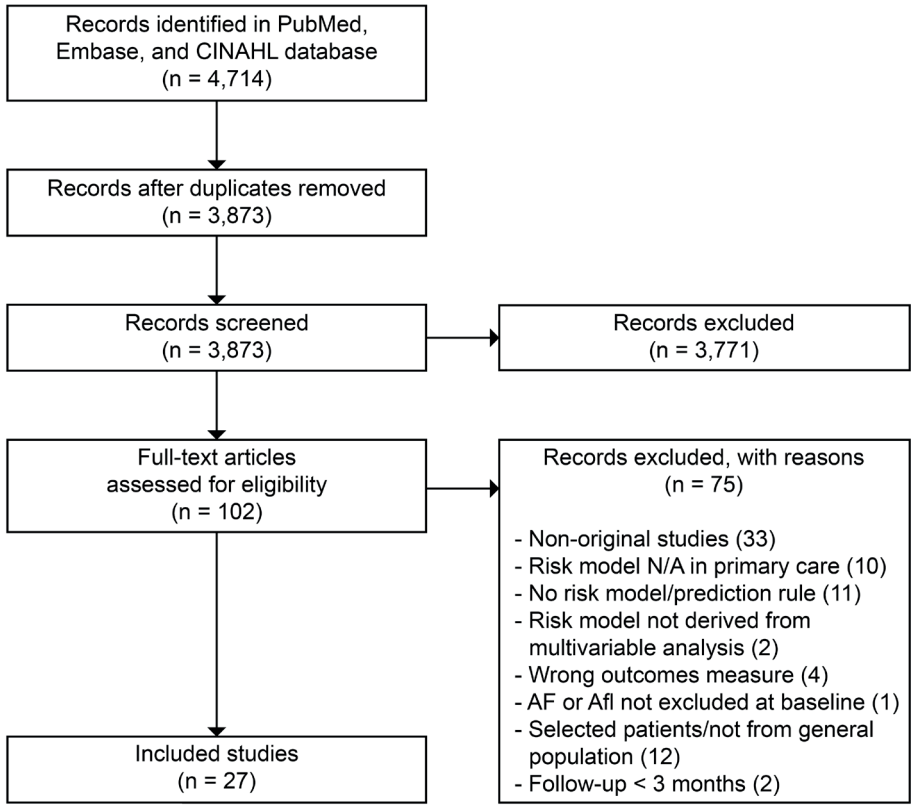
In our primary analysis, we assessed overall discrimination of all models that had ≥3 eligible cohorts with C-statistic data. In the secondary analysis we performed meta-analysis for each risk model that had ≥3 eligible cohorts reporting C-statistic data while applying a uniform prediction window, and grouped cohorts according to the applied risk prediction window (e.g. 5 or 10 years) since this is an important methodological considerations when wanting to translate summary risk model performance to clinical settings.[15]

We performed a sensitivity analysis in which we restricted the primary and secondary analyses to only those cohorts that had demonstrated sufficient calibration in order to assess overall discrimination among populations where the prediction model had also shown the ability to correctly classify absolute incident AF risk. Finally, we performed a sensitivity analysis in which we replaced primary and secondary meta-analyses data from 'double' cohorts (cohorts from the primary meta-analysis that had also been reported on in later studies) with data on that same cohort from any later study to assess whether later, possibly 'more complete' datasets could be of influence to our conclusions from the primary and secondary analyses.

RESULTS

We found a total of 3873 unique references, 102 of which we subjected to full-text screening. From these, we included 27 studies[20–46] for our final analysis (see Figure 1).

Figure 1. *Literature flow diagram.*



AF, atrial fibrillation; AFl, atrial flutter; CINAHL, Cumulative Index to Nursing and Allied Health Literature; NA, not applicable.

Characteristics of included studies

The 27 included studies were based on 20 different cohorts set in Europe (n = 8), East Asia (n = 5), North America (n = 5), and the Middle East (n = 2) (see Supplementary material, Table S2 for characteristics of the included studies and cohorts). Cohort size ranged from 646 to 1 062 073 patients, with a total of 2 978 659 unique participants. Mean age varied from 42 to 76 years, percentage of female participants ranged from 0% to 100%. Mean follow-up of the included cohorts varied from 3 to 20 years, with AF incidence during follow-up ranging from 0.2% to 24.5%. Ten cohorts used AF/Afl

as the outcome, and the other 10 cohorts described only AF. Thirteen of the 20 cohorts followed a prospective design, of which 6 cohorts applied prescheduled follow-up examinations to systematically identify AF.

Characteristics of included risk models

The included studies represented data on 21 multivariable prediction models. Ten models had specifically been derived for predicting incident AF (Table 1). Of these, nine had been derived in community cohorts[20–28] and one had been derived in a cohort of outpatients.[47] Five of the models derived for incident AF had also been externally validated.[20–24] The intended risk prediction window of models derived for incident AF varied between 5 and 11 years. The FHS-AF (Framingham Heart Study risk score for AF) model had originally been derived for predicting 10-year incident AF risk, but had later been recalibrated and subsequently externally validated for 5-year risk of incident AF.[33,40]

We identified seven risk models that had originally been derived for predicting other outcomes than incident AF but had been validated for this outcome in community cohorts[34,48-53] and a further four models that were incidentally employed to predict incident AF but that had not specifically been derived as a prediction model for that outcome[23,37,44,46] (Supplementary material, Table S3).

The number of variables incorporated into each of the included models varied between 5 and 18, with a total of 27 distinguishable variables/variable categories among all included risk models. Age was the only variable used in all models. Other common variables were hypertension history or treatment, heart failure history, sex, and blood pressure, incorporated into 16, 16, 14, and 14 of the 21 included models, respectively.

Risk model performance among included cohorts

Supplementary material, Table S4 shows the results on AF incidence, discrimination, and calibration of the included simple models among the cohorts in our search. All studies used the C-statistic to assess model discrimination for incident AF within their cohorts. Nine studies assessed calibration by providing both a P-value for a goodness-of-fit test and a calibration plot in at least one of the risk models that these studies reported on refs,[20,21,24,27,29,30,32,36,40] seven studies assessed calibration only by a P-value for goodness-of-fit test,[22,26,35,39,41,43,44] two studies assessed calibration only by a calibration plot,25,33 and nine studies reported neither of these calibration parameters.[23,28,31,34,37,38,42,45,46]

Table 1. *Characteristics of included risk models developed for incident AF*

Model	ARIC-AF	CHARGE-AF	C_2HEST	FHS-AF	Mayo	MHS	PREVEND	Seirei	Suita	WHS
Model type	Point-based	Cox regression	Point-based	Cox regression	Point-based	Point-based	Latent class analysis	Point-based	Point-based	Cox regression
Intended prediction window for incident AF (years)	10	5	11	5, 10	N/S	10	10	7	10	10
Model variables#										
Age	X	X	X	X	X	X	X	X	X	X
Sex				X	X	X	X	X	X	
Race	X	X					X			
Body measurements (height, weight, BMI)	X	X		X		X	X	X	X	X
Blood pressure (systolic, diastolic)	X	X		X		X	X	X		X
Heart rate							X	X		
Heart failure history	X	X	X	X	X	X	X		X	
Hypertension treatment or history	X	X	X	X	X	X	X		X	
Diabetes mellitus history	X	X			X		X			
Stroke history							X			
CHD or MI history	X	X	X				X			
Vascular disease history							X			
Alcohol use							X	X	X	X
Smoking	X	X					X		X	X
ECG parameters	X			X			X			
COPD			X			X				
Autoimmune or inflammatory disease history						X				
Significant murmur	X			X	X			X	X	
Serum lipids							X		X	
Glomerular filtration rate							X			
Urine albumin secretion							X			
Thyroid disease			X							

AF, atrial fibrillation; ARIC-AF, Atherosclerosis Risk In Communities score for Atrial Fibrillation; BMI, body mass index; CHADS2, Congestive heart failure, Hypertension, Age >75, Diabetes mellitus, prior Stroke or transient ischemic attack [2 points]; CHA_2DS_2-VASc, Congestive HF, Hypertension, Age >75 [2 points], Stroke/transient ischemic attack/thromboembolism [2 points], Vascular disease, Age 65-74, Sex category; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; CHD, coronary heart disease; C2HEST, Coronary artery disease/Chronic obstructive pulmonary disease [2 points], Hypertension, Elderly, Systolic heart failure, Thyroid disease; COPD, chronic obstructive pulmonary disease; ECG, electrocardiogram; FHS-AF, Framingham Heart Study score for Atrial Fibrillation; MHS, Maccabi Healthcare Services; MI, myocardial infarction; N/S, not specified; PREVEND, Prevention of Renal and Vascular End-stage Disease; WHS, Women's Health Study.

* Model not originally developed for incident AF, hence no intended risk prediction interval available for this outcome; # Depicted here are the variables in the simple (non-augmented) models.

Reported C-statistics for incident AF ranged from 0.58 (95% CI 0.55–0.61)[21] to 0.842 (95% CI 0.826–0.858).[44] The highest C-statistic while also showing sufficient calibration was reported in the FHS (Framingham Heart Study) cohort on the in the incidentally used FHS-Lubitz model with a C-statistic of 0.78 (95% CI 0.76–0.80) and P-value of the goodness-of-fit test of 0.11.[44]

Augmentation of included risk models

We identified augmentation data applicable to primary care settings for five of the included AF risk models (see Supplementary material, Table S5). Significant improvement was demonstrated in CHARGE-AF (Cohorts for Heart and Aging Research in Genomic Epidemiology) with addition of the P-wave axis[35] and brain natriuretic peptide (BNP) and/or C-reactive protein (CRP),[41] in the FHS-AF 10-year model with addition of BNP and CRP,[39] and in the Seirei model with addition of ECG parameters to the model.[27]

Risk of bias assessment

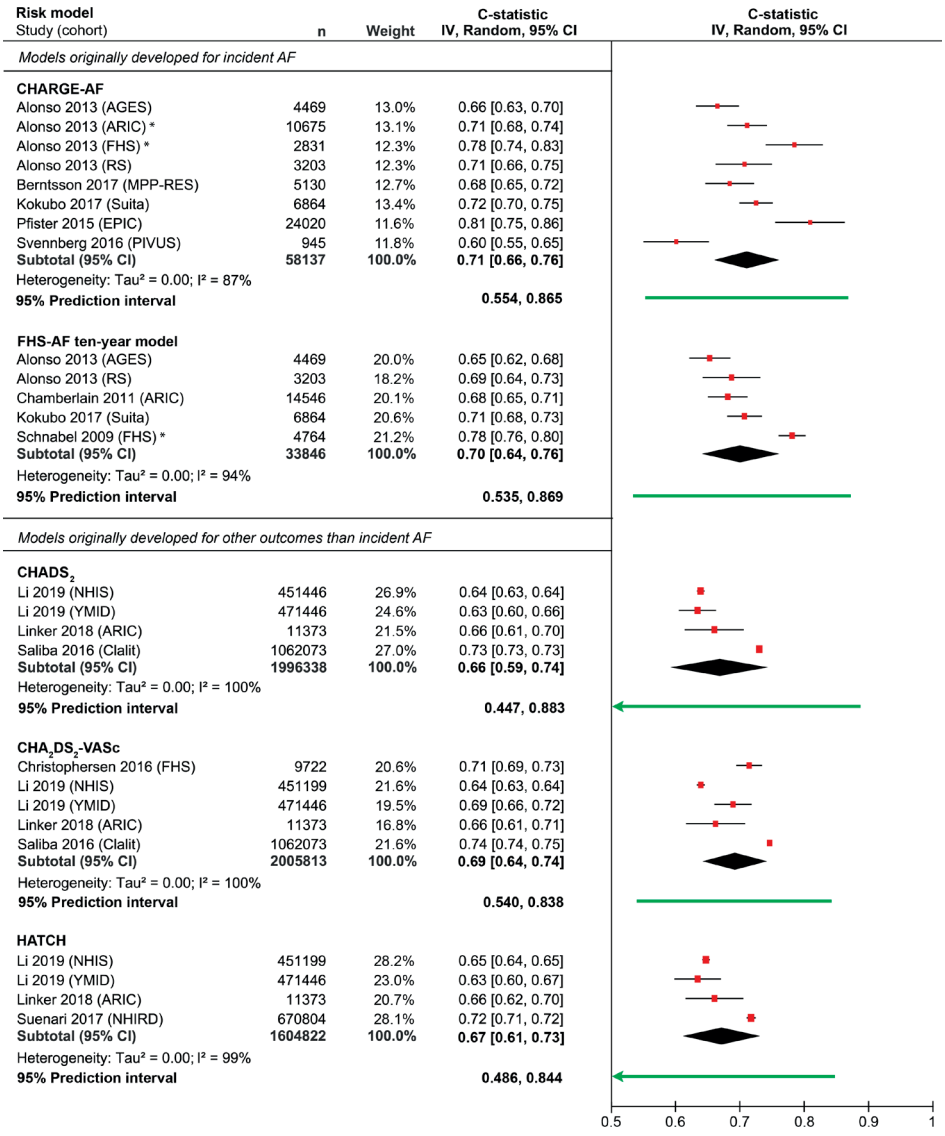
Supplementary material, Table S6 shows the results of the risk of bias assessment for each cohort in the included studies. We assessed the risk of bias of all domains as either low or unclear for all domains except for the participants domain. For this domain, we assessed eight cohorts employed by 10 studies as having a high risk of bias for excluding patients for reasons with a known association with risk of future AF.[20,26,27,29,33,40,41,43,45,46]

Meta-analyses

Five models were eligible for the primary meta-analysis, as shown in Figure 2. Of these, only CHARGE-AF and the FHS-AF for 10-year risk model had originally been derived for incident AF. All primary meta-analyses resulted in high heterogeneity for which we calculated a 95% PI. There were three models that resulted in a summary C-statistic with significant 95% PI in our primary meta-analysis: CHARGE-AF (summary C-statistic 0.71; 95% CI 0.66–0.76; I2 87%; 95% PI 0.554–0.865; n = 8 studies; n = 58 137 patients), the FHS-AF 10-year model (summary C-statistic 0.70; 95% CI 0.64–0.76; I2 94%; 95% PI 0.535–0.869; n = 5 studies; n = 33 846 patients), and CHA_2DS_2-VASc (summary C-statistic 0.69; 95% CI 0.64–0.74; I2 100%; 95% PI 0.540–0.838; n = 5 studies; n = 2 005 813 patients) (see Figure 3 for a comparison of these three models).

For our secondary analysis, we were able to meta-analyse CHARGE-AF and the FHS-AF 10-year model, each for a 5- and 10-year prediction window (Figure 4). Only the meta-analysis of CHARGE-AF with a 5-year prediction window resulted in significant overall discrimination (summary C-statistic 0.72; 95% CI 0.66–0.78; I2 85%; 95% PI 0.567–0.881; n = 6 studies; n = 50 328 patients).

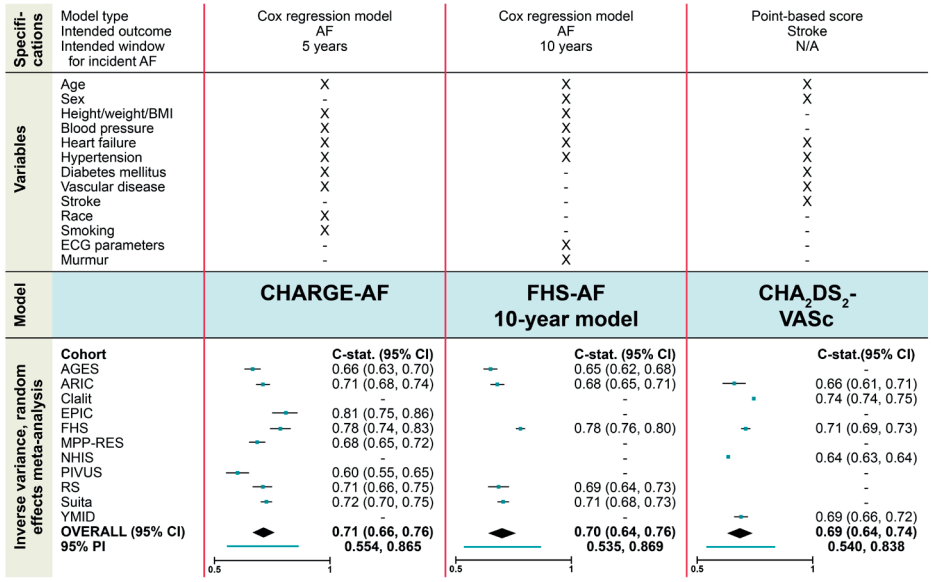
Figure 2. *Primary analysis: meta-analysis of C-statistics.*



AF, atrial fibrillation; AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; $CHADS_2$, Congestive heart failure, Hypertension, Age >75, Diabetes mellitus, prior Stroke or transient ischemic attack [2 points]; CHA_2DS_2-VASc, Congestive HF, Hypertension, Age >75 [2 points], Stroke/transient ischemic attack/thromboembolism [2 points], Vascular disease, Age 65-74, Sex category; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; CI, confidence interval; EPIC, European Prospective Investigation into Cancer and Nutrition; FHS, Framingham Heart Study; FHS-AF, Framingham Heart Study score for Atrial Fibrillation; HATCH, Hypertension, Age, stroke or Transient ischemic attack, Chronic obstructive pulmonary disease, Heart failure; IV, inverse variance; MPP-RES, Malmö Preventive Project Re-examination Study; NHIS, National Health Insurance Service; NHIRD, National Health Insurance Research Database; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; RS, Rotterdam Study; YMID, Yunnan Medical Insurance Database; 95% CI, 95% confidence interval.

* Derivation cohort.

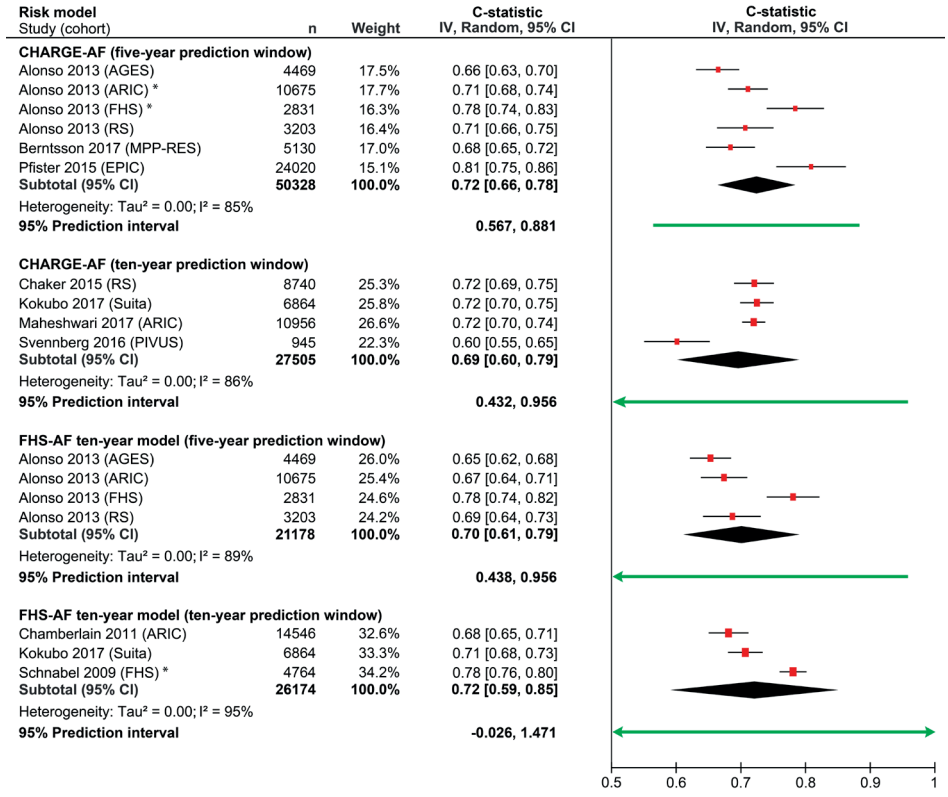
Figure 3. *Comparison of the three models that resulted in significant 95% prediction intervals in the primary meta-analysis.*



AF, atrial fibrillation; AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; BMI, body mass index; CHA_2DS_2-VASc, Congestive HF, Hypertension, Age >75 [2 points], Stroke/transient ischemic attack/thromboembolism [2 points], Vascular disease, Age 65-74, Sex category;
CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; C-stat., C-statistic; ECG, electrocardiogram; EPIC, European Prospective Investigation into Cancer and Nutrition; FHS, Framingham Heart Study; FHS-AF, Framingham Heart Study score for Atrial Fibrillation; MPP-RES, Malmö Preventive Project Re-examination Study; N/A, not applicable; NHIS, National Health Insurance Service; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; RS, Rotterdam Study; YMID, Yunnan Medical Insurance Database; 95% CI, 95% confidence interval; 95% PI, 95% prediction interval.
The meta-analyses of C-statistics for the outcome incident AF are grouped by cohort from which C-statistics were reported, allowing for a comparison of multiple models' performance within one cohort insofar as data are available.

In our sensitivity analysis of restricting primary and secondary analyses models to cohorts with sufficient calibration, we found no model with significant overall discrimination due to high heterogeneity (Supplementary material, Figures S1 and S2). Our sensitivity analysis on double cohorts in the primary and secondary analyses did not lead to different conclusions on overall discriminatory ability of meta-analysed models in all but one comparison (see Supplementary material, Table S7).

Figure 4. *Secondary analysis: meta-analysis of C-statistics grouped according to application of a uniform prediction window within a model.*



AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; CI, confidence interval; EPIC, European Prospective Investigation into Cancer and Nutrition; FHS, Framingham Heart Study; FHS-AF, Framingham Heart Study score for Atrial Fibrillation; IV, inverse variance; MPP-RES, Malmö Preventive Project Re-examination Study; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; RS, Rotterdam Study; 95% CI, 95% confidence interval.

* Derivation cohort.

DISCUSSION

In this systematic review and meta-analysis, we provided an overview of prediction models for incident AF risk that are applicable in and had been derived, validated, and/or augmented in community cohorts. We identified 21 risk models that met these criteria, of which 10 had specifically been derived for predicting AF incidence in the community. In meta-analysis of C-statistics, three models showed significant overall discrimination for AF incidence at any follow-up duration and with any calibration despite high heterogeneity. Two of those models were derived specifically for incident AF risk prediction: CHARGE-AF and the FHS-AF 10-year model. Only CHARGE-AF

showed significant overall discrimination among cohorts with a uniform prediction window (the model's originally intended 5-year window).

Clinical relevance

The outcomes of this systematic review and meta-analysis are highly relevant for the field of primary AF screening. Previous AF screening programmes showed only moderate efficiency in selecting at-risk patients from the community, with an estimated number needed to screen of 111 among 23 studies that had screened community cohorts for incident AF by various methods.[12] Patients were often selected for screening based only on age.[7-9,54] The age criterion in selecting patients for AF screening has its clinical merits since oral anticoagulation in AF patients is indicated in all women ≥65 and all men ≥75 years of age and should be considered in men aged ≥65 years in the absence of other risk factors.[5,55] Age as a criterion, however, should not be considered absolute in selecting patients for primary AF screening. Half of all AF cases detected in the Belgian Heart Rhythm Week were younger than 65 years of age.[56] Moreover, there is evidence that CHARGE-AF has higher discrimination among younger patients, although calibration here was lower due to lower absolute AF risk in this younger subgroup.[36] Finally, the two studies within our search that compared multivariable models with age alone as the predictor both found that the multivariable models had significantly higher C-statistics for incident AF.[26,34] We conclude, therefore, that the use of multivariable risk models in selecting patients for community AF screening is likely to result in more efficient screening than selecting based on age alone. Given that there is adequate stroke prevention therapy available once AF is detected, it is likely that the use of such models in AF screening will result in more efficient stroke prevention.[5] More work on the implementation of multivariable risk models in AF screening as well as on long-term follow-up of screening-detected AF cases, however, is necessary to test these hypotheses.

Whether an immediate start of anticoagulation therapy is warranted when AF is detected in younger patients with risk factors other than high age will subsequently depend on the number and nature of these other risk factors. However, as shown in Table 1, most AF risk prediction models include a multitude of the variables in the CHA_2DS_2-VASc score used to assess anticoagulant treatment indication.[5,55] In a younger patient selected for AF screening based on a multivariable prediction model due to presence of other risk factors than high age, an AF diagnosis is therefore likely to still be relevant in terms of the need for anticoagulant therapy, if not for prevention of other pathology associated with AF such as heart failure.[57]

One of the aims of this work was to determine which model may be best suitable for increasing efficiency of future primary AF screening efforts. Our work showed that there are ample AF risk models to choose from, however with one model that currently stands out between the others: CHARGE-AF. Despite heterogeneity in included cohorts, CHARGE-AF showed significant summary discrimination over a relatively short (5-year) risk prediction window. The model contains variables that are generally easy to extract from health records, and requires only body measurements that are easily obtainable (height, weight, and blood pressure). The FHS-AF model, in contrast, though performing nearly as well in overall discrimination, requires variables 'significant murmur' and ECG variables which are less easy to acquire or interpret for many care professionals. Concluding, CHARGE-AF currently seems the most suitable prediction model for incident AF, and likely has merits as a low-cost triage test for future primary AF screening efforts.

Derivation, validation, and augmentation

In risk models derived for incident AF in community cohorts, there was a trend that the derivation cohort had the highest C-statistic compared with external validation cohorts. The only exception was CHARGE-AF, where Pfister et al.[36] reported a C-statistic of 0.808. Calibration of CHARGE-AF in their cohort, however, was insufficient (P-value for goodness-of-fit test <0.001 and O: E ratio 0.47) due to a systematic overestimation of 5-year AF risk in all risk deciles. One explanation lies in the differences in demographics, as Pfister's cohort was younger and had lower baseline prevalence of diabetes mellitus than the CHARGE-AF derivation cohorts. Depending on whether one's aim is to distinguish high from low-risk patients, or to predict absolute 5-year incident AF risk, a researcher may use this knowledge to decide whether or not to recalibrate a model for his own target population. In augmentation studies, we saw that addition of BNP and CRP to a model seemed most promising in terms of improving risk classification. We note, however, that the significance of an added value of BNP and CRP to CHARGE-AF was not consistent, and that the augmentation studies provided no information on the added costs of augmentation parameters relative to those of acquiring the simple model risk score.29

Previous work

Previous systematic reviews have focused on individual predictors for AF,[58,59] on AF as a risk factor for other outcomes,[60,61] or on risk models for adverse outcomes in AF patients.[62,63] However, to our knowledge, this is the first systematic review and meta-analysis on performance of incident AF risk prediction models, and the first with a focus on such risk models validated in and applicable to the community.

Future work

Future studies could focus on finding optimal cut-offs for the more promising AF prediction models, and to find the most cost-effective use of multivariable models within various screening schemes. Researchers may opt here, e.g. for either a dichotomization into patients with higher and lower risk or assigning patients to one of multiple risk strata. Patients at higher risk could be offered a more intensive, sensitive screening scheme (e.g. multiple-point screening or Holter monitoring) when compared with patients at low risk (single point or no screening). Further research could also assess whether implementation of multivariable models in AF screening could be aided by software that automatically extracts patient data from health records, informs the physician of a patient's current risk category, and suggests parameters that should be updated for a more accurate current AF risk stratification.

Strengths and limitations

This study has a number of strengths. First, we included only studies performed in community cohorts, which contributed to the value of our results for primary care AF screening. Second, we included any risk model that was used to predict risk of incident AF. This enabled us to expand our scope to models that had originally not been intended for predicting incident AF, but that may have merits in predicting this outcome. Third, we attributed high bias to studies that excluded or over-represented patients based on factors that are likely to be associated with risk of incident AF, further contributing to the generalizability of our results to the community. Fourth, we included only the C-statistic of raw, non-bootstrapped data into meta-analysis in order to not bias the meta-analysis with potentially overly narrow confidence intervals. Finally, we refrained from meta-regression or subgroup meta-analysis based on e.g. a subdivision of cohorts' mean age, AF incidence or region, to explain the heterogeneity in our results. Such analyses from aggregate data are known to have a high risk of especially ecological bias and are inferior to subgroup results derived from individual participant data (IPD).[14] An IPD meta-analysis, however, was not the scope of the current study.

The primary limitation of our study is the high heterogeneity of included studies. We attempted to cope with this limitation by performing sensitivity analyses and by calculating a 95% PI in our meta-analyses with high heterogeneity. The outcomes of our meta-analyses with significant 95% PI can be considered generalizable to such populations as included into those meta-analyses, despite high heterogeneity. As a second limitation, we did not provide a meta-analysis on model calibration since such analyses are often challenging due to a lack of calibration measures reported among studies.[14] Indeed, we found that meta-analysable data on calibration was poorly

reported on among included studies (Supplementary material online, Table S4). Moreover, summarizing O:E data would have automatically excluded those models that were not originally intended for incident AF, since expected incident AF rates would never have been defined for such models. We addressed calibration by performing a sensitivity analysis among cohorts which had demonstrated sufficient calibration by their applied risk model(s). A third limitation is that we included both prospective and retrospective cohort studies. This may have introduced bias as AF is not always symptomatic[64] and asymptomatic patients are less likely to undergo rhythm evaluation when left to their physicians' discretion than when ECG is performed in the context of a prescheduled follow-up. The restriction of our search to studies written in English which we applied for quality-related as well as practical reasons, finally, has been found not to lead to significant bias.[65]

CONCLUSION

We provided an overview of prediction models for incident AF risk that are applicable in and have been derived, validated, and/or augmented in community cohorts. We identified 21 risk models that met these criteria. Of these, CHARGE-AF seemed the most robust in terms of performance as well as applicability in the community.

Funding: This work was supported by the Netherlands Organisation for Health Research and Development (ZonMw) [80-83910-98-13046]. The authors had full autonomy in design, conduct, and reporting of the manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 648131) and German Center for Cardiovascular Research (DZHK e.V.) (81Z1710103). Salary support for R.E.H. was provided by a Rubicon Fellowship grant by the Netherlands Organisation for Scientific Research (NWO) (grant No. 452173116).

Conflict of interest: None declared.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY METHODS

Applied interpretation of CHARMS checklist items in risk of bias assessment

A number of Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) domains differentiate between applicability to the study aims and risk of bias.[15] We simultaneously assessed applicability to our study aims and risk of bias to derive one score for risk of bias for each CHARMS domain. We did not score risk of bias for the Sample size domain since we could find no consensus or documentation on which to base the criteria 'low', 'unclear' or 'high' risk of bias for sample size and outcome incidence. For our interpretation of the other 10 CHARMS criteria, see below:

Source of data:

- Low: prospective cohort study in primary care/general population;
- Unclear: retrospective cohort study in primary care/general population;
- High: cohort collected from secondary care, whether in- or outpatients, whether prospective or retrospective.

Participants:

- Low: authors applied no criteria for patient inclusion that could have affected the comparability to the target population (primary care/general population);
- Unclear: authors applied criteria for patient inclusion that may have affected the comparability to the target population (primary care/general population);
- High: authors applied criteria for patient inclusion that have with high probability affected the comparability to the target population (primary care/general population), including but not limited to exclusion of patients with heart failure, high cardiovascular risk, or diabetes.

Outcome:

- Low: AF was identified using ECG recordings or usual care data (applicability); there was systematic follow-up of participants for the outcome AF (each participant had the same probability of receiving follow-up ECG or rhythm monitoring);
- Unclear: there was non-systematic follow-up of participants (follow-up ECG or rhythm monitor at the physician's discretion);
- High: AF was identified by other means, e.g. by patient questionnaire, and not by using ECG recordings or usual care data.

Predictors:

- Low: predictors used in the multivariable model were well defined and applicable to primary care settings;
- Unclear: one or more predictors used in the multivariable model were unclearly defined;
- High: one or more predictors used in the multivariable model were not applicable to primary care settings.

Missing data:

- Low: percentage of missing data reported was ≤10% AND missing variables were imputed and imputation method was well documented OR reasons for missing data were documented and likely not associated with risk of AF;
- Unclear: >10% of patients were excluded for missing data and/or it was unclear if reasons for missing data were associated with risk of AF;
- High: >10% of patients were excluded for missing data and/or exclusion reasons for missing data were likely associated with risk of AF.

Model development:

- Low: authors reported on all items on the development of the prediction model as stated in the CHARMS checklist;
- Unclear: the weight of ≥1 variable in the multivariable prediction model remained unknown or unclear;
- N/A: authors performed a validation of an existing model, however no derivation.

Model evaluation:

- Low: authors reported on all items on the evaluation of the prediction model as stated in the CHARMS checklist; authors validated the applied risk model using the intended follow-up duration;
- Unclear: the prediction model was recalibrated, however authors did not report the new weight of ≥1 variable; authors validated the applied risk model using a follow-up duration other than the intended follow-up duration;
- N/A: authors performed a derivation of a new model, however no validation.

Model performance:

- Low: authors provided a clear assessment of both discrimination and calibration;
- Unclear: authors provided a clear assessment of either discrimination or calibration;
- High: authors provided a clear assessment of neither discrimination nor calibration.

Results:

- Low: authors provided a clear representation of the results, with all elements within the methods section addressed;
- Unclear: authors provided all elements within the methods section addressed, however some of the results were unclearly reported;
- High: authors did not address all elements within the methods section and/or most results were unclearly reported.

Discussion:

- Low: authors addressed all relevant discussion items carefully, and in accordance with the results from their study;
- Unclear: authors did not address all relevant discussion items carefully, however the discussed topics were in accordance with the results from their study;
- High: one or more of the discussed topics were not in accordance with the results from the authors' study.

SUPPLEMENTARY TABLES

Supplementary Table S1. *Search strategy*

PubMed search (1 August 2019)		
#	Search	Hits
#1	((risk score[tiab] OR decision support techniques[mesh] OR prediction model*[tiab] OR decision aid*[tiab] OR clinical prediction rule*[tiab] OR decision model*[tiab] OR risk prediction model*[tiab] OR risk-scoring system*[tiab] OR risk model*[tiab] OR prediction aid[tiab] OR prediction tool) AND "english"[Filter] AND "humans"[Filter])	88,627
#2	((atrial fibrillation OR atrial fibrillation[MeSH Terms])AND "english"[Filter] AND "humans"[Filter])	54,581
#3	((atrial flutter OR atrial flutter[MeSH Terms])AND "english"[Filter] AND "humans"[Filter])	6,219
#4	((#2 OR #3) AND #1AND "english"[Filter] AND "humans"[Filter])	979
Ovid/Embase search (1 August 2019)		
#	Search	Hits
#1	(atrial fibrillation or atrial flutter).af.	139,002
#2	(risk score or decision support techniques or prediction model or decision aid or clinical prediction rule or decision model or risk prediction model or risk-scoring system or risk model or prediction aid or prediction tool).af.	56,627
#3	1 and 2	1,811
#4	Limit 3 to (human and English language)	1,643
CINAHL search (1 August 2019)		
#	Search	Hits
#1	MH atrial fibrillation OR TX atrial fibrillation OR MH atrial flutter OR TX atrial flutter	40,109
#2	MH risk score OR TX risk score OR MH decision support techniques OR TX decision support techniques OR MH prediction model OR TX prediction model OR MW decision aid OR TX decision aid OR MH risk model OR TX risk model OR MW clinical prediction rule OR TX clinical prediction rule OR MH prediction tool OR TX prediction tool	61,117
#3	S1 AND S2	2,093
#4	S3 - Restricting to English	2,092

Supplementary Table S2. *Characteristics of included studies and cohorts*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Alonso 2013	ARIC (USA)	D	Randomly selected general population participants from 4 centers in the US, those attending the 4th examination cycle (P)	10,675	63 ± 6	56.7	28.8 ± 5.6	13.7	36.0	4.6	AF or Afl, identified by examinations every 3 years or medical records	1996-2005 (N/A)
	CHS (USA)	D	Men and women 65 years or older from 4 US communities with an oversampling of African Americans to increase minority representation (P)	5,043	73 ± 5	59.8	26.7 ± 4.7	16.3	46.3	3.8	AF or Afl, identified by annual examinations or medical records	1989-2005 (N/A)
	FHS (USA)	D	Offspring from the original FHS cohort, those attending the 6th examination cycle (P)	2,838	60 ± 8	54.5	27.9 ± 5.1	10.2	29.6	0.6	AF or Afl, identified by examinations every 4 to 8 years or medical records	1995-2005 (N/A)
	AGES (IS)	EV	Men and women born between 1907 and 1935 living in the greater Reykjavik area in 1967-1996 (P)	4,469	76 ± 6	60.4	27.1 ± 4.4	11.5	61.2	1.7	AF or Afl, identified by medical records	2002-2011 (N/A)
	RS (NL)	EV	Men and women aged 55 years and older living in the Rotterdam suburb of Ommoord in 1989-1993 (P)	3,203	72 ± 7	58.9	26.9 ± 3.9	10.2	37.1	3.5	AF or Afl, identified by examination every 3-5 years or medical records	1997-2005 (N/A)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Alonso 2016	MESA (USA)	EV	Men and women, 45 to 84 years of age, free of clinical cardiovascular disease from 6 communities across the US (P)	6,663	62 ± 10	52.8	28.7	12.5	36.9	0	AF or Afl, identified by hospitalizations	2000-2012 (10.2)
Aronson 2018	MHS (IL)	D	Men and women aged > 50 years randomly selected from ambulatory clinics database with data available prior to 1 January 2005 (R)	96,778	62 ± 9	53.7	28.2 ± 5.1	13.5	34.3	1.0	AF or Afl, identified by medical records	2005-2015 (10.0)
		EV	Men and women aged > 50 years randomly selected from ambulatory clinics database with data available prior to 1 January 2005 (R)	48,404	62 ± 9	54.1	28.1 ± 5.1	13.4	34.1	0.9	AF or Afl, identified by medical records	2005-2010 (10.0)
Berntsson 2017	MPP-RES (SE)	EV	Men born in 1921, 1926–1942, 1944, 1946 and 1948–1949 and women born in 1926, 1928, 1930–1936, 1938, 1941–1942 and 1949 in Malmö, Sweden (P)	5,130	69.2 ± 6.2	30.9	27.2	11.4	37.4	1.0	AF, identified by hospitalizations	2002-2010 (5.6)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Chaker 2015	RS (NL)	EV	Men and women aged 55 years and older living in the Rotterdam suburb of Ommoord in 1989-1993 (P)	8,740	65.0 ± 9.9	56.7	27.1 ± 4.1	8.4	51.7	N/A	AF or Afl, identified by examination every 3-5 years or medical records	1997-2012 (6.8)
Chamberlain 2011	ARIC (USA)	D	Randomly selected general population participants from 4 communities in the US (P)	14,546	54.6	55.3	N/A	12.0	30.6	4.6	AF or Afl, identified by triennial follow-up examinations, annual phone calls or medical records	1987-1998 (N/A)
Christophersen 2016	FHS (USA)	EV	Participants aged 45-95 from the original FHS cohort attending the 23rd, 26th and 29th examination, and offspring from the original FHS cohort attending the 5th, 7th and 8th examination (P)	4,548*	63.9 ± 10.6	56	27.9	11	51	1	AF or Afl, identified by scheduled follow-up examinations or medical records	1991-2013 (6.6)
Everett 2013	WHS (USA)	D, IV	American female health professionals enrolled in the Women's Genome Health Study, a subset of the WHS, aged 45 or older with no CVD or HF at baseline (P)	20,822	52.9	100	24.8	2.4	24.0	0	AF, identified by annual telephonic self-report, confirmed by medical records	1992-N/A (14.5)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Hamada 2019	Seirei (JP)	D, IV	Participants aged 40-79 years who underwent check-ups to identify non-communicable disease and its risk factors from April 2008 to March 2014 (R)	65,984	52.4	34.9	22.8	N/A	N/A	N/A	AF or Afl, identified by follow-up examinations, medical records or self-report	2008-2015 (5.5)
Kokubo 2017	Suita (JP)	D	Participants aged 30-79 years randomly selected in 1989 and 1996, as well as a volunteer group, from the municipality population registry of Suita City, Japan (R)	6,864	55.7	53.0	22.6	4.9	12.2	0.1	AF or Afl, identified by follow-up examinations or medical records	1989-2015 (13.9)
Kumarathurai 2017	CopHS (DK)	EV	Patients 55-75 years of age from two postal regions in Copenhagen with no history of prior AF, stroke, or CVD who self-reported ≥2 cardiovascular risk factors (P)	646	64.4 ± 6.8	41.6	26.3	11.1	28.2	N/A	AF, identified by medical records	N/A (14.4)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Li 2019	YMID (CN)	D, IV	Random sample from medical insurance database including >10 million individuals (R)	471,446	47.0	47.3	N/A	4.0	9.7	0.1	AF in hospital admission records	2001-2012 (4.1)
	NHIS (KR)	EV	Random sample of participants in screening programs by National Health Insurance Service (R)	451,199	56.1 ± 9.3	46.0	N/A	8.3	31.7	1.2	AF, identified by admission and/ or outpatient records	2002-2013 (7.3)
Linker 2018	ARIC (USA)	EV	Randomly selected general population participants from 4 centers in the US, baseline is 3rd examination (P)	11,373	60.0 ± 5.7	N/A	N/A	N/A	N/A	N/A	AF or Afl, identified by triennial follow-up examinations, annual phone calls or medical records	1996-2005 (N/A)
Lubitz 2010	FHS (USA)	IN, AU	Participants aged 45-95 from the original FHS cohort attending the 11th, 18th, 22nd and 26th examination, and offspring from the original FHS cohort attending the 1st, 3rd, 5th and 7th examination (P)	4,421†	53.9 ± 13.3	54	26.9	N/A	20	0.6	AF or Afl, identified by scheduled follow-up examinations or medical records	1968-2007 (N/A)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Magnani 2015	ARIC (USA)	IN, AU	Randomly selected general population participants from 4 centers in the US, baseline is 4th examination (P)	8,254	62.3 ± 5.6	57.3	28.3	13.9	N/A	0.6	AF or Afl, identified by triennial follow-up examinations, annual phone calls or medical records	1986-N/A (N/A)
	FHS (USA)	IN, AU	Participants aged 45-95 from the original FHS cohort attending the 20th, offspring from the original FHS cohort attending the 6th, and third generation cohort attending the 1st examination (P)	3,110	62.6 ± 9.8	56.9	27.2	8.1	N/A	0.3	AF or Afl, identified by scheduled follow-up examinations or medical records	1987-N/A (N/A)
Maheshwari 2017	ARIC (USA)	EV	Randomly selected general population participants from 4 centers in the US, baseline is 4th examination (P)	15,320	54.2 ± 5.7	55.2	27.7	11.8	30.5	4.6	AF or Afl, identified by triennial follow-up examinations, annual phone calls or medical records	1987-2013 (20.2)
Pfister 2015	EPIC (UK)	EV	Men and women aged 39-79 years, living in Norfolk, UK, included in 1993-1997 (P)	24,020	58.6	55	26.4	2.2	18.1	0.5	AF, using hospital record linkages	1993-2015 (12.5)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Rienstra 2014	FHS (USA)	IN, AU	Offspring from the original FHS cohort, those attending the 6th examination cycle (P)	3,217	59 ± 10	54	27.9	11	27	1	AF or Afl, identified by follow-up examinations or medical records	1995-2008 (N/A)
Rienstra 2016	FHS (USA)	V	Offspring from the original FHS cohort, those attending the 6th examination cycle (P)	3,162	58 ± 9	53.7	N/A	9.0	26.5	0.5	AF or Afl, identified by follow-up examinations or medical records	1995-2008 (N/A)
	PREVEND (NL)	D	Persons with UAE >10 mg/l and a random sample of persons UAE <10 mg/l who responded after inviting all inhabitants aged 28-75 years from Groningen, NL, included from 1997	8,265	49 ± 13	50.2	N/A	3.8	13.3	0.2	AF or Afl, identified by follow-up examinations or medical records	1997-N/A (9.2)
Rosenberg 2012	CHS (USA)	EV, AU	Men and women 65 years or older from 4 US communities with an oversampling of African Americans to increase minority representation (P)	5,117	72.9	59.1	26.7	16.4	66.0	N/A	AF or Afl, identified by annual examinations or medical records	1989-N/A (12.2)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Saliba 2016	ClaHS (IL)	EV	All adult subjects ages 50 years or older on January 1, 2012, from a not-for-profit health care provider covering more than half of the Israeli population (R)	1,062,073	65.7 ± 11.2	54.7	N/A	25.3	48.9	4.3	AF, identified by medical records	2012-2014 (2.9)
Schnabel 2009	FHS (USA)	D	Participants aged 45-95 from the original FHS cohort attending the 11th or 17th examination, and offspring from the original FHS cohort attending the 1st or 3rd examination (P)	4,764#	60.9 ± 9.9	55	26.3 ± 4.3	14	24	1	AF or Afl, identified by follow-up examinations or medical records	1968-2007 (9.2)
Schnabel 2010a	FHS (USA)	IV	Offspring from the original FHS cohort, those attending the 6th examination cycle (P)	3,120	58.4	54	28.0	N/A	27.4	0.5	AF or Afl, identified by follow-up examinations or medical records	1995-2007 (9.7)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Schnabel 2010b	AGES (IS)	EV	Men and women born between 1907 and 1935 living in the greater Reykjavik area in 1967-1996 (P)	4,238	76.3 ± 5.5	62.7	27.0 ± 4.5	N/A	59.8	1.3	AF or Afl, identified by medical records	2002-2011 (4.2)
	CHS (USA)	EV	Men and women 65 years or older from 4 US communities with an oversampling of African Americans to increase minority representation (P)	5,410[§]	75.1	60.1	26.7	N/A	41.9	5.5	AF or Afl, identified by annual examinations or medical records	1989-2005 (4.4)
	FHS (USA)	EV	Participants aged 45-95 from the original FHS cohort attending the 11th or 17th examination, and offspring from the original FHS cohort attending the 1st or 3rd examination (P)	4,764[#]	60.9 ± 9.9	55.4	26.3 ± 4.3	N/A	24	1.0	AF or Afl, identified by follow-up examinations or medical records	1968-1992 (4.8)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Sinner 2014	ARIC (USA)	AU	Randomly selected general population participants from 4 centers in the US, those attending the 4th examination cycle (P)	10,675	63 ± 8	55.6	28.8 ± 5.6	16.2	42.5	4.7	AF or Afl, identified by triennial follow-up examinations, annual phone calls or medical records	1996-2005 (N/A)
	CHS (USA)	AU	Men and women 65 years or older from 4 US communities with an oversampling of African Americans to increase minority representation (P)	5,043	73 ± 5	59.9	26.7 ± 4.7	16.3	46.3	3.7	AF or Afl, identified by annual examinations or medical records	1989-2000 (N/A)
	FHS (USA)	AU	Offspring from the original FHS cohort, those attending the 6th examination cycle (P)	2,838	60 ± 8	54.6	27.9 ± 5.1	10.2	29.6	0.6	AF or Afl, identified by examinations every 4 to 8 years or medical records	1995-2005 (N/A)
	AGES (IS)	AU	Men and women born between 1907 and 1935 living in the greater Reykjavik area in 1967-1996 (P)	4,467	76 ± 6	60.4	27.1 ± 4.4	11.5	61.2	1.7	AF or Afl, identified by medical records	2002-2011 (N/A)
	RS (NL)	AU	Men and women aged 55 years and older living in the Rotterdam suburb of Ommoord in 1989-1993 (P)	3,203	72 ± 7	58.9	26.9 ± 3.9	10.2	37.1	3.5	AF or Afl, identified by examination every 3-5 years or medical records	1997-2005 (N/A)

Supplementary Table S2. *Characteristics of included studies and cohorts (continued)*

Study	Cohort ID (CY)	Study aim	Setting (study design)	n	Age (mean ± SD)	Female (%)	BMI (mean ± SD)	DM (%)	HT (%)	HF (%)	Outcome measure	Enrolment period (mean F/U in years)
Suenari 2017	NHIRD (TW)	EV	Random sample from mandatory health insurance program that offers health insurance coverage to all Taiwanese residents (R)	670,840	42.4 ± 16.0	49.1	N/A	3.2	5.5	0.4	AF, identified by medical records	2000-2011 (9.0)
Svennberg 2016	ULSAM (SE)	EV	Men born between 1920 and 1924 living in Uppsala, Sweden, in 1991-1995 (P)	883	71	0	26.2 ± 3.4	10.3	32.7	N/A	AF, identified by medical records	1991-2001 (12.6)
	PIVUS (SE)	EV	Individuals aged 70 living in Uppsala, Sweden, cohort initiated in 2001 (P)	978	70	50	27 ± 4.3	11.7	30.8	0.6	AF, identified by medical records	2001-2011 (10.0)

AF, atrial fibrillation; Afl, atrial flutter; AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; AU, augmentation; CHS, Cardiovascular Health Study; ClaHS, Clalit Health Service; CN, China; CopHS, Copenhagen Holter Study; CVD, cardiovascular disease; CY, country; D, derivation; DK, Denmark; DM, diabetes mellitus; EMR, electronic medical records; EPIC, European Prospective Investigation into Cancer and Nutrition; EV, external validation; FHS, Framingham Heart Study; F/U, follow-up; HF, heart failure; HK, Hong Kong; HT, hypertension; ID, identity; IL, Israel; IN, incidental use of a risk model, with variables and coefficients similar to but differing from a risk model previously derived or validated for incident AF; IS, Iceland; IT, Italy; IV, internal validation; JP, Japan; KR, Republic of Korea; MESA, Multi-Ethnic Study of Atherosclerosis; MHS, Maccabi Healthcare Services; MPP-RES, Malmö Preventive Project Re-examination Study; NHIRD, National Health Insurance Research Database; NHIS, National Health Insurance Service; NL, Netherlands; N/A, not available; P, prospective study design; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; PREVEND, Prevention of Renal and Vascular End-stage Disease; R, retrospective study design; RS, Rotterdam Study; SD, standard deviation; SE, Sweden; TW, Taiwan; UK, United Kingdom; ULSAM, Uppsala Longitudinal Study of Adult Men; USA, United States of America; YMID, Yunnan Medical Insurance Database.

* In Christophersen 2016 FHS patients were eligible for inclusion into multiple five-year observation windows if they had not been censored and did not meet exclusion criteria at the end of an observation window. The study characteristics displayed in this table are from a total of 9,722 person-intervals from 4,548 unique participants; † In Lubitz 2010 FHS patients were eligible for inclusion into a second eight-year observation window if they had >16 years follow-up data and were free of AF at the end of the first eight-year observation window. The study characteristics displayed in this table are from a total of 11,971 person-intervals from 4,421 unique participants; # In Schnabel 2009 & Schnabel 2010b FHS patients were eligible for inclusion into a second ten-year observation window if they had >20 years follow-up data and were free of AF at the end of the first ten-year observation window. The study characteristics displayed in this table are from a total of 8,044 person-intervals from 4,764 unique participants; § In Schnabel 2010b CHS patients were eligible for inclusion into a second ten-year observation window if they had >20 years follow-up data and were free of AF at the end of the first ten-year observation window. The study characteristics displayed in this table are from a total of 9,766 person-intervals from 5,410 unique participants.

Supplementary Table S3. *Characteristics of included risk models that had been developed for other outcomes than incident AF or that had only incidentally been used for incident AF prediction*

Model	Models developed for outcomes other than incident AF							Models incidentally used for incident AF prediction			
	ARIC-CHD	**$CHADS_2$**	**CHA_2DS_2-VASc**	**FHS-CHD**	**FHS-hCHD**	**HATCH**	**SAAFE**	**CHARGE-Magnani**	**CHARGE-Rienstra 2014**	**CHARGE-Rienstra 2016**	**FHS-Lubitz**
Model type	Cox regression	Point-based	Point-based	Point-based	Cox regression	Point-based	Binary logistic regression	Cox regression	Cox regression	Cox regression	Cox regression
Originally derived for (outcome)	CHD	Ischemic stroke	Ischemic stroke	CHD	hCHD	pAF progression	Prevalent AF	Incident AF	Incident AF	Incident AF	Incident AF
Intended prediction window for incident AF (years)	N/A*	N/A*	N/A*	N/A*	N/A*	N/A*	N/A*	10	10	10	8
	Model variables#										
Age	X	X	X	X	X	X	X	X	X	X	X
Sex	X		X	X	X			X	X	X	X
Race	X						X	X		X	
Body measurements (height, weight, BMI)							X	X	X	X	X
Blood pressure (systolic, diastolic)	X			X	X			X	X	X	X
Heart rate								X			
Heart failure history		X	X			X	X	X	X	X	X
Hypertension treatment or history	X	X	X		X	X			X	X	X
Diabetes mellitus history		X	X	X			X	X	X	X	
Stroke history		X	X			X	X				
CHD or MI history							X	X	X	X	
Cardiac arrest history							X				
Vascular disease history			X								
Kidney transplant history							X				
Smoking	X			X	X			X	X	X	
ECG parameters								X			X
COPD						X	X				
Significant murmur											X
Serum lipids	X			X	X			X			
Serum BNP									X		
Serum CRP									X		
Recent hospitalization for cardio-vascular or pulmonary diagnosis							X				

AF, atrial fibrillation; BMI, body mass index; BNP, brain natriuretic peptide; $CHADS_2$, Congestive heart failure, Hypertension, Age >75, Diabetes mellitus, prior Stroke or transient ischemic attack [2 points]; CHA_2DS_2-VASc, Congestive HF, Hypertension, Age >75 [2 points], Stroke/transient ischemic attack/thromboembolism [2 points], Vascular disease, Age 65-74, Sex category; CHARGE-Magnani, model based on CHARGE-AF with adjusted coefficients as used in Magnani 2015; CHARGE-Rienstra 2014, model based on CHARGE-AF with adjusted coefficients as used in Rienstra 2014; CHARGE-Rienstra 2016, model based on CHARGE-AF with adjusted coefficients as used in Rienstra 2016; CHD, coronary heart disease; COPD, chronic obstructive pulmonary disease; ECG, electrocardiogram; FHS-CHD, Framingham Heart Study score for Coronary Heart Disease; FHS-hCHD, Framingham Heart Study score for hard Coronary Heart Disease; FHS-Lubitz, model based on FHS-AF with adjusted coefficients as used in Lubitz 2010; HATCH, Hypertension, Age, stroke or Transient ischemic attack, Chronic obstructive pulmonary disease, Heart failure; hCHD, hard coronary heart disease; MI, myocardial infarction; N/A, not applicable; pAF, paroxysmal AF; SAAFE, Screening for Asymptomatic Atrial Fibrillation Events; WHS, Women's Health Study.
* Model not originally developed for incident AF, hence no intended risk prediction interval available for this outcome; # Depicted here are the variables in the simple (non-augmented) models.

Supplementary Table S4. *Outcomes of studies reporting on included simple (non-augmented) prediction models*

Prediction Model	Study aim	Study (cohort)	Observed AF/total population (%)	Discrimination		Calibration		Risk prediction interval (years)	Overall RoB
				C-statistic	95%CI	P-value of GOF test	O:E ratio		
ARIC-AF[§]	D	Chamberlain 2011 (ARIC)	515/14,546 (3.5)	0.78	0.75-0.81[#]	0.67	0.99	10	U
	V	Chamberlain 2011* (ARIC)	515/14,546 (3.5)	0.77	0.75-0.78	N/A	N/A	10	U
		Kokubo 2017 (Suita)	311/6,898 (4.5)	0.712	0.672-0.752	N/A	N/A	10	U
		Linker 2018 (ARIC)	165/11,373 (1.5)	0.762	0.727-0.797	N/A	N/A	3	U
ARIC-CHD[†]	V	Chamberlain 2011 (ARIC)	515/14,546 (3.5)	0.58	0.55-0.61[#]	N/A	N/A	10	U
$CHADS_2$[†]	V	Alonso 2016 (MESA)	351/6,663 (5.3)	0.671	0.628-0.714	N/A	N/A	10	H
		Li 2019 (NHIS)	12,143/451,199 (2.7)	0.637	0.632-0.642	N/A	N/A	11	U
		Li 2019 (YMID)	921/471,446 (0.2)	0.632	0.604-0.660	N/A	N/A	11	U
		Linker 2018 (ARIC)	165/11,373 (1.5)	0.658	0.613-0.702	N/A	N/A	3	U
		Saliba 2016 (ClaHS)	23,223/1,062,073 (2.2)	0.728	0.725-0.731	N/A	N/A	3	U
CHA_2DS_2-VASc[†]	V	Alonso 2016 (MESA)	351/6,663 (5.3)	0.695	0.654-0.735	N/A	N/A	10	H
		Christophersen 2016 (FHS)	752/9,722 (7.7)	0.712	0.693-0.731	<0.001	0.90	5	U
		Li 2019 (NHIS)	12,143/451,199 (2.7)	0.637	0.632-0.642	N/A	N/A	11	U
		Li 2019 (YMID)	921/471,446 (0.2)	0.687	0.659-0.716	N/A	N/A	11	U
		Linker 2018 (ARIC)	165/11,373 (1.5)	0.66	0.615-0.706	N/A	N/A	3	U
		Saliba 2016 (ClaHS)	23,223/1,062,073 (2.2)	0.744	0.741-0.747	N/A	N/A	3	U

Supplementary Table S4. *Outcomes of studies reporting on included simple (non-augmented) prediction models (continued)*

Prediction Model	Study aim	Study (cohort)	Observed AF/total population (%)	Discrimination		Calibration		Risk prediction interval (years)	Overall RoB
				C-statistic	95%CI	P-value of GOF test	O:E ratio		
CHARGE-AF	D	Alonso 2013 (ARIC)	419/10,675 (3.9)	0.71	0.68-0.74	0.67	0.99[$]	5	L
		Alonso 2013 (CHS)	624/5,043 (12.4)	0.70	0.68-.073	0.045		5	H
		Alonso 2013 (FHS)	143/2,838 (5.0)	0.78	0.74-0.83	0.11		5	L
	V	Alonso 2013 (AGES)	408/4,469 (9.1)	0.664	0.632-0.697	0.18	N/A[$]	5	L
		Alonso 2013 (RS)	177/3,202 (5.5)	0.705	0.663-0.747	0.06	N/A[$]	5	L
		Alonso 2016 (MESA)	115/6,663 (1.7)	0.779	0.744-0.814	0.002	0.72	5	H
		Berntsson 2017 (MPP-RES)	362/5,130 (7.1)	0.683	0.65-0.72[#]	0.051	1.26	5	U
		Chaker 2015 (RS)	403/8,740 (4.6)	0.722	0.69-0.75[#]	N/A	N/A	10	U
		Christophersen 2016 (FHS)	752/9,722 (7.7)	0.757	0.741-0.762	0.69	0.98	5	U
		Kokubo 2017 (Suita)	311/6,898 (4.5)	0.724	0.699-0.749	N/A	N/A	10	U
		Linker 2018 (ARIC)	165/11,373 (1.5)	0.775	0.741-0.810	N/A	N/A	3	U
		Maheshwari 2017 (ARIC)	810/10,956 (7.4)	0.719	0.702-0.736	0.05	N/A	10	U
		Pfister 2015 (EPIC)	236/24,020 (1.0)	0.808	0.75-0.85	<0.001	0.47	5	U
		Sinner 2014 (ARIC+CHS+FHS)	1,186/18,556 (6.4)	0.765	0.748-0.781	N/A	N/A	5	H
		Svennberg 2016 (PIVUS)	145/945 (15.3)	0.60	0.55-0.65	0.578	N/A	10	U
		Svennberg 2016 (ULSAM)	113/883 (12.8)	0.62	0.57-0.66	0.276	N/A	10	H
CHARGE-Magnani	IN	Magnani 2015 (ARIC)	458/8,254 (5.5)	0.71	0.69-0.73	N/A	N/A	10	H
	IN	Magnani 2015 (FHS)	217/3,110 (7.0)	0.78	0.75-0.80	N/A	N/A	10	H

Supplementary Table S4. *Outcomes of studies reporting on included simple (non-augmented) prediction models (continued)*

Prediction Model	Study aim	Study (cohort)	Observed AF/total population (%)	Discrimination		Calibration		Risk prediction interval (years)	Overall RoB
				C-statistic	95%CI	P-value of GOF test	O:E ratio		
CHARGE-Rienstra 2014	IN	Rienstra 2014 (FHS)	242/3,217 (7.5)	0.803	0.777-0.830	N/A	N/A	10	U
CHARGE-Rienstra 2016	IN	Rienstra 2016 (FHS)	212/3,162 (6.7)	0.725	0.690–0.760	N/A	N/A	10	U
	IN	Rienstra 2016 (PREVEND)	250/8,265 (3.0)	0.842	0.820-0.864	N/A	N/A	10	U
C_2HEST	D	Li 2019 (YMID)	921/471,446 (0.2)	0.750	0.730-0.771	0.774	N/A	11	U
	V	Li 2019 (NHIS)	12,143/451,199 (2.7)	0.654	0.649-0.659	N/A	N/A	11	U
		Li 2019 (YMID)*	921/471,446 (0.2)	0.749	0.729-0.769	N/A	N/A	11	U
FHS-AF *Original ten-year model*	D	Schnabel 2009 (FHS)	457/4,764 (9.6)	0.78	0.76–0.80	0.09	1.00	10	U
	V	Alonso 2013 (AGES)	408/4,469 (9.1)	0.652	0.621-0.684	0.20	N/A	5	U
		Alonso 2013 (ARIC)	419/10,675 (3.9)	0.67	0.64-0.71	0.0002	N/A	5	U
		Alonso 2013 (CHS)	624/5,043 (12.4)	0.66	0.64-0.69	<0.0001	N/A	5	H
		Alonso 2013 (FHS)	143/2,838 (5.0)	0.78	0.74-0.82	<0.0001	N/A	5	U
		Alonso 2013 (RS)	177/3,202 (5.5)	0.686	0.642-0.729	0.49	N/A	5	U
		Alonso 2016 (MESA)	351/6,663 (5.3)	0.746	0.720-0.771	<0.0001	0.70	10	H
		Chamberlain 2011 (ARIC)	515/14,546 (3.5)	0.68	0.65-0.71#	N/A	N/A	10	U
		Kokubo 2017 (Suita)	311/6,898 (4.5)	0.706	0.680-0.731	N/A	N/A	10	U
		Kumarathurai 2017 (CopHS)	31/646 (4.8)	0.656	0.609-0.788	N/A	0.90	10	H
		Linker 2018 (ARIC)	165/11,373 (1.5)	0.626	0.578-0.673	N/A	N/A	3	U
		Rosenberg 2012* (CHS)	1,252/5,117 (24.5)	0.649	0.63-0.67#	N/A	N/A	10	H
		Schnabel 2009* (FHS)	457/4,764 (9.6)	0.76	0.75–0.77	0.09	N/A	10	U
		Schnabel 2010a (FHS)	209/3,120 (6.7)	0.78	0.75-0.81	0.0003	N/A	10	L

Supplementary Table S4. *Outcomes of studies reporting on included simple (non-augmented) prediction models (continued)*

Prediction Model	Study aim	Study (cohort)	Observed AF/total population (%)	Discrimination		Calibration		Risk prediction interval (years)	Overall RoB
				C-statistic	95%CI	P-value of GOF test	O:E ratio		
Recalibrated five-year model	R	Schnabel 2010b (FHS)	175/8,044 (0.4)	0.78	0.75-0.82	0.92	1.01	5	U
	V	Kumarathurai 2017 (CopHS)	10/646 (1.5)	0.71	0.61-0.82	N/A	N/A	5	H
		Schnabel 2010b (AGES)	226/4,238 (1.3)	0.67	0.64-0.71	0.02	0.97	5	L
		Schnabel 2010b (CHS B)	126/1,552 (1.8)	0.66	0.61-0.71	0.01	1.42	5	H
		Schnabel 2010b (CHS W)	832/8,254 (2.3)	0.68	0.66-0.70	<0.001	2.05	5	H
FHS-CHD[†]	V	Chamberlain 2011 (ARIC)	515/14,546 (3.5)	0.63	0.60-0.66[#]	N/A	N/A	10	U
FHS-hCHD[†]	V	Chamberlain 2011 (ARIC)	515/14,546 (3.5)	0.59	0.56-0.62[#]	N/A	N/A	10	U
FHS-Lubitz	IN	Lubitz 2010 (FHS)	440/11,971 (3.7)	0.842	0.826-0.858	0.11	N/A	8	U
HATCH[†]	V	Li 2019 (NHIS)	12,143/451,199 (2.7)	0.646	0.641-0.651	N/A	N/A	11	U
		Li 2019 (YMID)	921/471,446 (0.2)	0.633	0.598-0.667	N/A	N/A	11	U
		Linker 2018 (ARIC)	165/11,373 (1.5)	0.659	0.615-0.703	N/A	N/A	3	U
		Suenari 2017 (NHIRD)	9,174/670,804 (1.4)	0.716	0.710-0.723	N/A	N/A	9	U
Mayo[**]	V	Linker 2018 (ARIC)	165/11,373 (1.5)	0.721	0.683-0.759	N/A	N/A	3	U
MHS	D	Aronson 2018 (MHS)	5,660/96,778 (5.8)	0.743	0.737-0.749	N/A	0.97	10	U
	V	Aronson 2018 (MHS)	2,791/48,404 (5.8)	0.749	0.741–0.759	N/A	1.00	10	U
PREVEND	D	Rienstra 2016 (PREVEND)	250/8,265 (3.0)	0.830	0.806-0.853	N/A	N/A	10	U
	V	Rienstra 2016 (FHS)	212/3,162 (6.7)	0.704	0.666–0.742	N/A	N/A	10	U
SAAFE[†]	V	Linker 2018 (ARIC)	165/11,373 (1.5)	0.766	0.732-0.800	N/A	N/A	3	U

Supplementary Table S4. *Outcomes of studies reporting on included simple (non-augmented) prediction models (continued)*

Prediction Model	Study aim	Study (cohort)	Observed AF/total population (%)	Discrimination		Calibration		Risk prediction interval (years)	Overall RoB
				C-statistic	95%CI	P-value of GOF test	O:E ratio		
Seirei	D	Hamada 2019 (Seirei)	349/65,984 (0.5)	0.77	0.73-0.81	0.13	0.97	7	H
	V	Hamada 2019* (Seirei)	349/65,984 (0.5)	0.77	0.75-0.79	0.015	N/A	7	H
Suita	D	Kokubo 2017 (Suita)	311/6,898 (4.5)	0.749	0.724–0.774	N/A	N/A	10	U
WHS	D	Everett 2013 (WHS)	404/13,061 (3.1)	N/A	N/A	N/A	N/A	10	H
	V	Everett 2013* (WHS)	212/6,879 (3.1)	0.718	0.684-0.753	0.43	N/A	10	H

AF, atrial fibrillation; AGES, Age, Gene and Environment-Reykjavik Study; ARIC-AF, Atherosclerosis Risk In Communities score for Atrial Fibrillation; ARIC-CHD, Atherosclerosis Risk In Communities score for Coronary Heart Disease; B, blacks; $CHADS_2$, Congestive heart failure, Hypertension, Age >75, Diabetes mellitus, prior Stroke or transient ischemic attack [2 points]; CHA_2DS_2-VASc, Congestive HF, Hypertension, Age >75 [2 points], Stroke/transient ischemic attack/thromboembolism [2 points], Vascular disease, Age 65-74, Sex category; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; CHS, Cardiovascular Health Study; CopHS, Copenhagen Holter Study; D, model derivation; EPIC, European Prospective Investigation into Cancer and Nutrition; FHS-AF, Framingham Heart Study score for Atrial Fibrillation; FHS-CHD, Framingham Heart Study score for Coronary Heart Disease; FHS-hCHD, Framingham Heart Study score for hard Coronary Heart Disease; GOF, goodness-of-fit; H, high; HATCH, Hypertension, Age, stroke or Transient ischemic attack, Chronic obstructive pulmonary disease, Heart failure; IN, incidental use of a risk model, with variables and coefficients similar to but differing from a risk model previously derived or validated for incident AF; L, low; MHS, Maccabi Healthcare Services; MPP-RES, Malmö Preventive Project Re-examination Study; N/A, not available; NHIRD, National Health Insurance Research Database; NHIS, National Health Insurance Service; O:E, observed versus expected events; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; PREVEND, Prevention of Renal and Vascular End-stage Disease; R, recalibration; RoB, risk of bias; RS, Rotterdam Study; SAAFE, Screening for Asymptomatic Atrial Fibrillation Events; U, unclear; ULSAM, Uppsala Longitudinal Study of Adult Men; V, model validation; W, whites; WHS, Women's Health Study; YMID, Yunnan Medical Insurance Database; 95%CI, 95% confidence interval.

§ Data are based on the ARIC-AF Cox regression model; * validation using bootstrap resampling; # 95%CI not reported in article, estimated from the reported C-statistic according to methods described by Debray *et al.* 2017; $ Alonso 2013 provided only a calibration plot for the combined derivation cohorts, and none for the validation cohorts; † Risk model not originally developed for incident AF risk prediction, thus no derivation cohort data available.; ** The Mayo model was originally intended for incident AF risk prediction, however was derived from a secondary care cohort, preventing us from including these data into this review.

Supplementary Table S5. *Augmentation of included models*

Study (cohort)	Model	Discrimination		Calibration		NRI (95%CI)	Risk prediction interval (years)
		C-statistic	**95%CI**	**P-value of GOF test**	**O:E ratio**		
				CHARGE-AF			
Alonso 2013 (ARIC + CHS + FHS)	Simple model	0.765	0.748-0.781	0.41	N/A	Reference	5
	SM + ECG parameters	0.767	0.750-0.783	0.80	N/A	-0.0032 (-0.0178-0.0113)	
Alonso 2013 (AGES)	Simple model	0.664	0.632-0.697	0.18	N/A	Reference	5
	SM + ECG parameters	0.665	0.633-0.697	0.053	N/A	N/A	
Alonso 2013 (RS)	Simple model	0.705	0.663-0.747	0.06	N/A	Reference	5
	SM + ECG parameters	0.716	0.680-0.761	0.34	N/A	N/A	
Alonso 2016 (MESA)	Simple model	0.779	0.744,-0.814	0.002	N/A	Reference	5
	SM + hsCRP	0.784	0.747- 0.821	0.08	N/A	-0.037 (-0.102-0.020)	
	SM + NT-proBNP	0.825	0.791- 0.859	0.36	N/A	0.045 (-0.049-0.147)	
	SM + hsCRP + NT-proBNP	0.825	0.791- 0.860	0.55	N/A	0.041 (-0.055, 0.147)	
Berntsson 2017 (MPP-RES)	Simple model	0.683	0.65-0.72#	0.051	1.26	Reference	5
	SM + MR-proANP	0.738	0.70-0.77#	0.734	1.00	0.267	
Chaker 2015 (RS)	Simple model	0.722	0.69-0.75#	N/A	N/A	Reference	10
	SM + FT4	0.729	0.70-0.76#	N/A	N/A	N/A	
Linker 2018	Simple model	0.775	0.741-0.810	N/A	N/A	Reference	3
	SM + ECG parameters	0.785	0.750-0.819	N/A	N/A	N/A	
Maheshwari 2017 (ARIC)	Simple model	0.719	0.702-0.736	0.05	N/A	Reference	10
	SM + P-wave axis	0.722	0.705-0.739	0.08	N/A	0.021 (0.001-0.040)*	

Supplementary Table S5. *Augmentation of included models (continued)*

Study (cohort)	Model	Discrimination		Calibration		NRI (95%CI)	Risk prediction interval (years)
		C-statistic	**95%CI**	**P-value of GOF test**	**O:E ratio**		
Sinner 2014 (ARIC+CHS+FHS)	Simple model	0.765	0.748-0.781	N/A	N/A	Reference	5
	SM + BNP	0.790	0.77-0.81#	0.05	N/A	0.389 (0.322-0.455)*	
	SM + CRP	0.768	0.75-0.79#	0.31	N/A	0.154 (0.081-0.228)*	
	SM + BNP + CRP	0.791	0.77-0.81#	0.16	N/A	0.375 (0.303-0.448)*	
Sinner 2014 (AGES)	Simple model	0.664	0.632-0.697	0.18	N/A	Reference	5
	SM + BNP	0.722	0.69-0.75#	0.76	N/A	0.612 (0.497-0.734)*	
	SM + CRP	0.669	0.64-0.70#	0.52	N/A	0.142 (0.018-0.269)*	
	SM + BNP + CRP	0.723	0.69-0.75#	0.48	N/A	0.633 (0.517-0.751)*	
Sinner 2014 (RS)	Simple model	0.705	0.663-0.747	0.06	N/A	Reference	5
	SM + BNP	0.746	0.70-0.79#	0.06	N/A	0.449 (0.248-0.623)*	
	SM + CRP	0.700	0.65-0.75#	0.08	N/A	0.011 (−0.178-0.184)	
	SM + BNP + CRP	0.744	0.70-0.79#	<0.01	N/A	0.470 (0.270-0.655)*	
Svennberg 2016 (ULSAM)	Simple model	0.62	0.56-0.68	N/A	N/A	Reference	10
	SM + NT-proBNP	0.66	0.59-0.72	N/A	N/A	0.01	
Svennberg 2016 (PIVUS)	Simple model	0.61	0.55-0.67	N/A	N/A	Reference	10
	SM + NT-proBNP	0.64	0.59-0.70	N/A	N/A	0.24	
CHARGE-Magnani							

Supplementary Table S5. *Augmentation of included models (continued)*

Study (cohort)	Model	Discrimination		Calibration		NRI (95%CI)	Risk prediction interval (years)
		C-statistic	**95%CI**	**P-value of GOF test**	**O:E ratio**		
Magnani 2015 (ARIC)	Simple model	0.71	0.69-0.73	N/A	N/A	Reference	10
	SM + PR interval >200ms	0.71	0.69-0.74	N/A	N/A	0.3	
	SM + P duration >120ms	0.72	0.69-0.74	N/A	N/A	-0.1	
	SM + P area ≥95th percentile	0.71	0.69-0.74	N/A	N/A	1.5	
	SM + P terminal force >4000 μV*ms	0.72	0.70-0.74	N/A	N/A	2.0	
Magnani 2015 (FHS)	Simple model	0.78	0.75-0.80	N/A	N/A	Reference	10
	SM + PR interval >200ms	0.78	0.75-0.80	N/A	N/A	-0.2	
	SM + P duration >120ms	0.78	0.75-0.80	N/A	N/A	2.9	
	SM + P area ≥95th percentile	0.78	0.75-0.80	N/A	N/A	-0.1	
	SM + P terminal force >4000 μV*ms	0.78	0.75-0.81	N/A	N/A	0.04	
		FHS-AF ten-year model					
Kumarathurai 2017 (CopHS)	Simple model	0.656	0.609-0.788	N/A	0.90	Reference	10
	SM + PAC	0.726	0.667-0.847	N/A	1.00	N/A	
	SM + NT-proBNP	0.684	0.619-0.814	N/A	0.85	N/A	
	SM + PAC + NT-proBNP	0.723	0.664-0.842	N/A	0.80	N/A	
Rosenberg 2012 (CHS)	Simple model	0.649	0.63-0.67#	N/A	N/A	Reference	10
	SM + height	0.659	0.64-0.68#	N/A	N/A	N/A	
Schnabel 2010a (FHS)	Simple model	0.78	0.75-0.81	0.99	N/A	Reference	10
	SM + BNP	0.80	0.78-0.83	0.44	N/A	0.06 (-0.01-0.14)	
	SM + CRP	0.78	0.75-0.81	0.29	N/A	0.009 (-0.04-0.06)	
	SM + BNP + CRP	0.81	0.78-0.84	0.65	N/A	0.11 (0.04-0.19)*	

Supplementary Table S5. *Augmentation of included models (continued)*

Study (cohort)	Model	Discrimination		Calibration		NRI (95%CI)	Risk prediction interval (years)
		C-statistic	**95%CI**	**P-value of GOF test**	**O:E ratio**		
				FHS-Lubitz			
Lubitz 2010 (FHS)	Simple Model	0.842	0.826-0.860	0.11	N/A	Reference	8
	SM + No. of first degree relatives with AF	0.844	0.828-0.860	N/A	N/A	−0.022 (−0.046-0.003)	
	SM + familial AF	0.844	0.828-0.860	0.08	N/A	−0.029 (−0.057-0.000)	
	SM + familial AF + age at onset of youngest affected relative	0.846	0.830-0.862	N/A	N/A	−0.010 (−0.041-0.021)	
	SM + premature familial AF	0.846	0.831-0.862	N/A	N/A	0.011 (−0.021-0.042)	
				Seirei			
Hamada 2019 (Seirei)	Simple model	0.77	0.73-0.81	0.13	0.97	Reference	7
	SM + ECG parameters	0.78	0.74-0.82	0.37	0.97	0.108 (0.006-0.153)*	

AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; BNP, brain natriuretic peptide; CHS, cardiovascular Health Study; CopHS, Copenhagen Holter Study; CRP, C-reactive protein; ECG, electrocardiogram; FHS, Framingham Heart Study; FHS-AF, Framingham Heart Study score for Atrial Fibrillation; FT4, free T4; GDF-15, Growth/differentiation factor 15; GOF, goodness-of-fit; hsCRP, high-sensitive C-reactive protein; ECG, electrocardiogram; hsTnI, high-sensitive troponin I; MESA, Multi-Ethnic Study of Atherosclerosis; MPP-RES, Malmö Preventive Project Re-examination Study; MR-proANP, mid-regional pro-atrial natriuretic peptide; N/A, not available; NRI, net reclassification improvement index; NRI, net reclassification improvement index; NT-proBNP, N-terminal pro-B-type natriuretic peptide; O:E, observed versus expected; PAC, premature atrial contraction; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; RS, Rotterdam Study; SM, simple model; ST2, suppression of tumorigenicity 2; ULSAM, Uppsala Longitudinal Study of Adult Men; 95%CI, 95% confidence interval.

95%CI not reported in article, estimated from the reported C-statistic according to methods described by Debray *et al.* 2017; * significant net reclassification improvement

Supplementary Table S6. *Risk of bias and applicability assessment*

Study	Cohort	Source of data	Partici-pants	Outcome	Prediction variables	Missing data	Model develop-ment	Model evaluation	Model perfor-mance	Results	Interpre-tation and discussion	Overall RoB
Alonso 2013	AGES	L	L	L	L	U	N/A	L*	L	L	L	U
	ARIC	L	L	L	L	L	L	L*	L	L	L	L
	CHS	L	H	L	L	L	L	L*	L	L	L	H
	FHS	L	L	L	L	L	L	L*	L	L	L	L
	RS	L	L	L	L	U	N/A	L*	L	L	L	U
Alonso 2016	MESA	L	H	U	L	U	N/A	L	L	L	L	H
Aronson 2018	MHS	U	L	U	L	L	L	U	L	L	L	U
Berntsson 2017	MPP-RES	L	L	U	L	U	N/A	L	L	L	L	U
Chaker 2015	RS	L	L	L	L	U	N/A	U	L	L	L	U
Chamberlain 2011	ARIC	L	L	L	L	U	L	N/A	L	L	L	U
Christophersen 2016	FHS	L	L	L	L	U	N/A	L	L	L	L	U
Everett 2013	WHS	L	H	U	L	U	L	L	L	L	L	H
Hamada 2019	Seirei	U	H	U	L	L	L	L	U	L	L	H
Kokubo 2017	Suita	L	U	L	L	U	L	L#	U	L	L	U
Kumarathurai 2017	CopHS	L	H	U	L	U	N/A	U	L	L	L	H
Li 2019	NHIS	U	L	U	L	U	N/A	L	L	L	L	U
	YMID	U	U	U	L	U	L	U	L	L	L	U
Linker 2018	ARIC	L	L	L	L	L	N/A	U	L	L	L	U
Lubitz 2010	FHS	L	L	L	L	U	U	U	L	L	L	U
Magnani 2015	ARIC	L	H	L	L	U	L	U	L	L	L	H
	FHS	L	H	L	L	U	L	U	L	L	L	H
Maheshwari 2017	ARIC	L	L	L	L	U	N/A	U	L	L	L	U
Pfister 2015	EPIC	L	U	U	L	U	N/A	L	L	L	L	U
Rienstra 2014	FHS	L	L	L	L	U	U	U	L	L	L	U
Rienstra 2016	FHS	L	L	L	L	U	N/A	L	L	L	L	U
	PREVEND	L	L	L	L	U	L	N/A	L	L	L	U
Rosenberg 2012	CHS	L	H	L	L	U	N/A	U	L	L	L	H
Saliba 2016	ClaHS	U	L	U	U	U	N/A	L	L	L	L	U
Schnabel 2009	FHS	L	U	L	L	U	L	N/A	L	L	L	U
Schnabel 2010a	FHS	L	L	L	L	L	N/A	L	L	L	L	L

Supplementary Table S6. *Risk of bias and applicability assessment* (*continued*)

Study	Cohort	Source of data	Partici-pants	Outcome	Prediction variables	Missing data	Model develop-ment	Model evaluation	Model perfor-mance	Results	Interpre-tation and discussion	Overall RoB
Schnabel 2010b	AGES	L	L	L	L	L	N/A	L	L	L	L	L
	CHS	L	H	L	L	L	N/A	L	L	L	L	H
	FHS	L	U	L	L	L	N/A	L	L	L	L	U
Sinner 2014	AGES	L	U	L	L	L	N/A	L	L	L	L	U
	ARIC	L	U	L	L	L	N/A	L	L	L	L	U
	CHS	L	H	L	L	L	N/A	L	L	L	L	H
	FHS	L	U	L	L	L	N/A	L	L	L	L	U
	RS	L	U	L	L	L	N/A	L	L	L	L	U
Suenari 2017	NHIRD	U	L	U	L	U	N/A	L	L	L	L	U
Svennberg 2016	PIVUS	L	L	U	L	L	N/A	U	U	L	L	U
	ULSAM	L	H	U	L	L	N/A	U	U	L	L	H

AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; CHS, Cardiovascular Health Study; ClaHS, Clalit Health Service; CopHS, Copenhagen Holter Study; EPIC, European Prospective Investigation into Cancer and Nutrition; FHS, Framingham Heart Study; H, high; L, low; MESA, Multi-Ethnic Study of Atherosclerosis; MHS, Maccabi Healthcare Services; MPP-RES, Malmö Preventive Project Re-examination Study; NHIRD, National Health Insurance Research Database; NHIS, National Health Insurance Service; N/A, not applicable; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; PREVEND, Prevention of Renal and Vascular End-stage Disease; RoB, Risk of Bias; RS, Rotterdam Study; U, unclear; ULSAM, Uppsala Longitudinal Study of Adult Men; YMID, Yunnan Medical Insurance Database.

* Alonso 2013 validated the FHS-AF risk score for ten-year AF risk on their cohorts, however follow-up duration within these cohorts was five years. The Model evaluation domain of all five cohorts in Alonso 2013 on FHS-AF therefore had an unclear risk of bias. The Model evaluation domain had low risk of bias on comparisons of other models in Alonso 2013; # Kokubo 2017 used ten-year follow-up data to validate CHARGE-AF, which is for intended five-year risk prediction. In validation of CHARGE-AF the Model evaluation domain therefore had an unclear risk of bias, while validation of other cohorts had low risk of bias.

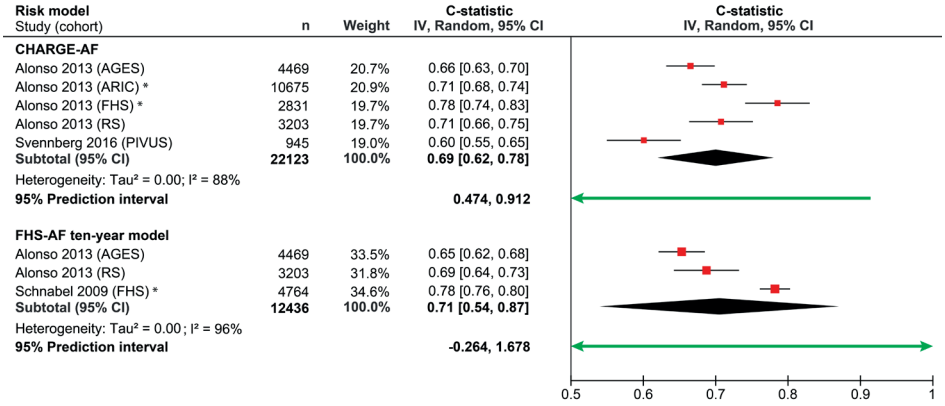
Supplementary Table S7. *Sensitivity analysis replacing double cohorts from the primary and secondary meta-analyses by later published data of the same cohort*

	Comparison	Summary C-statistic	95%CI	I^2 (%)	Tau^2	95%PI	n patients
Primary analyses	**CHARGE-AF (n = 8 cohorts)**						
	Primary meta-analysis	0.71	0.66-0.76	87	0.00	0.554-0.865	58,137
	FHS data by Christophersen 2016, not Alonso 2013	0.71	0.66-0.76	90	0.00	0.562-0.852	65,028
	ARIC data by Linker 2018, not Alonso 2013	0.72	0.66-0.78	90	0.00	0.550-0.886	58,835
	ARIC data by Maheshwari 2017, not Alonso 2013	0.71	0.66-0.76	87	0.00	0.556-0.866	58,418
	RS data by Chaker 2015, not Alonso 2013	0.71	0.66-0.76	87	0.00	0.556-0.866	63,674
	FHS-AF ten-year model (n = 5 cohorts)						
	Primary analysis	0.70	0.64-0.76	95	0.00	0.535-0.869	33,846
	FHS data by Alonso 2013, not Schnabel 2009	0.70	0.64-0.76	85	0.00	0.545-0.855	31,913
	FHS data by Schnabel 2010a, not Schnabel 2009	0.70	0.64-0.76	90	0.00	0.539-0.863	32,202
	ARIC data by Alonso 2013, not Chamberlain 2011	0.70	0.64-0.76	94	0.00	0.530-0.871	29,975
	ARIC data by Linker 2018, not Chamberlain 2011	0.69	0.62-0.77	94	0.00	0.491-0.893	30,673
Secondary analyses	**CHARGE-AF, five-year prediction window (n = 6 cohorts)**						
	Secondary meta-analysis	0.72	0.66-0.78	85	0.00	0.567-0.881	50,328
	FHS data by Christophersen 2016, not Alonso 2013	0.72	0.67-0.77	89	0.00	0.579-0.861	57,219
	CHARGE-AF, ten-year prediction window (n = 4 cohorts)						
	No double cohorts	-	-	-	-	-	-
	FHS-AF ten-year model, five-year prediction window (n = 4 cohorts)						
	No double cohorts	-	-	-	-	-	-
	FHS-AF ten-year model, ten-year prediction window (n = 3 cohorts)						
	Secondary meta-analysis	0.72	0.59-0.85	95	0.00	-0.026-1.471	26,174
	FHS data by Schnabel 2010a, not Schnabel 2009	0.72	0.66-0.78	91	0.00	-0.011-1.455	24,530

ARIC, Atherosclerosis Risk In Communities; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; FHS, Framingham Heart Study; RS, Rotterdam Study; 95%CI, 95% confidence interval; 95%PI, 95% prediction interval.

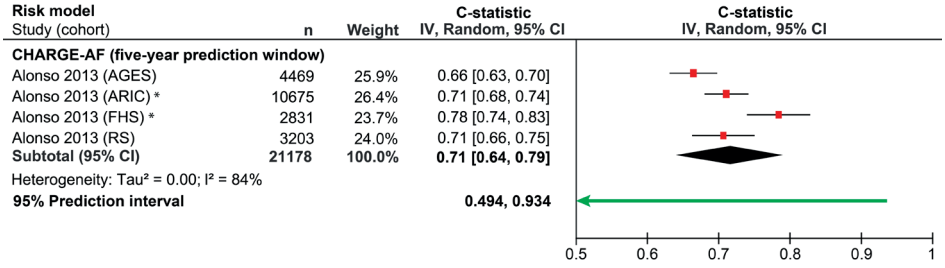
SUPPLEMENTARY FIGURES

Supplementary Figure S1. *Sensitivity analysis restricting the primary analysis to models with ≥3 eligible cohorts that had demonstrated sufficient calibration*



AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; CI, confidence interval; FHS, Framingham Heart Study; FHS-AF, Framingham Heart Study score for Atrial Fibrillation; IV, inverse variance; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; RS, Rotterdam Study; 95% CI, 95% confidence interval.
* Derivation cohort.

Supplementary Figure S2. *Sensitivity analysis restricting the secondary analysis to models with ≥3 eligible cohorts that had demonstrated sufficient calibration*



AGES, Age, Gene and Environment-Reykjavik Study; ARIC, Atherosclerosis Risk In Communities; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology; CI, confidence interval; FHS, Framingham Heart Study; IV, inverse variance; RS, Rotterdam Study; 95% CI, 95% confidence interval.
* Derivation cohort.

REFERENCES

1 Chugh SS, Roth GA, Gillum RF, Mensah GA. Global burden of atrial fibrillation in developed and developing nations. Glob Heart 2014;9:113–9.

2 Magnussen C, Niiranen TJ, Ojeda FM, Gianfagna F, Blankenberg S, Njolstad I et al. Sex differences and similarities in atrial fibrillation epidemiology, risk factors, and mortality in community cohorts: results from the BiomarCaRE Consortium (Biomarker for Cardiovascular Risk Assessment in Europe). Circulation 2017;136:1588–97.

3 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.

4 Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991;22:983–8.

5 Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Europace 2016;18:1609–78.

6 Freedman B, Camm J, Calkins H, Healey JS, Rosenqvist M, Wang JG et al. Screening for atrial fibrillation: a report of the AF-SCREEN International Collaboration. Circulation 2017;135:1851–67.

7 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.

8 Jacobs MS, Kaasenbrood F, Postma MJ, van Hulst M, Tieleman RG. Cost-effectiveness of screening for atrial fibrillation in primary care with a handheld, single-lead electrocardiogram device in the Netherlands. Europace 2018;20:12–8.

9 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.

10 Lowres N, Neubeck L, Redfern J, Freedman SB. Screening to identify unknown atrial fibrillation. A systematic review. Thromb Haemost 2013;110:213–22.

11 Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart 2012;98:683–90.

12 Mairesse GH, Moran P, Van Gelder IC, Elsner C, Rosenqvist M, Mant J et al. Screening for atrial fibrillation: a European Heart Rhythm Association (EHRA) consensus document endorsed by the Heart Rhythm Society (HRS), Asia Pacific Heart Rhythm Society (APHRS), and Sociedad Latinoamericana de Estimulacion Cardiaca y Electrofisiologia (SOLAECE). Europace 2017;19:1589–623.

13 Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009;62:1006–12.

14 Debray TP, Damen JA, Snell KI, Ensor J, Hooft L, Reitsma JB et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ 2017;356:i6460.

15 Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11:e1001744.

16 Partlett C, Riley RD. Random effects meta-analysis: coverage performance of 95% confidence and prediction intervals following REML estimation. Stat Med 2017;36:301–17.
17 Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011;342:d549.
18 Higgins JPT, Green S. (eds). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. www.handbook.cochrane.org.
19 Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc 2009;172:137–59.
20 Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB et al. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc 2013;2:e000102.
21 Chamberlain AM, Agarwal SK, Folsom AR, Soliman EZ, Chambless LE, Crow R et al. A clinical risk score for atrial fibrillation in a biracial prospective cohort (from the Atherosclerosis Risk in Communities [ARIC] study). Am J Cardiol 2011;107:85–91.
22 Li YG, Pastori D, Farcomeni A, Yang PS, Jang E, Joung B et al. A Simple Clinical Risk Score (C2HEST) for predicting incident atrial fibrillation in Asian subjects: derivation in 471,446 Chinese subjects, with internal validation and external application in 451,199 Korean subjects. Chest 2019;155:510–8.
23 Rienstra M, Geelhoed B, Yin XY, Siland JE, Vermond RA, Mulder BA et al. Cluster individuals based on phenotype and determine the risk for atrial fibrillation in the PREVEND and Framingham Heart Study Populations. PLoS One 2016;11:e0165828.
24 Schnabel RB, Sullivan LM, Levy D, Pencina MJ, Massaro JM, D'Agostino RBSr et al. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. Lancet 2009;373:739–45.
25 Aronson D, Shalev V, Katz R, Chodick G, Mutlak D. Risk score for prediction of 10-year atrial fibrillation: a community-based study. Thromb Haemost 2018;118:1556–63.
26 Everett BM, Cook NR, Conen D, Chasman DI, Ridker PM, Albert CM. Novel genetic markers improve measures of atrial fibrillation risk prediction. Eur Heart J 2013;34:2243–51.
27 Hamada R, Muto S. Simple risk model and score for predicting of incident atrial fibrillation in Japanese. J Cardiol 2019;73:65–72.
28 Kokubo Y, Watanabe M, Higashiyama A, Nakao YM, Kusano K, Miyamoto Y. Development of a basic risk score for incident atrial fibrillation in a Japanese general population—the Suita study. Circ J 2017;81:1580–8.
29 Alonso A, Roetker NS, Soliman EZ, Chen LY, Greenland P, Heckbert SR. Prediction of atrial fibrillation in a racially diverse cohort: the Multi-Ethnic Study of Atherosclerosis (MESA). J Am Heart Assoc 2016;5;1–8.
30 Berntsson J, Smith JG, Nilsson PM, Hedblad B, Melander O, Engstrom G. Pro-atrial natriuretic peptide and prediction of atrial fibrillation and stroke: the Malmo Preventive Project. Eur J Prev Cardiol 2017;24:788–95.
31 Chaker L, Heeringa J, Dehghan A, Medici M, Visser WE, Baumgartner C et al. Normal thyroid function and the risk of atrial fibrillation: the Rotterdam study. J Clin Endocrinol Metab 2015;100:3718–24.
32 Christophersen IE, Yin XY, Larson MG, Lubitz SA, Magnani JW, McManus DD et al. A comparison of the CHARGE-AF and the CHA(2)DS(2)-VASc risk scores for prediction of atrial fibrillation in the Framingham Heart Study. Am Heart J 2016;178:45–54.

33 Kumarathurai P, Mouridsen MR, Mattsson N, Larsen BS, Nielsen O, Gerds TA et al. Atrial ectopy and N-terminal pro-B-type natriuretic peptide as predictors of atrial fibrillation: a population-based cohort study. Europace 2017;19:364–70.

34 Linker DT, Murphy TB, Mokdad AH. Selective screening for atrial fibrillation using multivariable risk models. Heart 2018;104:1492–9.

35 Maheshwari A, Norby FL, Soliman EZ, Koene R, Rooney M, O'Neal WT et al. Refining prediction of atrial fibrillation risk in the general population with analysis of P-wave axis (from the Atherosclerosis Risk in Communities Study). Am J Cardiol 2017;120:1980–4.

36 Pfister R, Bragelmann J, Michels G, Wareham NJ, Luben R, Khaw KT. Performance of the CHARGE-AF risk model for incident atrial fibrillation in the EPIC Norfolk cohort. Eur J Prev Cardiol 2015;22:932–9.

37 Rienstra M, Yin XY, Larson MG, Fontes JD, Magnani JW, McManus DD et al. Relation between soluble ST2, growth differentiation factor-15, and high-sensitivity troponin I and incident atrial fibrillation. Am Heart J 2014;167:109–15.e2.

38 Saliba W, Gronich N, Barnett-Griness O, Rennert G. Usefulness of CHADS2 and CHA2DS2-VASc scores in the prediction of new-onset atrial fibrillation: a population-based study. Am J Med 2016;129:843–9.

39 Schnabel RB, Larson MG, Yamamoto JF, Sullivan LM, Pencina MJ, Meigs JB et al. Relations of biomarkers of distinct pathophysiological pathways and atrial fibrillation incidence in the community. Circulation 2010;121:200–7.

40 Schnabel RB, Aspelund T, Li G, Sullivan LM, Suchy-Dicey A, Harris TB et al. Validation of an atrial fibrillation risk algorithm in whites and African Americans. Arch Intern Med 2010;170:1909–17.

41 Sinner MF, Stepas KA, Moser CB, Krijthe BP, Aspelund T, Sotoodehnia N et al. B-type natriuretic peptide and C-reactive protein in the prediction of atrial fibrillation risk: the CHARGE-AF Consortium of community-based cohort studies. Europace 2014;16:1426–33.

42 Suenari K, Chao TF, Liu CJ, Kihara Y, Chen TJ, Chen SA. Usefulness of HATCH score in the prediction of new-onset atrial fibrillation for Asians. Medicine (Baltimore) 2017;96:e5597.

43 Svennberg E, Lindahl B, Berglund L, Eggers KM, Venge P, Zethelius B et al. NT-proBNP is a powerful predictor for incident atrial fibrillation—validation of a multimarker approach. Int J Cardiol 2016;223:74–81.

44 Lubitz SA, Yin XY, Fontes JD, Magnani JW, Rienstra M, Pai M et al. Association between familial atrial fibrillation and risk of new-onset atrial fibrillation. JAMA 2010;304:2263–9.

45 Rosenberg MA, Patton KK, Sotoodehnia N, Karas MG, Kizer JR, Zimetbaum PJ et al. The impact of height on the risk of atrial fibrillation: the Cardiovascular Health Study. Eur Heart J 2012;33:2709–17.

46 Magnani JW, Zhu L, Lopez F, Pencina MJ, Agarwal SK, Soliman EZ et al. P-wave indices and atrial fibrillation: cross-cohort assessments from the Framingham Heart Study (FHS) and Atherosclerosis Risk in Communities (ARIC) study. Am Heart J 2015;169:53–61.e1.

47 Brunner KJ, Bunch TJ, Mullin CM, May HT, Bair TL, Elliot DW et al. Clinical predictors of risk for atrial fibrillation: implications for diagnosis and monitoring. Mayo Clin Proc 2014;89:1498–505.

48 Chambless LE, Folsom AR, Sharrett AR, Sorlie P, Couper D, Szklo M et al. Coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC) study. J Clin Epidemiol 2003;56:880–90.

49 Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. JAMA 2001;285:2864–70.

50 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.

51 Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837–47.

52 D'Agostino RBSr, Grundy S, Sullivan LM, Wilson P; CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001;286:180–7.

53 de Vos CB, Pisters R, Nieuwlaat R, Prins MH, Tieleman RG, Coelen R-JS et al. Progression from paroxysmal to persistent atrial fibrillation clinical correlates and prognosis. J Am Coll Cardiol 2010;55:725–31.

54 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.

55 Jones C, Pollit V, Fitzmaurice D, Cowan C, Grp GD. The management of atrial fibrillation: summary of updated NICE guidance. BMJ 2014;348.

56 Proietti M, Mairesse GH, Goethals P, Scavee C, Vijgen J, Blankoff I et al. ; Belgian Heart Rhythm Week Investigators. A population screening programme for atrial fibrillation: a report from the Belgian Heart Rhythm Week screening programme. Europace 2016;18:1779–86.

57 Singh JP, Thomas SS. Atrial fibrillation and heart failure prevention: do we need a risk score? JACC Heart Fail 2017;5:53–5.

58 Jones NR, Taylor KS, Taylor CJ, Aveyard P. Weight change and the risk of incident atrial fibrillation: a systematic review and meta-analysis. Heart 2019;105:1799–805.

59 Himmelreich JCL, Lucassen WAM, Heugen M, Bossuyt PMM, Tan HL, Harskamp RE et al. Frequent premature atrial contractions are associated with atrial fibrillation, brain ischaemia, and mortality: a systematic review and meta-analysis. Europace 2019;21:698–707.

60 He WQ, Chu YJ. Atrial fibrillation as a prognostic indicator of myocardial infarction and cardiovascular death: a systematic review and meta-analysis. Sci Rep 2017;7:1–13.

61 Ruddox V, Sandven I, Munkhaugen J, Skattebu J, Edvardsen T, Otterstad JE. Atrial fibrillation and the risk for myocardial infarction, all-cause mortality and heart failure: a systematic review and meta-analysis. Eur J Prev Cardiol 2017;24:1555–66.

62 Proietti M, Farcomeni A, Romiti GF, Di Rocco A, Placentino F, Diemberger I et al. Association between clinical risk scores and mortality in atrial fibrillation: systematic review and network meta-regression of 669,000 patients. Eur J Prev Cardiol 2018;doi:10.1177/2047487318817662.

63 Caldeira D, Costa J, Fernandes RM, Pinto FJ, Ferreira JJ. Performance of the HAS-BLED high bleeding-risk category, compared to ATRIA and HEMORR2HAGES in patients with atrial fibrillation: a systematic review and meta-analysis. J Interv Card Electrophysiol 2014;40:277–84.

64 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.

65 Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care 2012;28:138–44.

3

CHARGE-AF in a national routine primary care electronic health records database in the Netherlands: validation for 5-year risk of atrial fibrillation and implications for patient selection in atrial fibrillation screening

Jelle C.L. Himmelreich, Wim A.M. Lucassen, Ralf E. Harskamp, Claire Aussems, Henk C.P.M. van Weert, Mark M.J. Nielen

Open Heart, Volume 8, Issue 1,

October 2021, Pages 1-11,

*https://doi.org/*10.1136/openhrt-2020-001459

ABSTRACT

Aims: To validate a multivariable risk prediction model (Cohorts for Heart and Aging Research in Genomic Epidemiology model for atrial fibrillation (CHARGE-AF)) for 5-year risk of atrial fibrillation (AF) in routinely collected primary care data and to assess CHARGE-AF's potential for automated, low-cost selection of patients at high risk for AF based on routine primary care data.

Methods: We included patients aged ≥40 years, free of AF and with complete CHARGE-AF variables at baseline, 1 January 2014, in a representative, nationwide routine primary care database in the Netherlands (Nivel-PCD). We validated CHARGE-AF for 5-year observed AF incidence using the C-statistic for discrimination, and calibration plot and stratified Kaplan-Meier plot for calibration. We compared CHARGE-AF with other predictors and assessed implications of using different CHARGE-AF cut-offs to select high-risk patients.

Results: Among 111 475 patients free of AF and with complete CHARGE-AF variables at baseline (17.2% of all patients aged ≥40 years and free of AF), mean age was 65.5 years, and 53% were female. Complete CHARGE-AF cases were older and had higher AF incidence and cardiovascular comorbidity rate than incomplete cases. There were 5264 (4.7%) new AF cases during 5-year follow-up among complete cases. CHARGE-AF's C-statistic for new AF was 0.74 (95% CI 0.73 to 0.74). The calibration plot showed slight risk underestimation in low-risk deciles and overestimation of absolute AF risk in those with highest predicted risk. The Kaplan-Meier plot with categories <2.5%, 2.5%–5% and >5% predicted 5-year risk was highly accurate. CHARGE-AF outperformed CHA_2DS_2-VASc (Cardiac failure or dysfunction, Hypertension, Age >=75 [Doubled], Diabetes, Stroke [Doubled]-Vascular disease, Age 65-74, and Sex category [Female]) and age alone as predictors for AF. Dichotomisation at cut-offs of 2.5%, 5% and 10% baseline CHARGE-AF risk all showed merits for patient selection in AF screening efforts.

Conclusion: In patients with complete baseline CHARGE-AF data through routine Dutch primary care, CHARGE-AF accurately assessed AF risk among older primary care patients, outperformed both CHA_2DS_2-VASc and age alone as predictors for AF and showed potential for automated, low-cost patient selection in AF screening.

KEY QUESTIONS

What is already known about this subject?

Patient selection in atrial fibrillation (AF) screening studies has so far been based mainly on high age. There are indications, however, that multivariable risk prediction models are better at discriminating for high and low risk of AF in the community than age alone. A recent systematic review and meta-analysis showed that Cohorts for Heart and Aging Research in Genomic Epidemiology model for atrial fibrillation (CHARGE-AF) may be the best suitable risk model for this purpose in community cohorts.

What does this study add?

Previous validations of CHARGE-AF have been performed mainly in prospective community cohorts with high completeness of data. If the model were to be used for low-cost, automated patient selection in AF screening, however, it is more likely that researchers will turn to readily available routine primary care data, without a costly baseline visit for each eligible patient. This study is the first to provide detailed information on how selecting at different cut-offs of CHARGE-AF risk would translate into numbers of patients to be screened and percentage of AF yield to be expected while using a large European routine primary care dataset.

How might this impact on clinical practice?

Outcomes of this work are relevant to the prospect of using clinical risk models as triage test for AF screening, while also maintaining low cost in their risk assessment efforts. We showed that those with complete CHARGE-AF variables as per routine primary care constitute a small but highly relevant subset for AF screening. CHARGE-AF's high accuracy in predicting absolute 5-year year risk for predefined risk categories suggests that the model can be used to reliably differentiate between low and high AF risk among cases with complete CHARGE-AF data through routine primary care. Moreover, CHARGE-AF can do so with higher accuracy than two predictors that are currently used as triage tests for AF screening: age alone and the congestive heart failure, hypertension, age, diabetes and previous stroke or transient ischaemic attack, vascular disease and female sex category (CHA_2DS_2-VASc) score. This work therefore encourages researchers in the field of community AF screening to consider CHARGE-AF as a triage test for patient selection.

INTRODUCTION

Atrial fibrillation (AF) is a common arrhythmia increasing in incidence with age.[1] It is associated with a higher risk of ischaemic stroke for which effective prophylactic treatment is available.[2] There is increasing interest in more efficient strategies for early AF detection in the ageing community.[3] One approach is the use of multivariable risk models for patient selection in AF screening: longer or more frequent follow-up in patients with higher risk and less stringent regimes in the lower risk strata.[4]

The Cohorts for Heart and Aging Research in Genomic Epidemiology model for atrial fibrillation (CHARGE-AF) model predicts an individual's 5-year risk of new AF using relatively easily obtainable variables: age, ethnicity, height, weight, systolic blood pressure (SBP), diastolic blood pressure (DBP), current smoking, antihypertensive medication use, diabetes mellitus (DM), heart failure and myocardial infarction (MI).[5] CHARGE-AF was derived and calibrated in community-dwelling older subjects of European and African descent. It has been validated in various community cohorts[5-10] and appears to be the most viable prediction model for patient selection in future community AF screening.[11]

To further increase efficiency of risk model-assisted AF screening efforts, minimal resources should be required to adequately perform baseline risk stratification.[3] One eligible data source for this purpose are primary care electronic health records (EHRs). However, while age and cardiovascular morbidities can be deduced from primary care EHRs with high completeness, other CHARGE-AF variables may not be as frequently recorded. Most notably, the body measurements required in CHARGE-AF—height, weight, SBP and DBP—have been shown to often be incomplete in real-world primary care data, with selective reporting favouring those with higher comorbidity rates.[12,13]

If CHARGE-AF were shown to be a valid risk stratification tool within the subset of patients with readily available complete data for CHARGE-AF risk assessment, and if this subset were to constitute a population with clinical significance for AF screening, this could point to a reduced necessity for a baseline visit prior to risk stratification in these patients. We therefore set out to perform a retrospective cohort study using a nationwide primary care EHR database with three aims:

1. To study the subgroup of primary care patients with recent and complete baseline data for the CHARGE-AF variables in terms of relevance for AF screening.
2. To validate CHARGE-AF for 5-year AF risk and to compare it with other established predictors for AF in complete CHARGE-AF cases.

3. To explore how a choice of baseline CHARGE-AF risk cut-offs could affect patient selection and potential AF yield in future AF screening among complete CHARGE-AF cases.

METHODS

We reported this study in accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis statement.[14]

Netherlands Institute for Health Services Research Primary Care Database (Nivel-PCD)

The Nivel-PCD consists of routine primary care EHR data from over 1.8 million patients from over 500 general practices across the Netherlands in 2019. The database includes information on diagnoses, consultations, prescribed medication and (laboratory) measurements.

In the Netherlands, all non-institutionalised inhabitants are obligatorily registered with one general practitioner (GP) as their primary care provider. In general practices, all encounters are linked to International Classification of Primary Care version 1 (ICPC-1) diagnostic codes in the EHR.[15] Since GPs have a central role in Dutch primary care as the gatekeepers of referrals to specialised care, all specialists report their findings back to the GP. The GP then links this correspondence to either an existing or a new ICPC-1 code. Therefore, GPs have a complete overview of morbidity of their patients. Nivel-PCD constructs episodes of illness with associated start and end date using multiple markers of diagnostic information in the EHRs (see Supplementary Methods for details). This process has been described previously and has been shown to provide an accurate assessment of morbidity rates.[16]

Prescriptions are recorded according to the Anatomical Therapeutic Chemical classification system. Since GPs in the Netherlands are often tasked with providing repeat prescriptions for medication initiated by specialists, Nivel-PCD widely covers prescriptions for chronic morbidities initiated by both GPs and specialists. Other data including but not limited to sex, age, smoking status and body measurements are stored as separate parameters. Due to prohibitions by Dutch law, information on ethnic background is not systematically recorded in EHRs.[17]

Data extraction

We used data from 1 January 2013 to 31 December 2018. Baseline was 1 January 2014, with the EHR data recorded during the calendar year 2013 serving as baseline data in order to include only recent measurement and medication data. When multiple entries for one variable were available in 2013, we used the recorded entry closest to baseline, 1 January 2014. Detailed operational definitions for the CHARGE-AF variables are shown in the Supplementary Methods.

We assumed absence of baseline morbidity or smoking when no episode of illness or status as active smoker was recorded for a disease prior to baseline.[18] Age and sex were available for all patients. When a patient had no recorded height, weight, SBP or DBP during calendar year 2013, we considered these measurements as missing. We applied no imputation techniques for missing CHARGE-AF measurement variables since we expected these data not to be missing at random.

Study population

We included patients aged 40 years or older and free of AF at baseline who were registered at one of the Nivel-PCD associated practices during the full calendar year 2013. We excluded patients from practices without follow-up data beyond 2013 since inclusions of such data would automatically render patients without follow-up data. Among included patients, we distinguished those with missing data for one or more of the four body measurements included in the CHARGE-AF model (height, weight, SBP and DBP) – 'incomplete cases' – and those with baseline data available for all these measurements – 'complete cases'.

Outcomes

The primary outcome was newly diagnosed AF. We defined AF as the recording of the ICPC-1 code K78 'AF or atrial flutter' or any recording of a treating physician for AF or participation in AF care programme. We defined the date of AF diagnosis as the first date associated with either of these AF entries. We were unable to ascertain death as the reason for loss of follow-up, since date and cause of death are not validly recorded in primary care EHRs.

Follow-up

Patient registration at a Nivel-PCD associated practice is assessed quarterly. Reasons for loss of follow-up in Nivel-PCD are death, exclusion of practice due to low quality data, technical failure of data extraction or a patient moving away from their Nivel-PCD associated practice. We defined loss to follow-up as the first day of a period of four or more consecutive quarters of absent data, or the first day of a period of con-

secutive quarters of absent data that included the last quarter of calendar year 2018. We censored follow-up in our analyses at time of AF diagnosis, loss to follow-up or end of the 5-year observation window (31 December 2018), whichever occurred first.

The CHARGE-AF model

We calculated each individual's CHARGE-AF predicted 5-year AF risk using the formula from the original derivation article[5]: 1–0.9718412736^exp (ΣbX – 12.5815600). Here, ΣbX is calculated as: (age in years/5) * 0.5083 + ethnicity (Caucasian/white) * 0.46491 + (height in centimetres/10) * 0.2478 + (weight in kg/15) * 0.1155 + (SBP in mm Hg/20) * 0.1972 – (DBP in mm Hg/10) * 0.1013 + current smoking * 0.35931 + antihypertensive medication use * 0.34889 + DM * 0.23666 + heart failure * 0.70127 + MI * 0.49659.

The Dutch population is ~95% Caucasian/white,[19] and Nivel-PCD contains a representative sample of Dutch inhabitants.[20] In absence of ethnicity data in Nivel-PCD, we therefore assumed ethnicity as Caucasian/white for all Nivel-PCD subjects. We chose this approach in accordance with previous work and because the CHARGE-AF formula results in a prediction of an individual's absolute 5-year AF risk. Leaving ethnicity out of the formula would lead to a systematic underestimation of absolute risk by the model.[21]

We assessed the relative contribution of each CHARGE-AF variable to an increase in baseline CHARGE-AF score by multiplying the mean value of each risk factor by its CHARGE-AF coefficient within successive strata of baseline CHARGE-AF risk.

Statistical analysis

We reported continuous variables as means ± SD, ordinal variables as median and interquartile range (IQR), and dichotomous variables as number and percentages. We assessed differences in baseline parameters using the unpaired t-test with Welch's approximation, the Wilcoxon rank-sum test and the χ^2 test where appropriate. We assessed significance in all analyses at the 0.05 level.

We estimated the cumulative 5-year AF incidence using survival analysis and presented it as number and percentages as well as incidence per 1000 person years using survival-time analysis. We plotted the cumulative AF incidence using a Kaplan-Meier failure plot.

In validation of the CHARGE-AF model for 5-year AF risk, we assessed discrimination by the C-statistic and 95% CI. We assessed calibration by the calibration plot according

to deciles of baseline CHARGE-AF risk,[22] by the calibration slope of the linear predictor and its 95% CI[22] and by the Hosmer-Lemeshow goodness-of-fit test modified for survival analyses by D'Agostino and Nam.[23] A Nam-D'Agostino χ^2 with p value <0.05 indicated insufficient calibration.[24] A calibration slope significantly smaller than 1 indicated overfitting of the CHARGE-AF model when applied to our cohort.[22] Finally, we assessed calibration by the Kaplan-Meier failure function stratified according to baseline CHARGE-AF risk. For this, we used categories <2.5%, 2.5%–5% and >5% predicted risk in accordance with the original CHARGE-AF publication.[5]

We compared CHARGE-AF's discriminatory abilities for risk of newly diagnosed AF with that of two other easily obtainable predictors that have previously been shown to predictive of new AF: age alone as continuous linear variable and the CHA_2DS_2-VASc score[25] as a categorical variable.[4,6,26-29] We assessed net reclassification improvement (NRI) by the NRI index and 95% CI for 5-year AF of CHARGE-AF versus age alone as well as CHARGE-AF versus CHA_2DS_2-VASc using 200 bootstrap samples in low, intermediate and high AF risk categories with cut-offs at 2.5% and 5% predicted AF risk.[22] Data for age and CHA_2DS_2-VASc score were complete in all participants.

We performed stratified analyses according to age, sex and CHA_2DS_2-VASc score in all validation analyses in order to assess whether CHARGE-AF, CHA_2DS_2-VASc score and age would perform better among clinically relevant subgroups, and whether different predictors for newly diagnosed AF outperformed others in any of these subgroups.

Finally, we assessed the clinical implications of applying different cut-offs for dichotomisation of baseline CHARGE-AF risk into high-risk and low-risk groups. We applied cut-offs 2.5%, 5% and 10% baseline CHARGE-AF risk and assessed for each cut-off: the proportion of patients that would be counted as high risk; the proportion of total 5-year AF cases that would be among high-risk patients; 5-year AF incidence among those counted as high-risk patients; the proportion of high-risk patients with a CHA_2DS_2-VASc score ≥2 (corresponding with the need for oral anticoagulation therapy[2]); and the proportion of high-risk 5-year AF cases with a CHA_2DS_2-VASc score ≥2. In order to formally test whether the applied cut-offs were able to discriminate between high and low risk of 5-year AF incidence, we provided the unadjusted HR for 5-year AF incidence of high-risk patients with low-risk patients as reference using a Cox proportional hazards model.

We used Stata V.15.0[30] and R V.1.1.463[31] using the haven, nricens, polspline, rms, survival and survminer packages for our analyses.

Ethics and study approval

Dutch law allows the use of EHRs for research purposes under certain conditions. According to this legislation, neither obtaining informed consent from patients nor approval by a medical ethics committee is obligatory for this type of observational studies containing no directly identifiable data (Dutch Civil Law, Article 7:458).[17]

RESULTS

We included 668 955 patients aged ≥40 years from 328 Nivel-PCD practices with follow-up data available for ≥1 year after baseline. Of these, 551 655 patients had missing data for ≥1 of the CHARGE-AF measurements height, weight, SBP and DBP during 2013. Of the 117 300 patients with complete CHARGE-AF baseline data, 5825 (4.97%) had prevalent AF at baseline. The remaining 111 475 patients free of AF and with complete CHARGE-AF variables at baseline (17.2% of all patients aged ≥40 years and free of AF) constituted the validation sample of complete cases (see study flowchart in Supplementary Figure S1).

Patients with complete CHARGE-AF baseline data

Among complete cases, mean age was 65.5 ± 11.4 years, 52.5% were female and median CHA_2DS_2-VASc was 3 (IQR 2–4) (Table 1). The distribution of baseline CHARGE-AF risk was skewed with more than half of all patients with complete baseline CHARGE-AF data having a predicted 5-year AF risk <5% (Supplementary Figure S2, panel A). Age was the major factor driving an increase in baseline CHARGE-AF risk (Supplementary Figure S2, panel B).

Compared with those who remained free of AF, patients who were diagnosed with new AF during follow-up were older and had higher overall cardiovascular burden, except for DBP, burden of hypercholesterolaemia and proportion of current smokers that were lower. For a comparison between patients with and those without complete baseline CHARGE-AF data, see Supplementary Results.

AF incidence and follow-up

There were 5264 cases of new AF among complete CHARGE-AF cases during the 5-year follow-up window (4.7%; 13.6/1000 person-years; see Supplementary Figure S3, panel A, for the Kaplan-Meier plot). Mean follow-up in the sample was 3.5 ± 1.7 years. Main reason for loss to follow-up was practices' data being excluded from further analysis due to low quality data (see Supplementary Figure S3, panel B, for the number of practices and patients at risk during follow-up).

Table 1. *Baseline characteristics of the study sample with complete baseline CHARGE-AF data*

	All (n=111,475)	AF during follow-up (n=5,264)	No AF during follow-up (n=106,211)	p-value for difference*
Age, years	65.5 ± 11.4	73.1 ± 9.4	65.2 ± 11.4	<0.001
Female	58,549 (52.5%)	2,572 (48.9%)	55,977 (52.7%)	<0.001
SBP, mmHg	137.3 ± 16.3	139.5 ± 17.3	137.2 ± 16.2	<0.001
DBP, mmHg	80.5 ± 10.5	78.8 ± 10.8	80.6 ± 10.5	<0.001
Height, cm	170.0 ± 9.9	170.3 ± 9.9	170.0 ± 9.9	0.01
Weight, kg	82.5 ± 16.8	83.8 ± 17.2	82.4 ± 16.8	<0.001
Antihypertensive medication	79,057 (70.9%)	4,494 (85.4%)	74,563 (70.2%)	<0.001
Hypertension	74,149 (66.5%)	3,864 (73.4%)	70,285 (66.2%)	<0.001
Diabetes mellitus	47,557 (42.7%)	2,514 (47.8%)	45,043 (42.4%)	<0.001
Heart failure	4,693 (4.2%)	562 (10.7%)	4,131 (3.9%)	<0.001
Myocardial infarction	5,404 (4.9%)	391 (7.4%)	5,013 (4.7%)	<0.001
Current smoking	15,774 (14.2%)	600 (11.4%)	15,174 (14.3%)	<0.001
Stroke	7,462 (6.7%)	472 (9.0%)	6,990 (6.6%)	<0.001
TIA	3,339 (3.0%)	224 (4.3%)	3,115 (2.9%)	<0.001
Pulmonary embolism	506 (0.5%)	31 (0.6%)	475 (0.4%)	0.14
Angina pectoris	10,167 (9.1%)	750 (14.3%)	9,417 (8.9%)	<0.001
CHA_2DS_2-VASc **CHA_2DS_2-VASc ≥2**	3 (IQR 2-4) 88,538 (79.4%)	4 (IQR 3-5) 4,866 (92.4%)	3 (IQR 2-4) 83,672 (78.8%)	<0.001 <0.001
Asthma	13,262 (11.9%)	652 (12.4%)	12,610 (11.9%)	0.26
COPD	12,523 (11.2%)	879 (16.7%)	11,644 (11.0%)	<0.001
Atherosclerosis	6,367 (5.7%)	416 (7.9%)	5,951 (5.6%)	<0.001
Hypercholesterol-aemia	19,427 (17.4%)	694 (13.2%)	18,733 (17.6%)	<0.001
Gout	7,639 (6.9%)	589 (11.2%)	7,050 (6.6%)	<0.001
Enrolled in care program for:				
Asthma	1,846 (1.7%)	77 (1.5%)	1,769 (1.7%)	0.26
COPD	4,777 (4.3%)	335 (6.4%)	4,442 (4.2%)	<0.001
Diabetes mellitus	35,640 (32.0%)	1,943 (36.9%)	33,697 (31.7%)	<0.001
Any care program	40,468 (36.3%)	2,212 (42.02%)	38,256 (36.0%)	<0.001

Data are number (percentage), mean ± standard deviation or median (IQR).
AF, atrial fibrillation; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age, Diabetes and previous Stroke or Transient Ischaemic Attack, Vascular disease and female Sex category; COPD, chronic obstructive pulmonary disorder; DBP, diastolic blood pressure; IQR, interquartile range; SBP, systolic blood pressure; TIA, transient ischaemic attack.
* Difference between those with and without AF during follow-up

CHARGE-AF validation

Validation of CHARGE-AF among all patients with complete baseline CHARGE-AF data resulted in a C-statistic of 0.736 (95% CI 0.727 to 0.744), a Nam-D'Agostino χ^2 of 901.8 (p<0.001) and a calibration slope of 0.69 (95% CI 0.67 to 0.71) (Table 2). The calibration plot showed a slight underestimation of AF risk among lower deciles of

CHARGE-AF risk but strong overestimation of AF risk in the higher CHARGE-AF deciles (Figure 1, panel A). The Kaplan-Meier plot stratified by risk categories <2.5%, 2.5%–5% and >5% CHARGE-AF predicted 5-year risk indicated an accurate estimation of observed 5-year AF risk in the overall sample of complete cases (Figure 1, panel B).

CHARGE-AF showed superior discrimination to CHA_2DS_2-VASc as well as age alone as the predictor in both the overall and all stratified analyses. Results of the stratified analyses on CHARGE-AF are shown in theSupplementary Results. CHARGE-AF resulted in significant reclassification improvement versus both CHA_2DS_2-VASc (NRI index: 0.24; 95% CI 0.22 to 0.25) and age alone (NRI index: 0.05; 95% CI 0.04 to 0.06).

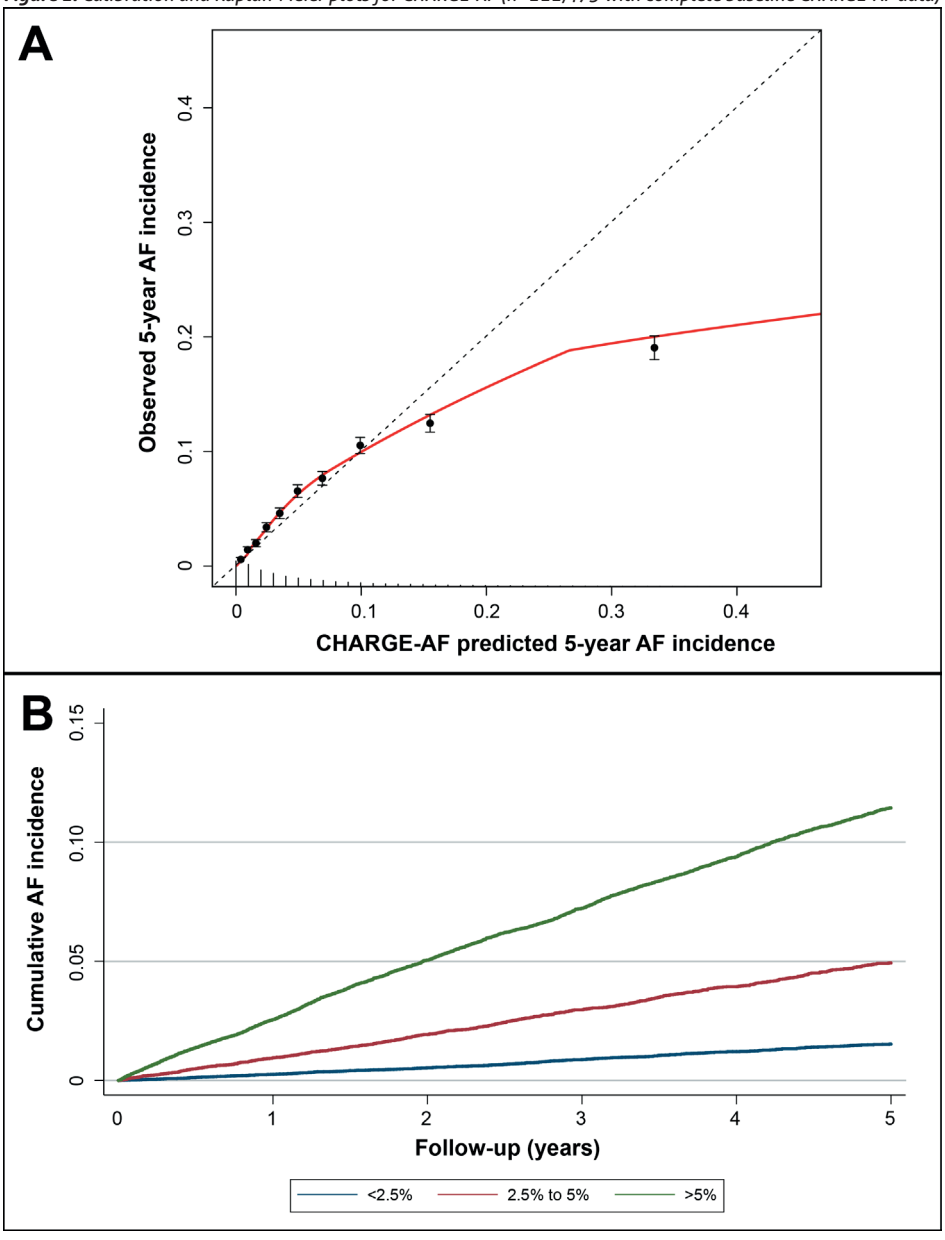
Application of different CHARGE-AF cut-offs

Figure 2 shows the analysis on dichotomisation of CHARGE-AF risk at cut-offs 2.5%, 5% and 10%. The high-risk groups showed significantly higher AF incidence over time in all comparisons as assessed by the unadjusted HRs for high-risk versus low-risk patients. Cut-offs at 2.5%, 5% and 10% CHARGE-AF risk would have classified 65%, 45% and 25% of patients with complete CHARGE-AF baseline data as 'high risk', respectively. Routine care 5-year AF incidence among the high-risk patients at these cut-offs was 6.7%, 8.0% and 9.8%, respectively. In all high-risk groups, >95% observed AF cases had CHA_2DS_2-VASc ≥2 at baseline ($p<0.001$ for difference with proportion of CHA_2DS_2-VASc ≥2 among low-risk AF cases in all comparisons).

DISCUSSION

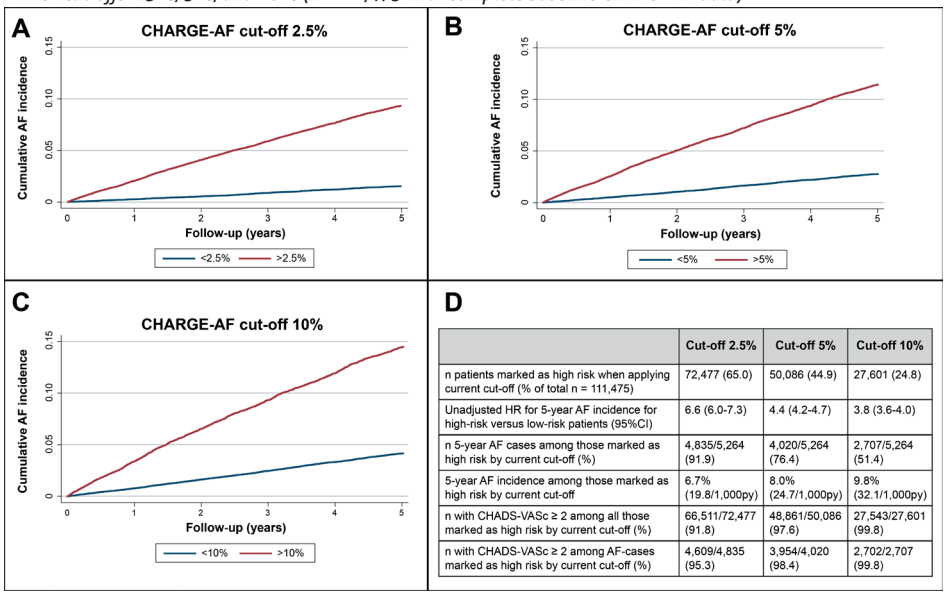
In a routine primary care EHR database representative of the Netherlands, one in six patients aged 40 years and older was free of AF and had complete baseline CHARGE-AF data. These patients had significantly higher 5-year AF incidence and cardiovascular morbidity than those with ≥1 missing CHARGE-AF variables. Validation of CHARGE-AF among complete cases showed that despite overestimation of absolute 5-year AF risk in those with the highest baseline CHARGE-AF scores, the model had overall sufficient discrimination for 5-year AF risk and was able to accurately group patients according to predefined risk categories. CHARGE-AF had superior discrimination for 5-year risk of AF compared with CHA_2DS_2-VASc and age alone. Explorative analyses on the application of different CHARGE-AF cut-offs for patient selection indicated that cut-offs at 2.5%, 5% and 10% all have potential merits for use in AF risk stratification.

Figure 1. *Calibration and Kaplan-Meier plots for CHARGE-AF (n=111,475 with complete baseline CHARGE-AF data)*



AF, atrial fibrillation; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology.
Panel A, Calibration plot for CHARGE-AF. The points indicate intersects of observed and expected for each decile of baseline CHARGE-AF risk, with brackets indicating the 95%CI of observed AF probability during 5-year follow-up in each decile. The red line indicates the trend for CHARGE-AF calibration in the sample. When the intersect of observed and expected AF incidence exceeds the dotted line, this indicates underestimation of AF risk by CHARGE-AF for that decile. When the intersect of observed and expected AF incidence is below the dotted line, this indicates overestimation of AF risk by CHARGE-AF for that decile. The spikes on the x axis indicate the distribution of AF-free survivors by CHARGE-AF risk; Panel B, Kaplan-Meier plot of AF incidence stratified according to baseline CHARGE-AF predicted risk categories <2.5%, 2.5% to 5%, and >5%

Figure 2. *Kaplan-Meier plots and outcomes table of AF incidence when dichotomized according to baseline CHARGE-AF risk cut-offs 2.5%, 5%, and 10% (n=111,475 with complete baseline CHARGE-AF data)*



	Cut-off 2.5%	Cut-off 5%	Cut-off 10%
n patients marked as high risk when applying current cut-off (% of total n = 111,475)	72,477 (65.0)	50,086 (44.9)	27,601 (24.8)
Unadjusted HR for 5-year AF incidence for high-risk versus low-risk patients (95%CI)	6.6 (6.0-7.3)	4.4 (4.2-4.7)	3.8 (3.6-4.0)
n 5-year AF cases among those marked as high risk by current cut-off (%)	4,835/5,264 (91.9)	4,020/5,264 (76.4)	2,707/5,264 (51.4)
5-year AF incidence among those marked as high risk by current cut-off	6.7% (19.8/1,000py)	8.0% (24.7/1,000py)	9.8% (32.1/1,000py)
n with CHADS-VASc ≥ 2 among all those marked as high risk by current cut-off (%)	66,511/72,477 (91.8)	48,861/50,086 (97.6)	27,543/27,601 (99.8)
n with CHADS-VASc ≥ 2 among AF-cases marked as high risk by current cut-off (%)	4,609/4,835 (95.3)	3,954/4,020 (98.4)	2,702/2,707 (99.8)

AF, atrial fibrillation; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age, Diabetes and previous Stroke or Transient Ischaemic Attack, Vascular disease and female Sex category; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology model for AF; HR, hazard ratio; py, person years; Nivel-PCD, Netherlands Institute for Health Services Research Primary Care Database; 95%CI, 95% confidence interval.
Panel A, Kaplan-Meier (KM) plot of AF incidence dichotomized according to baseline CHARGE-AF predicted risk cut-off 2.5%; Panel B, KM plot of AF incidence dichotomized according to baseline CHARGE-AF predicted risk cut-off 5%; Panel C, KM plot of AF incidence dichotomized according to baseline CHARGE-AF predicted risk cut-off 10%; Panel D, table of outcomes if CHARGE-AF risk cut-offs 2.5%, 5% and 10%, respectively, had been applied for patient selection.

Clinical implications

Outcomes of this work are relevant to the prospect of using clinical risk models as triage test for AF screening, while maintaining low cost in their risk assessment efforts. We showed that those with complete CHARGE-AF variables as per routine primary care constitute a small but highly relevant subset for AF screening. The model's high accuracy in predicting absolute 5-year risk for predefined risk categories suggests that the model can be used to reliably differentiate between low and high AF risk among complete cases. Moreover, CHARGE-AF outperformed two other predictors that have been employed to select for AF screening eligibility, as assessed by both the C-statistic and NRI index. This work therefore encourages researchers in the field of community AF screening to consider CHARGE-AF as a triage test for patient selection.

We provided data on how the choice for a baseline CHARGE-AF cut-off for classifying patients as 'high risk' could translate into actual patient selection for screening. The sensitivity of 'baseline CHARGE-AF' as a triage test for 5-year observed new AF ranged

between 51% at CHARGE-AF cut-off 0.1% and 92% at CHARGE-AF cut-off 0.025. Since these findings are based on simple routine care EHR data acquired without imputation or text mining techniques, CHARGE-AF showed its potential for low-cost automated, remote AF risk stratification. This suggests a lower need for a baseline visit prior to screening. The model could also be used as an alert for clinicians to check for AF in the subset of patients with complete data through routine care.

We emphasise that the outcome in our work was 5-year risk of an AF diagnosis acquired through routine care. To our knowledge, there have been no clinical studies on the efficacy of CHARGE-AF as a triage test for patient selection for screening. Although our work does not provide concrete recommendations to practising GPs on whether and how to best use CHARGE-AF in selecting patients for further rhythm analysis, it points to CHARGE-AF as a model with the highest potential for this purpose.

Comparison with previous work

This study diverges from previous CHARGE-AF validation studies in that it made an explicit attempt to bridge the gap between model validation and subsequent application as a tool for patient selection in community AF screening. To our knowledge, we were the first to provide detailed information on how selecting at different cut-offs would translate into numbers of patients to be screened and percentage of AF yield to be expected in a large routine primary care dataset.

The C-statistic for CHARGE-AF in our study (0.74) was lower than in the aggregate CHARGE-AF derivation cohorts (0.77) but higher than the summary C-statistic in a recent meta-analysis of CHARGE-AF for 5-year AF risk in community cohorts (0.72).[5,11] Possible explanations for difference with the original CHARGE-AF article are that the model was calibrated to fit the derivation data, that our dataset had a lower percentage of women in whom CHARGE-AF performed better than in men and that the ethnic diversity was lower in Nivel-PCD. Applying the same age restrictions to our dataset as were used in the derivation article (46–94 years) resulted in the same C-statistic as the current overall analysis (data not shown).

A recent study validated CHARGE and CHA_2DS_2-VASc based on a large routine care EHR dataset from seven hospitals in the USA from which they excluded patients with non-complete measurement data.[18] Results of validation of CHARGE-AF and CHA_2DS_2-VASc were similar to ours. The main difference between this study and ours is the population. Since Dutch primary care EHR data covers all non-institutionalised inhabitants, with all secondary care facilities reporting back to GPs, Nivel-PCD is likely to have a wider coverage of the population than a regional agglomeration of hospitals. The

percentage of patients with complete measurements, however, was greater in Hulme et al's[18] hospital-derived dataset where measurements may be more routinely taken. Both studies, however, provide evidence that routine care data can be used to assess risk of AF in patients with complete measurement data at baseline, with each study having its own merits in terms of generalisability to different care settings.

Although our patient selection differed from the derivation study as well as previous validation studies that were performed in largely unselected community cohorts, a number of observations are common among validation studies of CHARGE-AF, age alone and CHA_2DS_2-VASc for new AF. Mainly, these studies, like ours, found that CHARGE-AF outperformed CHA_2DS_2-VASc and age alone as predictors for new AF and that CHARGE-AF showed higher C-statistics among lower risk subgroups within their sample.[4-10,26-29,32-34]

Our study corroborates the findings that patients with complete recent baseline measurement data as per routine care were older and had higher burden of cardiovascular comorbidity than those with missing measurements.[12] Our study expands on that by showing that having complete measurements through routine primary care is also associated with higher 5-year risk of AF.

We were unable to validate a number of other models developed for AF risk prediction in community cohorts due to restrictions in data availability in Dutch primary care EHRs.[6,8,18,26,35,36] We refrained from recalibration and augmentation of CHARGE-AF to better fit our sample, since our aim was to validate CHARGE-AF, not to improve its risk prediction in a specific population.[4,5,7,10,27,32-34,37]

Future work

Our work relied heavily on the assumption that AF risk through routine care is correlated with AF yield through active screening. Although there are few studies to assess the validity of this hypothesis, one recent pilot study that selected individuals with both age ≥65 years and high CHA_2DS_2-VASc score for screening with continuous ECG monitoring found promising results.[38] Post hoc analyses on the added value of multivariable risk models in previous AF screening studies would be welcomed.

Our work shows that higher completeness of primary care EHR data is needed. Since such data completeness will likely not be achieved in the foreseeable future, research should focus on ways of handling missing data in primary care EHRs while still achieving accurate risk prediction. Until then, models that do not rely on measurement variables may be the model of choice for remote, automatic AF risk assessment in

primary care settings. Finally, the ethical implications of using EHR data to remotely brand individuals as 'at high risk of AF and stroke' deserve further research.[3]

Strengths and limitations

This work had a number of strengths. First, our validation of CHARGE-AF in patients with complete data through routine primary care enabled an assessment of CHARGE-AF's merits as a potential triage test for AF screening without the need for a resource-intensive baseline visit for data collection. Second, given the use of a large dataset that encompasses a representative sample of primary care patients in the Netherlands, and considering the role of GPs in the Netherlands where all inhabitants are registered at a GP and where all secondary care providers report health outcomes back to GPs, results from this study are likely generalisable to similar settings.[20] Third, we included a comparison of patients with and without complete baseline CHARGE-AF measurements. This enabled us to show that patients with complete baseline parameters had higher AF risk and higher cardiovascular comorbidity and more often had a CHA_2DS_2-VASc score ≥2. An AF diagnosis in these patients is therefore both more likely and more often relevant in terms of anticoagulation initiation.[2] Finally, we provided researchers interested in using CHARGE-AF as a selection tool for AF screening among complete cases with ample data to assess which baseline CHARGE-AF cut-off may be most viable for such purposes.

Our study's primary strength was also its most prominent limitation. Due to its restriction to patients with complete CHARGE-AF measurements, results of this study are not generalisable to the community at large. Additional work is therefore required to assess how CHARGE-AF can be used to reliably assess risk for incident AF in the larger community while still refraining from the need to perform baseline visits. Second, the nature of a routine primary care database dictates that diagnosis and correct registration of morbidities had been at treating physicians' discretion. Most notably, this may increase the risk of verification bias in diagnosing incident AF as well as underestimation of prevalence of baseline comorbidities.[39,40] Third, one of CHARGE-AF's variables—ethnicity—was missing altogether from the database due to restrictions in Dutch primary healthcare regulations. Although our evaluation of the relative contribution of variables to increments in baseline risk showed ethnicity to play only a minor role in overall AF risk assessment when assumed as Caucasian/white in all individuals, it is unclear how information on this variable might have influenced the validity of predictions in non-Caucasian individuals. Finally, it is unclear whether the classification of AF and MI diagnoses as non-chronic episodes in Nivel-PCD, with a patient's AF or MI episode being inactivated after a contact-free period of 1 year, may have affected AF prevalence and CHARGE-AF score before baseline and AF incidence during follow-

up.[16] Prior work on Nivel-PCD showed that extending this period from 1 to 2 years did not lead to significantly different incidence rates.[16] We sought to further ameliorate this limitations by using a 1-year baseline window, which has been shown to lead to a more accurate representation of disease prevalence in routine care EHRs than point prevalence.[20] We hereby effectively extended the non-contact window after which AF and MI patients would become false-negative from 1 to 2 years before baseline.

Ethics statements: Patient consent for publication: Not required.

Ethics approval: This study has been approved according to the governance code of Nivel-PCD under number NZR-00318.043.

Acknowledgments: We would like to thank Wim Busschers for R code templates for risk model validation and Evert Karregat for input in early stages of manuscript preparation.

Correction notice: This article has been corrected since it first published. The provenance and peer review statement has been included.

Funding: This work was supported by the Netherlands Organisation for Health Research and Development (ZonMw) (80-83910-98-13046) and the European Research Council under the European Union's Horizon 2020 research and innovation programme (648 131). The authors had full autonomy in design, conduct and reporting of the manuscript.

Contributors: JCLH performed data preparation, data analysis and data presentation and was primarily responsible for manuscript preparation. MMJN supervised data preparation, data analysis and data presentation at Nivel-PCD and provided valuable input to the manuscript. CA supervised statistical analysis and provided valuable input to the manuscript. REH, WAML and HCPMvW were responsible for project supervision at AUMC and provided valuable input to the manuscript.

Conflict of interest: None declared.

Data availability statement: Data are deidentified routine primary care electronic health records licensed by the Netherlands Institute for Health Services Research Primary Care Database. For requests for and information on data usage: directie@nivel.nl.

Provenance and peer review: Not commissioned; externally peer reviewed.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY METHODS

Nivel-PCD episodes of illness construction in current dataset

Episodes of illness that are deemed 'chronic' in Nivel-PCD, including e.g. hypertension and diabetes mellitus (DM), remain active throughout extractions. This allowed for inclusion of all recorded chronic episodes of illness prior to 1 January, 2014, including those with their latest GP encounters prior to calendar year 2013. Episodes of illness that are classified as 'long-lasting reversible diseases' in Nivel-PCD, a category that includes AF and MI, are available in an annual extraction if the last GP encounter was up to 1 year prior to extraction. We were thus able to include all long-lasting reversible diseases of which a patient's EHR contained a recorded GP encounter (physical or administrative) on or later than 1 January, 2012. Prior Nivel-PCD analyses have shown that extension of this 1-year contact-free interval does not lead to significant differences in long-lasting reversible disease incidence.[16]

Operational variable definitions

- Atrial fibrillation (AF): entry of ICPC-1 code K78 (AF/flutter) and/or data codes 3451 (treating physician for AF) or 3838 (enrolment in care program for AF);
- Age: the discrete number of years attained in the year 2013 since year of birth;
- Sex: male or female;
- Systolic blood pressure (SBP): latest recording in 2013 of data codes 1744 (SBP), 2055 (SBP home measurement), 2668 (mean SBP in 24-hour measurement), 3336 (mean SBP in 30-minute measurement), 1745 (SBP lying down), 2189 (SBP standing), or 1794 (SBP of the arm when used for ankle-brachial index test). We applied a hierarchy in which code to use, in the order of aforementioned data codes. We first looked at entries for data code 1744 and when available we used the latest entry in 2013. If there was no entry for data code 1744, we looked at entries for data code 2055. If there was no entry for data code 2055, we looked at data code 2668, etc. until data code 1794. In order to prevent inclusion of values erroneously entered by GP personnel, we included only SBP values 25-250mmHg;
- Diastolic blood pressure (DBP): latest recording in 2013 of data codes 1740 (DBP), 2056 (DBP home measurement), 2669 (mean DBP in 24-hour measurement), 3337 (mean DBP in 30-minute measurement), 1741 (DBP lying down), or 2188 (DBP standing). We applied a hierarchy in which code to use, in the order of aforementioned data codes. We first looked at entries for data code 1740 and when available we used the latest entry in 2013. If there was no entry for data code 1740, we looked at entries for data code 2056. If there was no entry for data code

2056, we looked at data code 2669, etc. until data code 2188. In order to prevent inclusion of values erroneously entered by GP personnel, we included only SBP values 25-250mmHg;

- Weight: latest recording in 2013 of data codes 357 (weight) or 2408 (weight home measurement). When entries for these data codes were absent in 2013, but data codes 560 (height) and 1272 (body mass index, BMI) were present, we calculated weight as BMI*weight2 and used the latest recordings in 2013. In order to prevent inclusion of values erroneously entered by GP personnel, we included only weight values 30-300kg;
- Height: latest recording in 2013 of data code 560 (height). When an entry for data code 560 was absent, but data codes for weight and BMI were both present in 2013, we calculated height in centimeters as 100*√(weight/BMI) and used the latest recordings in 2013. In order to include only realistic values, and to prevent inclusion of values erroneously entered by GP personnel, we included only height values 130-230cm. Values below 130 were multiplied by 100 in order to include data entered as meters instead of centimeters. We subsequently applied the same limits of 130-230cm;
- Antihypertensive medication: ATC subcodes for C02 (antihypertensives) and/or C03 (diuretics), C04 (peripheral vasodilators), C05 (vasoprotectives), C07 (beta blocking agents), C08 (calcium channel blockers), or C9 (agents acting on the renin-angiotensin system);
- Hypertension: entry of ICPC-1 codes K86 (uncomplicated hypertension) and/or K87 (hypertension with involvement target organs) or data code 1694 (hypertension comorbidity);
- Diabetes mellitus (DM): entry of ICPC-1 code T90 (DM) and/or data code 2206 (treating physician for DM);
- Heart failure (HF): entry of ICPC-1 code K77 (HF) and/or data codes 3016 (treating physician for HF), 2722 (NYHA severity of HF symptoms) or 1643 (HF comorbidity);
- Myocardial infarction (MI): entry of ICPC-1 code K75 (acute MI) and/or data code 1693 (MI comorbidity);
- Current smoking: classified as current smoker when indicated as smoker as per data codes 1739 (smoking) and/or 1992 (number of (rolling tobacco) cigarettes per day), 1993 (number of cigarettes per day), 1996 (wants to quit smoking in short term) or 2405 (motivation to quit smoking), and not followed in time (but before 01-01-2014) by an indication of having quit smoking as per data codes 1739 (smoking) and/or 2003 (quit smoking since);
- Stroke: entry of ICPC-1 code K90 (stroke/cerebrovascular accident) and/or lab code 2132 (cerebral ischaemia history comorbidity);

- Transient ischemic attack (TIA): entry of ICPC-1 code K89 (transient cerebral ischaemia);
- Pulmonary embolism (PE): entry of ICPC-1 code K93 (PE);
- Angina pectoris: entry of ICPC-1 code K74 (angina pectoris);
- Vascular disease: entry of ICPC-1 codes K74 (angina pectoris) and/or K91 (atherosclerosis), K92 (other arterial obstruction/peripheral vascular disease) or MI as defined above;
- Congestive heart failure, Hypertension, Age, Diabetes and previous Stroke or Transient Ischaemic Attack, Vascular disease and female Sex category (CHA_2DS_2-VASc): 1 point for each of female sex, HF, hypertension, DM, vascular disease or age 65-74 years, plus 2 points for each of (stroke, TIA or PE) or age ≥75 years;
- Asthma: entry of ICPC-1 code R96 (asthma) and/or indication for asthma as per data codes 1598 (asthma diagnosed by) and/or 1599 (asthma goals attained), 1618 (medication adherence asthma), 1621 (avoids provoking factors asthma), 1716 (reason for failure to achieve asthma goals), 1776 (asthma management), 1806 (change asthma medication), 1822 (asthma severity), 1824 (asthma self-management), 1826 (appointment for asthma self-management), 1877 (asthma comorbidity), 2406 (treating physician for asthma), 3018 (adverse effects asthma medication), 3608 (degree of control in asthma management), 3338 (ACQ question 1), 3339 (ACQ question 2), 3340 (ACQ question 3), 3341 (ACQ question 4), 3345 (C-ACT question 1), 3346 (C-ACT question 2), 3347 (C-ACT question 3), 3348 (C-ACT question 4), 3349 (C-ACT question 5), 3828 (enrolment in care program for asthma);
- Chronic obstructive pulmonary disease (COPD): entry of ICPC-1 code R95 (COPD) and/or indication for COPD as per data codes 1779 (medication adherence COPD) and/or 1785 (COPD management), 1786 (causes for COPD exacerbation), 1807 (change COPD medication), 1818 (reason not to enrol in COPD care program), 1909 (reasons for not attaining COPD goals), 1911 (COPD diagnosed by), 2209 (GOLD classification COPD), 2399 (mean symptom score CCQ COPD), 2400 (mean function score CCQ COPD), 2401 (mean psychological score CCQ COPD), 2402 (mean limitations score CCQ COPD), 2407 (treating physician COPD), 2676 (cachexia COPD), 3013 (COPD disease burden), 3019 (adverse effects COPD medication);
- Atherosclerosis: entry of ICPC-1 code K91 (atherosclerosis);
- Hypercholesterolaemia: entry of data code 2053 (hypercholesterolaemia comorbidity) and/or value for data code 181 (cholesterol/HDL ratio) ≥5 mmol/L;
- Gout: entry of ICPC-1 code T92 (gout);
- Enrolment in care program for asthma: indication for enrolment in care program for asthma as per data codes 2406 (treating physician for asthma) and/or 3828 (enrolment in care program for asthma);

- Enrolment in care program for COPD: indication for enrolment in care program for COPD as per data codes 2407 (treating physician for COPD) and/or 3829 (enrolment in care program for COPD);
- Enrolment in care program for DM: Enrolment in care program for COPD: indication for enrolment in care program for DM as per data codes 2206 (treating physician for DM) and/or 3827 (enrolment in care program for DM);
- Enrolment in care program for any care program: indication for enrolment in one or more care programs of asthma, COPD or DM as defined above, or for indication for enrolment in care program for HF as per data codes 3016 (treating physician for HF) and/or 3833 (enrolment in care program for HF), or for indication for enrolment in care program for thyroid disease as per data codes 3040 (treating physician for thyroid disease) and/or 3835 (enrolment in care program for thyroid disease).

SUPPLEMENTARY RESULTS

Comparison of patients with and without complete baseline CHARGE-AF data

Supplementary Table S1 shows a comparison between those free of AF at baseline with complete baseline CHARGE-AF data and those free of AF at baseline without complete baseline CHARGE-AF data (n=538,308). Five-year AF incidence was significantly lower among incomplete CHARGE-AF cases (2.10%, p<0.001). Patients with complete CHARGE-AF baseline data were significantly older and had significantly higher burden of cardiovascular comorbidities than patients with incomplete CHARGE-AF variables at baseline. The percentage of missing CHARGE-AF measurements varied from 69.3% (SBP) to 81.3% (height). Patients with at least 1 but not all 4 CHARGE-AF measurements recorded in the EHR in 2013 had a higher mean SBP, DBP and height, but lower weight, than patients with complete baseline CHARGE-AF measurements.

Additional CHARGE-AF validation analyses

In the stratified analyses on CHARGE-AF, discrimination was consistently higher in the lower risk groups (women, age <65 years and CHA_2DS_2-VASc <2), with highest C-statistic in the subgroup of women (0.751; 95%CI: 0.740-0.763). Calibration of CHARGE-AF was insufficient in all subgroups as assessed by the Nam-D'Agostino χ^2, and the calibration slope significantly deviated from 1 in all subgroups except in patients younger than 65 and in patients with CHA_2DS_2-VASc <2 (see Table 2 in main text).

Calibration plots for the stratified CHARGE-AF analyses were similar to that of the overall analysis, except in the subgroups age <65 years and CHA_2DS_2-VASc <2. In these lower risk strata, risk prediction was accurate for all deciles, without overestimation in the highest deciles seen in the other analyses (Supplementary Figure S4).

SUPPLEMENTARY TABLES

Supplementary Table S1. *Comparison between baseline characteristics of patients free of AF at baseline within all extracted Nivel-PCD participants and those with complete CHARGE-AF variables*

	All (n = 649,783)	Complete CHARGE-AF variables at baseline (n = 111,475)	Incomplete CHARGE-AF variables at baseline (n = 538,308)	p-value for difference*
Age, years	58.2 ± 12.6	65.5 ± 11.4	56.7 ± 12.2	<0.001
Female	335,155 (51.6%)	58,549 (52.5%)	276,606 (51.38%)	<0.001
AF during 5-year follow-up	16,581 (2.55%)	5,264 (4.7%)	11,317 (2.10%)	<0.001
SBP, mmHg	137.9 ± 17.1 *450,044 (69.3%) missing*	137.3 ± 16.3	138.7 ± 18.1 *(from n = 88,264 non-missing)*	<0.001
DBP, mmHg	81.1 ± 10.8 *450,848 (69.4%) missing*	80.5 ± 10.5	81.9 ± 11.0 *(from n = 87,460 non-missing)*	<0.001
Height, cm	170.1 ± 9.9 *528,047 (81.3%) missing*	170.0 ± 9.9	171.2 ± 9.9 *(from n = 10,261 non-missing)*	<0.001
Weight, kg	82.2 ± 17.2 *516,993 (79.6%) missing*	82.5 ± 16.8	80.4 ± 18.6 *(from n = 21,315 non-missing)*	<0.001
Antihypertensive medication	188,122 (29.0%)	79,057 (70.9%)	109,065 (20.26%)	<0.001
Hypertension	177,537 (27.3%)	74,149 (66.5%)	103,388 (19.21%)	<0.001
Diabetes mellitus	72,467 (11.2%)	47,557 (42.7%)	24,910 (4.63%)	<0.001
Heart failure	12,753 (2.0%)	4,693 (4.2%)	8,060 (1.50%)	<0.001
Myocardial infarction	14,572 (2.2%)	5,404 (4.9%)	9,168 (1.70%)	<0.001
Current smoking	21,036 (3.2%)	15,774 (14.2%)	5,262 (0.98%)	<0.001
Stroke	19,380 (3.0%)	7,462 (6.7%)	11,918 (2.21%)	<0.001
TIA	8,630 (1.3%)	3,339 (3.0%)	5,291 (0.98%)	<0.001
Pulmonary embolism	2,208 (0.3%)	506 (0.5%)	1,702 (0.32%)	<0.001
Angina pectoris	28,328 (4.4%)	10,167 (9.1%)	18,161 (3.37%)	<0.001
CHA_2DS_2-VASc **CHA_2DS_2-VASc ≥2**	1 (IQR 0-2) 247,694 (38.1%)	3 (IQR 2-4) 88,538 (79.4%)	1 (IQR 0-2) 159,156 (29.6%)	<0.001 <0.001
Asthma	57,929 (8.9%)	13,262 (11.9%)	44,667 (8.30%)	<0.001
COPD	35,252 (5.4%)	12,523 (11.2%)	22,729 (4.22%)	<0.001
Atherosclerosis	13,759 (2.1%)	6,367 (5.7%)	7,392 (1.37%)	<0.001
Hypercholesterolaemia	34,135 (5.3%)	19,427 (17.4%)	14,708 (2.73%)	<0.001
Gout	23,516 (3.6%)	7,639 (6.9%)	15,877 (2.95%)	<0.001

Supplementary Table S1. *Comparison between baseline characteristics of patients free of AF at baseline within all extracted Nivel-PCD participants and those with complete CHARGE-AF variables* (*continued*)

	All (n = 649,783)	Complete CHARGE-AF variables at baseline (n = 111,475)	Incomplete CHARGE-AF variables at baseline (n = 538,308)	p-value for difference*
Enrolled in care program for:				
Asthma	4,374 (0.7%)	1,846 (1.7%)	2,528 (0.47%)	<0.001
COPD	8,572 (1.3%)	4,777 (4.3%)	3,795 (0.70%)	<0.001
Diabetes mellitus	38,969 (6.0%)	35,640 (32.0%)	3,329 (0.62%)	<0.001
Any care program	49,820 (7.7%)	40,468 (36.3%)	9,352 (1.74%)	<0.001

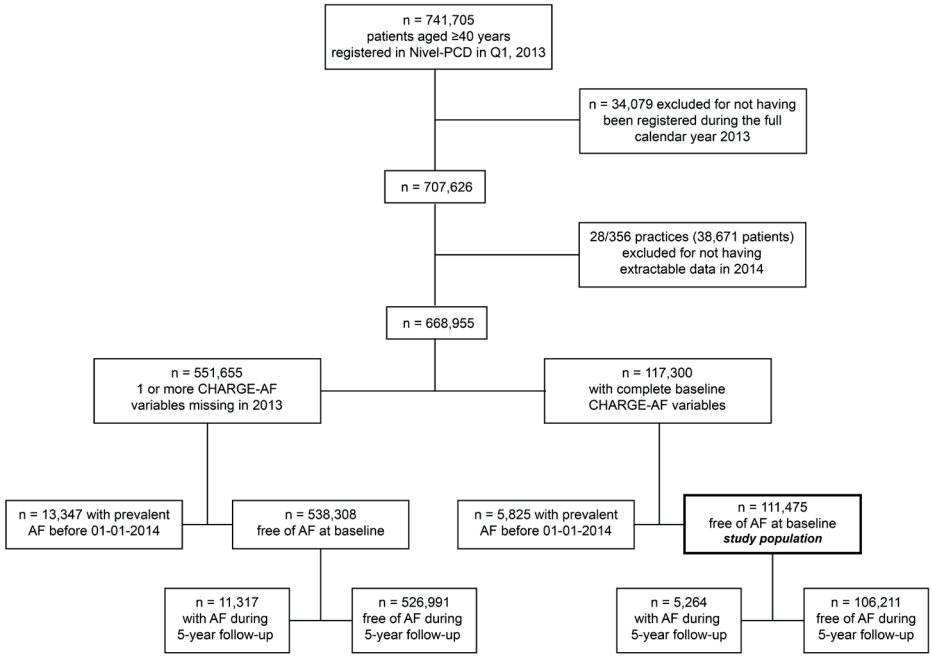
AF, atrial fibrillation; COPD, chronic obstructive pulmonary disorder; DBP, diastolic blood pressure; IQR, interquartile range; SBP, systolic blood pressure; TIA, transient ischaemic attack.

Data are number (percentage), mean ± standard deviation or median (IQR).

* Difference between those with and without complete baseline CHARGE-AF measurements

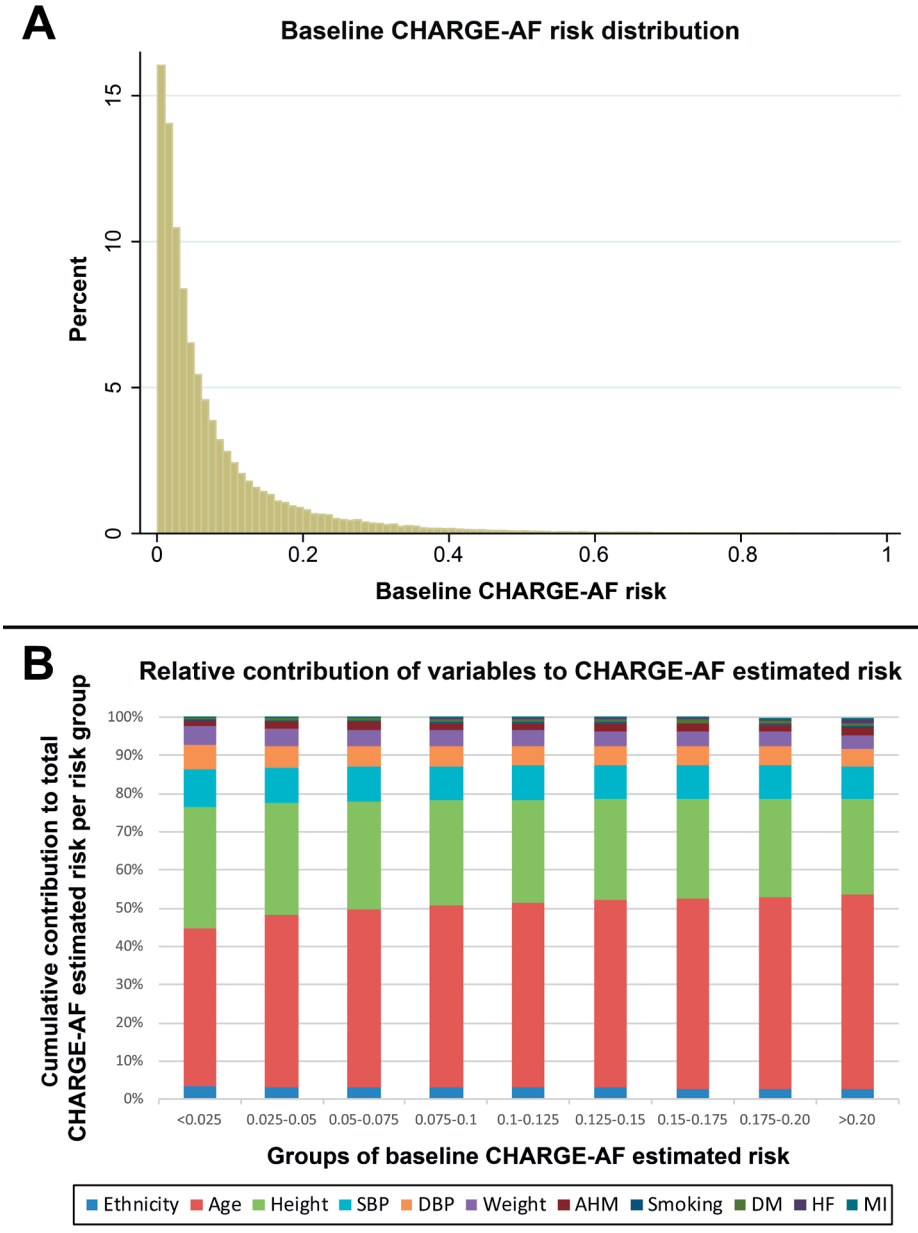
SUPPLEMENTARY FIGURES

Supplementary Figure S1. *Study flowchart*



AF, atrial fibrillation; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology model for AF; Nivel-PCD, Netherlands Institute for Health Services Research Primary Care Database; Q1, first quarter.

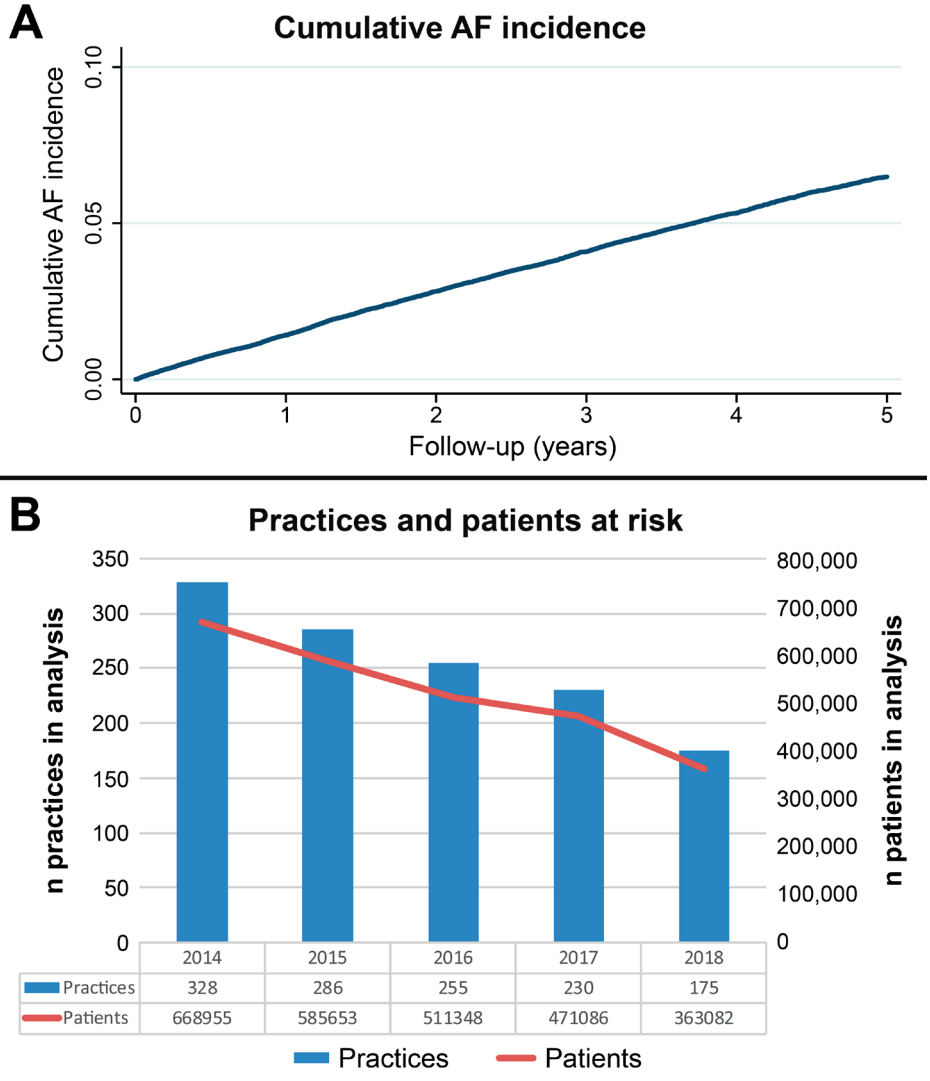
Supplementary Figure S2. *Baseline CHARGE-AF risk distribution in the sample and relative contribution of CHARGE-AF risk factors to increments in baseline risk (n = 111,475 with complete baseline CHARGE-AF data)*



AHM, antihypertensive medication use; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology model for AF; DBP, diastolic blood pressure; DM, diabetes mellitus; HF, heart failure; MI, myocardial infarction; SBP, systolic blood pressure.

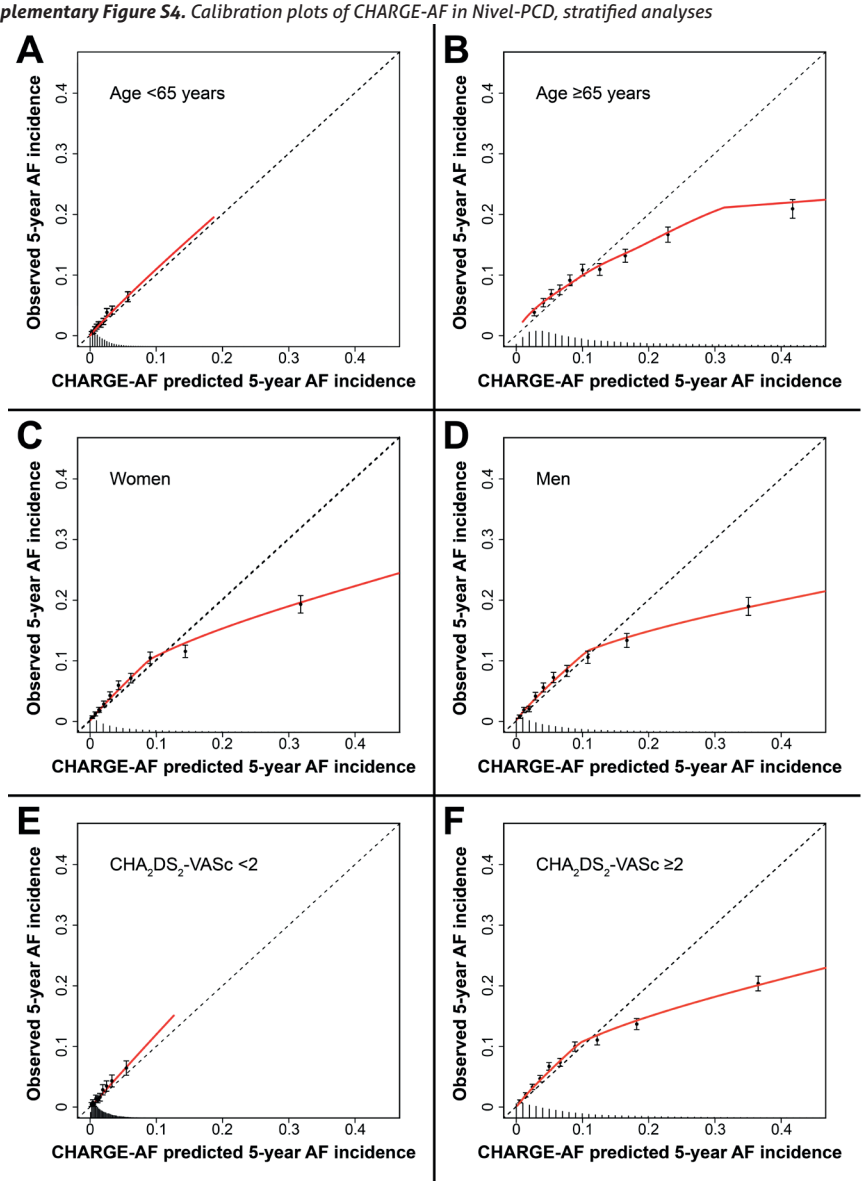
Panel A, Baseline CHARGE-AF risk distribution; Panel B, Relative contribution of CHARGE-AF risk factors to mean baseline CHARGE-AF risk score in successive strata of increased CHARGE-AF risk. Since DBP has a negative coefficient in the CHARGE-AF formula, DBP is depicted as such in this graph.

Supplementary Figure S3. *Cumulative AF incidence and number of practices included in the analysis during follow-up*



AF, atrial fibrillation; Nivel-PCD, Netherlands Institute for Health Services Research Primary Care Database. Panel A, Kaplan-Meier plot of cumulative AF incidence for all n=111,475 free of AF and complete CHARGE-AF data at baseline; Panel B, Number of Nivel-PCD practices (blue bars) and patients (red line) at risk during each Nivel-PCD extraction year.

Supplementary Figure S4. *Calibration plots of CHARGE-AF in Nivel-PCD, stratified analyses*



AF, atrial fibrillation; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age, Diabetes and previous Stroke or Transient Ischaemic Attack, Vascular disease and female Sex category; CHARGE-AF, Cohorts for Heart and Aging Research in Genomic Epidemiology model for AF.

Panel A, analysis including all aged <65 years (n = 50,947); Panel B, analysis including all aged ≥65 years (n = 60,528); Panel C, analysis including all women (n = 58,549); Panel D, analysis including all men (n = 52,926); Panel E, analysis including all CHA_2DS_2-VASc <2 (n = 88,538); Panel F, analysis including all CHA_2DS_2-VASc ≥2 (n = 88,538).

The points indicate intersects of observed and expected for each decile of baseline CHARGE-AF risk, with brackets indicating the 95% confidence intervals of observed AF probability during 5-year follow-up in each decile. The red line indicates the trend for CHARGE-AF calibration in the sample. The spikes on the x axis indicate the distribution of AF-free survivors by CHARGE-AF risk.

REFERENCES

1. Magnussen C, Niiranen TJ, Ojeda FM, Gianfagna F, Blankenberg S, Njolstad I, et al. Sex Differences and Similarities in Atrial Fibrillation Epidemiology, Risk Factors, and Mortality in Community Cohorts: Results From the BiomarCaRE Consortium (Biomarker for Cardiovascular Risk Assessment in Europe). Circulation. 2017;136(17):1588-97.
2. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Europace. 2016;18(11):1609-78.
3. Freedman B, Camm J, Calkins H, Healey JS, Rosenqvist M, Wang JG, et al. Screening for Atrial Fibrillation A Report of the AF-SCREEN International Collaboration. Circulation. 2017;135(19):1851-67.
4. Linker DT, Murphy TB, Mokdad AH. Selective screening for atrial fibrillation using multivariable risk models. Heart. 2018;104(18):1492-9.
5. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB, et al. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc. 2013;2(2):e000102;1-11.
6. Li YG, Pastori D, Farcomeni A, Yang PS, Jang E, Joung B, et al. A Simple Clinical Risk Score (C2HEST) for Predicting Incident Atrial Fibrillation in Asian Subjects: Derivation in 471,446 Chinese Subjects, With Internal Validation and External Application in 451,199 Korean Subjects. Chest. 2019;155(3):510-8.
7. Berntsson J, Smith JG, Nilsson PM, Hedblad B, Melander O, Engstrom G. Pro-atrial natriuretic peptide and prediction of atrial fibrillation and stroke: The Malmo Preventive Project. European Journal of Preventive Cardiology. 2017;24(8):788-95.
8. Kokubo Y, Watanabe M, Higashiyama A, Nakao YM, Kusano K, Miyamoto Y. Development of a Basic Risk Score for Incident Atrial Fibrillation in a Japanese General Population- The Suita Study. Circ J. 2017;81(11):1580-8.
9. Pfister R, Bragelmann J, Michels G, Wareham NJ, Luben R, Khaw KT. Performance of the CHARGE-AF risk model for incident atrial fibrillation in the EPIC Norfolk cohort. European Journal of Preventive Cardiology. 2015;22(7):932-9.
10. Svennberg E, Lindahl B, Berglund L, Eggers KM, Venge P, Zethelius B, et al. NT-proBNP is a powerful predictor for incident atrial fibrillation - Validation of a multimarker approach. Int J Cardiol. 2016;223:74-81.
11. Himmelreich JCL, Veelers L, Lucassen WAM, Schnabel RB, Rienstra M, van Weert H, et al. Prediction models for atrial fibrillation applicable in the community: a systematic review and meta-analysis. Europace. 2020;22(5):684–94.
12. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, Petersen I. Issues in multiple imputation of missing data for large general practice clinical databases. Pharmacoepidemiol Drug Saf. 2010;19(6):618-26.
13. Verberne LDM, Nielen MMJ, Leemrijse CJ, Verheij RA, Friele RD. Recording of weight in electronic health records: an observational study in general practice. BMC Fam Pract. 2018;19(1):174.
14. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162(1):55-63.

15. Lamberts H, Wood M. International Classification of Primary Care. 1st ed. Oxford: Oxford University Press; 1987.
16. Nielen MMJ, Spronk I, Davids R, Korevaar JC, Poos R, Hoeymans N, et al. Estimating Morbidity Rates Based on Routine Electronic Health Records in Primary Care: Observational Study. JMIR Med Inform. 2019;7(3):e11929.
17. Nivel. Netherlands Institute for Health Services Research 2018 [Available from: https://www.nivel.nl/en].
18. Hulme OL, Khurshid S, Weng LC, Anderson CD, Wang EY, Ashburner JM, et al. Development and Validation of a Prediction Model for Atrial Fibrillation Using Electronic Health Records. JACC Clin Electrophysiol. 2019;5(11):1331-41.
19. Vermond RA, Geelhoed B, Verweij N, Tieleman RG, Van der Harst P, Hillege HL, et al. Incidence of Atrial Fibrillation and Relationship With Cardiovascular Events, Heart Failure, and Mortality: A Community-Based Study From the Netherlands. J Am Coll Cardiol. 2015;66(9):1000-7.
20. Spronk I, Korevaar JC, Poos R, Davids R, Hilderink H, Schellevis FG, et al. Calculating incidence rates and prevalence proportions: not as simple as it seems. BMC Public Health. 2019;19(1):512.
21. Held U, Kessels A, Garcia Aymerich J, Basagana X, Ter Riet G, Moons KG, et al. Methods for Handling Missing Variables in Risk Prediction Models. Am J Epidemiol. 2016;184(7):545-51.
22. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the Performance of Prediction Models A Framework for Traditional and Novel Measures. Epidemiology. 2010;21(1):128-38.
23. D'Agostino R, Nam BH. Evaluation of the performance of survival analysis models: Discrimination and calibration measures. In: Balakrishnan N, Rao CR, editors. Handbook of statistics. 23. Amsterdam: Elsevier; 2004. p. 1-25.
24. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. Stat Med. 2015;34(10):1659-80.
25. 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(2):263-72.
26. Everett BM, Cook NR, Conen D, Chasman DI, Ridker PM, Albert CM. Novel genetic markers improve measures of atrial fibrillation risk prediction. Eur Heart J. 2013;34(29):2243-51.
27. Alonso A, Roetker NS, Soliman EZ, Chen LY, Greenland P, Heckbert SR. Prediction of Atrial Fibrillation in a Racially Diverse Cohort: The Multi-Ethnic Study of Atherosclerosis (MESA). J Am Heart Assoc. 2016;5(2):1-8.
28. Christophersen IE, Yin XY, Larson MG, Lubitz SA, Magnani JW, McManus DD, et al. A comparison of the CHARGE-AF and the CHA(2)DS(2)-VASc risk scores for prediction of atrial fibrillation in the Framingham Heart Study. Am Heart J. 2016;178:45-54.
29. Saliba W, Gronich N, Barnett-Griness O, Rennert G. Usefulness of CHADS2 and CHA2DS2-VASc Scores in the Prediction of New-Onset Atrial Fibrillation: A Population-Based Study. Am J Med. 2016;129(8):843-9.
30. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC, 2017.
31. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ .

32. Chaker L, Heeringa J, Dehghan A, Medici M, Visser WE, Baumgartner C, et al. Normal Thyroid Function and the Risk of Atrial Fibrillation: the Rotterdam Study. J Clin Endocrinol Metab. 2015;100(10):3718-24.
33. Maheshwari A, Norby FL, Soliman EZ, Koene R, Rooney M, O'Neal WT, et al. Refining Prediction of Atrial Fibrillation Risk in the General Population With Analysis of P-Wave Axis (from the Atherosclerosis Risk in Communities Study). Am J Cardiol. 2017;120(11):1980-4.
34. Sinner MF, Stepas KA, Moser CB, Krijthe BP, Aspelund T, Sotoodehnia N, et al. B-type natriuretic peptide and C-reactive protein in the prediction of atrial fibrillation risk: the CHARGE-AF Consortium of community-based cohort studies. Europace. 2014;16(10):1426-33.
35. Schnabel RB, Sullivan LM, Levy D, Pencina MJ, Massaro JM, D'Agostino RB, Sr., et al. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. Lancet. 2009;373(9665):739-45.
36. Hamada R, Muto S. Simple risk model and score for predicting of incident atrial fibrillation in Japanese. J Cardiol. 2019;73(1):65-72.
37. Schnabel RB, Aspelund T, Li G, Sullivan LM, Suchy-Dicey A, Harris TB, et al. Validation of an atrial fibrillation risk algorithm in whites and African Americans. Arch Intern Med. 2010;170(21):1909-17.
38. Pessoa-Amorim G, Casadei B, Jones NR, A'Court C, Bulbulia R, Buck G, et al. Active Monitoring for AtriaL FIbrillation (AMALFI): protocol and pilot from a mail-based randomized trial of screening for subclinical atrial fibrillation in high-risk individuals. ESC Heart & Stroke. 2020:9.
39. Barkhuysen P, de Grauw W, Akkernnans R, Donkers J, Schers H, Biermans M. Is the quality of data in an electronic medical record sufficient for assessing the quality of primary care? J Am Med Inform Assoc. 2014;21(4):692-8.
40. de Groot JA, Bossuyt PM, Reitsma JB, Rutjes AW, Dendukuri N, Janssen KJ, et al. Verification problems in diagnostic accuracy studies: consequences and solutions. BMJ. 2011;343:d4770.

4

Frequent premature atrial contractions are associated with atrial fibrillation, brain ischaemia, and mortality: a systematic review and meta-analysis

Jelle C.L. Himmelreich, Wim A.M. Lucassen, Martijn Heugen, Patrick M.M. Bossuyt, Hanno L. Tan, Ralf E. Harskamp, Faridi S. van Etten-Jamaludin, Henk C.P.M. van Weert

EP Europace, Volume 21, Issue 5,

May 2019, Pages 698–707,

https://doi.org/10.1093/europace/euy276

ABSTRACT

Aims: Premature atrial contractions (PACs) are a common cardiac phenomenon, traditionally considered to be of little clinical significance. Recent studies, however, suggest that PACs are associated with atrial fibrillation (AF), as well as ischaemic stroke, transient ischaemic attack, and mortality. This systematic review aims to investigate the association between PACs on standard electrocardiogram (ECG) as well as PAC-count on Holter monitor and future detection of AF, brain ischaemia, and all-cause mortality in patients without a history of AF.

Methods and results: We searched PubMed, Embase (OVID), and Cochrane Database of Systematic Reviews from inception through 11 April 2018 and performed a systematic review and meta-analysis. We assessed risk of bias using a modified Quality In Prognosis Studies tool. The primary expression of associations in meta-analysis was the unadjusted hazard ratio (HR) using a random effects model. We identified 33 eligible studies including 198 876 patients from Western and East Asian populations with mean age ranging 52–76 years. Frequent PACs on 24–48 h Holter was associated with AF (HR 2.96, 95% confidence interval [CI] 2.33–3.76; 15 cohorts, n = 16 613), first stroke (HR 2.54, 95% CI 1.68–3.83; 3 cohorts, n = 1468), and all-cause mortality (HR 2.14, 95% CI 1.94–2.37; 6 cohorts, n = 7571). There was insufficient evidence to conclude that presence of ≥1 PAC on standard 12-lead ECG is associated with future AF detection.

Conclusion: In older patients without a history of AF, frequent PACs on 24–48 h Holter are significantly associated with AF, first stroke, and mortality.

INTRODUCTION

Premature atrial contractions (PACs) are a common cardiac phenomenon, occurring at least once per 24 h in 99% of the general adult population.[1] These supraventricular ectopic beats have traditionally been considered to be of little clinical significance when seen on a standard electrocardiogram (ECG) or continuous ECG-monitor (Holter).[2,3] However, recent evidence suggests a positive relation between baseline PACs frequency and risk of incident atrial fibrillation (AF), ischaemic stroke or transient ischaemic attack (TIA), and mortality among older patients without known (paroxysmal) AF.[2]

Kamel et al.[4] outlined the theoretical framework and suggested that there is both a relation between PACs and AF, as well as a relation between PACs and stroke, and subsequently, mortality, beyond AF. The authors propose that AF and other atrial ECG-anomalies, among them PACs, must be seen as expressions of atrial cardiomyopathy (aCMP). While some forms of aCMP may be more likely thrombogenic—with AF as the clinically most established variant[5–7]—there are other expressions of aCMP such as PACs that may be independently related to clinical outcomes as well.[4]

A more thorough understanding of the alleged positive relation between PACs and subsequent risk of AF in older patients could be used to increase efficiency of AF screening and detection. Clinicians might consider referring patients for prolonged monitoring to detect paroxysmal AF when their 12-lead ECG or Holter shows a PAC-count over a certain clinically relevant threshold ('frequent PACs') in the absence of continuous AF.

Establishing the association between PACs and future AF, as well as stroke or TIA, and mortality may lead to a revision of PACs as a benign finding on 12-lead ECG or Holter monitor. This potential clinical relevance warrants a synthesis of the available evidence.[8] We conducted a systematic review and meta-analysis to evaluate the association between baseline PAC-count, as established on 12-lead ECG or on Holter monitor, in patients without a known history of AF, as a predictor for AF, as primary outcome of interest, and/or ischaemic stroke, TIA, or all-cause mortality, as secondary outcomes of interest.

METHODS

We reported this systematic review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[9] and within the framework of the Cochrane Prognosis Methods Group exemplar protocol for systematic reviews on prognostic factors.[10] We published the protocol before the search date at the International Prospective Register of Systematic Reviews (PROSPERO), ID CRD42017055311.[11]

Data sources and searches

We searched PubMed, Embase (OVID), and Cochrane Database of Systematic Reviews from database inception through to 11 April 2018. We included only studies written in English, Dutch, French, German, Italian, or Spanish. We used keywords PACs, AF, ischaemic stroke, TIA, and mortality (see Supplementary material, Table S1 for full search strategy).

Study selection

Two investigators (J.C.L.H. and M.H.) identified potentially eligible studies, while a third (W.A.M.L.) resolved any disagreements. We used an online systematic review platform (Covidence, Veritas Health Innovation Ltd, Melbourne, Australia). To be eligible for inclusion, studies had to be an original systematic review, randomized trial, or observational study (prospective or retrospective) and had to report in a full text article on PACs as a prognostic factor for AF (primary outcome) and/or ischaemic stroke, TIA, or mortality (secondary outcomes) in patients ≥18 years of age. We required a follow-up of at least 3 months, since we were primarily interested in the predictive value of PACs for the outcome AF, not whether PACs are indicative of previously undetected AF already present at baseline. Studies were only eligible if they had excluded patients with a known history of AF based on medical records check and/or baseline rhythm recording. We made an exception to the latter criterion if authors presented separate analyses for patients without a known history of AF, either within the original article or upon our request for additional data. For studies on ischaemic stroke or TIA, we required that these diagnoses were clearly distinguishable from, and not incorporated in a composite endpoint with, haemorrhagic stroke. We excluded studies with cohorts defined by a common history of (recent) catheter ablation for AF, coronary artery bypass graft, or percutaneous coronary intervention.

Data extraction and quality assessment

One investigator (J.C.L.H.) extracted data on study population, number of participants, exclusion criteria, follow-up duration, methods of ascertainment of both predictor

and outcome, the incidence of the studied outcome(s) within the cohort, information on the statistical model(s) applied in deriving the association between PACs and the outcome(s), and the numeric outcome of the statistical analysis. A second investigator (W.A.M.L.) independently reviewed these data for accuracy. Three investigators (J.C.L.H., W.A.M.L., and M.H.) assessed the risk of bias of included studies with a modified Quality In Prognosis Studies (QUIPS) tool for prognosis studies[12] on a consensus basis (see Supplementary material, Methods for modifications made to the interpretation of the QUIPS tool). Risk of bias for each of the six QUIPS domains, as well as overall risk of bias (low, moderate, or high) was assigned on a consensus basis as well (see Table 1).

Data synthesis and analysis

The primary expression of associations in meta-analysis was the unadjusted hazard ratio (HR): between PACs, either on 12-lead ECG as a dichotomized variable or on Holter as a dichotomized, ordinal, or continuous variable, and the outcomes AF, TIA, ischaemic stroke, or mortality, respectively. We used a random effects inverse variance model for meta-analysis of log HRs, enabling us to present the summary unadjusted HRs. Analyses were performed in Review Manager (RevMan version 5.3, The Cochrane Collaboration). We evaluated statistical significance in all analyses at the 0.05 level. In each analysis, we calculated the mean as the summary effect measure, its 95% confidence interval (CI), and the I2 statistic as an expression of the heterogeneity between studies.[10,13] A minimum of three studies is required for a reliable assessment of the overall effect and CI in random effects meta-analysis.[14] We, therefore, conducted meta-analysis for each of the outcomes of interest only when three or more studies of low or moderate overall risk of bias reported unadjusted HRs on a similar statistical approach to PAC-count as the predictor, i.e. PAC-count as a dichotomous, ordinal, or continuous variable on Holter, or as a dichotomous variable on ECG. In case of high statistical heterogeneity—defined as I2 >30%[15]—we provided a 95% prediction interval (PI) in order to allow for a better interpretation of the results of the random effects meta-analysis.[14] We derived the PI using the methods described by IntHout et al.[16] A significant 95% PI led us to uphold the conclusion that there is a significant association between predictor and outcome. A non-significant 95% PI, despite a significant 95% CI, led us to conclude that there is still insufficient evidence to suggest an association in the particular comparison.

We first selected for meta-analysis those studies with the most commonly used definition of PAC-count as a predictor within our sample, i.e. with the most similar cut-off for dichotomization, the most similar scale (e.g. linear or log) applied in PAC-count as a continuous variable, or the most similar categorization in PAC-count as an ordinal

***Table 1**. Risk of bias assessment for the six domains of the modified QUIPS tool and overall risk of bias assessment based on predefined criteria for all included studies*

Study	QUIPS Domain						Overall Risk of Bias
	1: Study Participation	2: Study Attrition	3: Prognostic Factor Measure-ment	4: Outcome Measure-ment for All Outcomes	5: Study Confounding	6: Statistical Analysis and Reporting	
Acharya 2015	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Binici 2010	LOW	LOW	LOW	MOD	LOW	LOW	MOD
Blanch Gracia 2013	LOW	MOD	MOD	MOD	LOW	MOD	HIGH
Cabrera 2016	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Chong 2012	LOW	LOW	LOW	MOD	MOD	LOW	MOD
Chun 2016	HIGH	LOW	LOW	MOD	MOD	LOW	HIGH
Dewland 2013	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Engstrom 2000	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Folkeringa 2006	LOW	MOD	LOW	MOD	MOD	*HIGH*	HIGH
Gladstone 2015	LOW	MOD	LOW	MOD	LOW	LOW	MOD
Inohara 2013	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Johnson 2015	LOW	LOW	LOW	MOD	LOW	LOW	MOD
Kochhauser 2014	LOW	LOW	LOW	LOW	LOW	MOD	MOD
Lin 2015	LOW	LOW	LOW	MOD	LOW	LOW	MOD
Marinheiro 2017	MOD	LOW	LOW	MOD	MOD	LOW	MOD
Murakoshi 2015	LOW	MOD	LOW	LOW	LOW	LOW	MOD
Nguyen 2017	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Nortamo 2017	LOW	LOW	LOW	MOD	LOW	LOW	MOD
O'Neal 2016	LOW	MOD	LOW	LOW	LOW	LOW	MOD
O'Neal 2017	LOW	MOD	LOW	LOW	LOW	LOW	MOD
Perez 2009	LOW	LOW	LOW	MOD	MOD	LOW	MOD
Pinho 2015	LOW	MOD	LOW	LOW	LOW	LOW	MOD
Qureshi 2014	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Raman 2017	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Suzuki 2013	LOW	MOD	LOW	MOD	LOW	LOW	MOD
Thijs 2016	LOW	LOW	LOW	LOW	LOW	LOW	LOW
Vinther 2016	LOW	LOW	LOW	MOD	MOD	LOW	MOD
Vinther 2017	LOW	LOW	LOW	MOD	MOD	LOW	MOD
Wallmann 2003	MOD	HIGH	LOW	HIGH	LOW	MOD	HIGH
Wallmann 2007	MOD	MOD	LOW	LOW	LOW	LOW	MOD
Weber-Krüger 2017	LOW	LOW	LOW	LOW	MOD	LOW	MOD
Yamada 2000	LOW	MOD	LOW	LOW	LOW	LOW	MOD
Yodogawa 2013	MOD	MOD	LOW	LOW	LOW	LOW	MOD

LOW, low risk of bias; HIGH, high risk of bias; MOD, moderate risk of bias; QUIPS, Quality In Prognosis Studies.

variable. When both the 95% CI and 95% PI in this homogeneous sample showed a significant association in a particular comparison, we proceeded to perform an overall meta-analysis that included all eligible studies for the respective comparison. We chose this approach because the statistical definition of PAC-count as the predictor has previously been suggested to play a major role in PAC-count's ability to accurately predict the outcome.[17] When both the homogeneous and the more heterogeneous overall meta-analysis found a significant association as determined by both 95% CI and 95% PI, we reported the result of the overall meta-analysis as the final result for that particular comparison between PAC-count at a certain statistical approach (dichotomous, ordinal, or continuous) and the studied outcome. When the analysis of the more heterogeneous sample resulted in a non-significant 95% CI and/or 95% PI, we reported the outcomes of the more homogeneous meta-analysis as the final results for that particular comparison, generalizable only to the particular statistical definition of the predictor applied in the homogeneous sample.

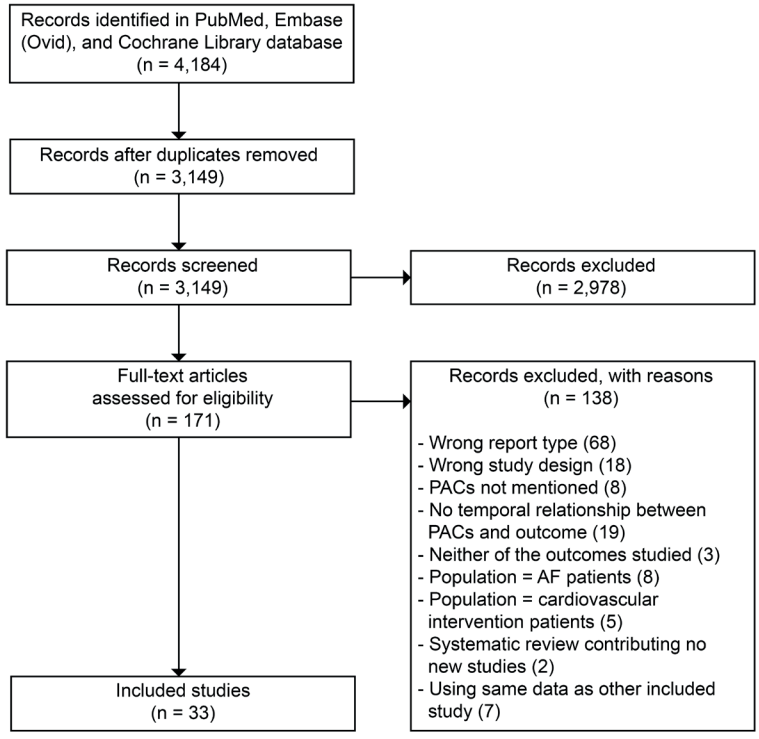
We performed further subgroup analyses according to overall risk of bias (categories: 'low'; 'moderate'), population (categories: 'general population, not necessarily assigned to baseline ECG, or Holter for cardiac symptoms'; 'general population, assigned to baseline ECG, or Holter for cardiac symptoms'; 'post-stroke population'), and mean follow-up duration (categories: '<5 years'; '≥5 years') to see if these subgroups showed different results compared with the findings among their respective overall analyses.

When a study presented HRs for both a singular PACs-based cut-off as well as a runs-of-PACs-based cut-off, we incorporated into meta-analysis only the data on the singular PACs-based cut-off, since we were primarily interested in the role of singular PAC-count as a predictor. When a study presented only HRs for subgroups within the cohort instead of a composite HR for the entire cohort, we performed a random effects inverse variance meta-analysis to calculate the summary HR and 95% CI for that study's entire cohort. In further analyses, we used this overall HR and 95% CI as representative of that study. Since not all studies used the same base of the logarithm for the log transformed continuous baseline PAC-count we adjusted those coefficients where necessary to the most commonly used base of 10 and their CIs accordingly. We created funnel plots of all meta-analyses that contained 10 or more studies to assess the risk of reporting bias.[10]

RESULTS

The search identified 3149 unique studies of which we assessed 171 reports in full-text. Among the 42 eligible studies, we excluded two systematic reviews that added no new studies[18,19] and seven studies for reporting on similar outcomes based on the same database as other included studies.[20–26] We included four studies that used data from two databases, since these studies either reported on different outcomes[27,28] or used different recording devices for baseline PAC-count assessment.[29,30] We eventually included 33 studies representing 32 databases for data synthesis (Figure 1).

Figure 1. *Literature search flow diagram.*



AF, atrial fibrillation; PACs, premature atrial contractions.

The characteristics of included studies and their results on the association between PACs and the studied outcome(s) are listed in Supplementary material, Tables S2 and S3, respectively. Studies represented out- and inpatient as well as community-based populations from North America, Europe, and East Asia. The number of participants varied from 68 to 42 751, with a total of 198 876. Average age ranged from 52 up to 76 years of age at baseline. Follow-up time ranged from 6 months up to 13 years.

Among eight studies that reported on 12-lead ECG as the baseline recording device ('ECG studies') seven reported on standard 10–15 s ECG[27,28,30–34] and one reported on a 2 mins ECG strip.[35] We included 25 studies that reported on continuous ECG monitoring as the baseline recording device ('Holter studies'); 21 on 24-h Holter,[17,29,36–54] two on 48-h Holter,[55,56] one on exercise test continuous ECG,[57] and one on polysomnography continuous ECG.[58] All ECG studies used the dichotomization 'one or more PACs' vs. 'no PACs' for the association with the studied outcome.[27,28,30–35] Among the 25 Holter studies, a number of studies reported on similar cut-offs for dichotomization, whether by coincidence (e.g. '≈100 singular PACs/24 h'[36,38,46,52]) or by design [e.g. cut-off = excessive supraventricular ectopic activity (ESVEA) as defined by Binici et al.[41,55]]. Some Holter studies based their cut-off on a percentile of PAC-count (e.g. cut-off = lower bound of upper quartile[29,38,39,42,44,46–49,58]) or on the derived optimum for outcome prediction within the cohort.[36,43] Other Holter studies provided no substantiation for the chosen cut-off value.[37,51,52] Most studies reported HR as the primary measure of association.[27,29,30,32–41,43–48,51–56,58] Others reported only relative risk,[42] odds ratio,[28,49,50] logistic regression,[17,42] or crude incidences;[31,57] we were unable to calculate HRs for these cohorts and could not include these studies in meta-analysis.

Association between premature atrial contractions and atrial fibrillation

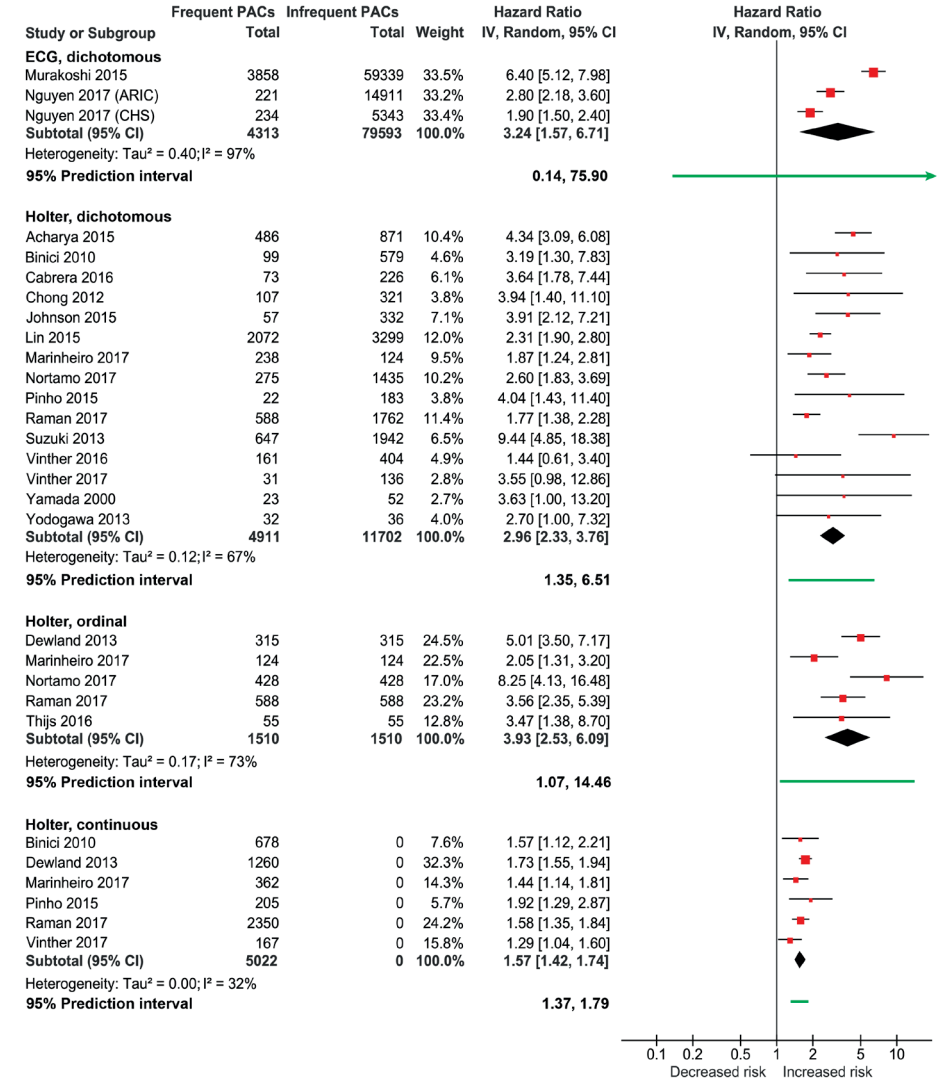
Association between dichotomized premature atrial contraction-count on electrocardiogram and atrial fibrillation

Two studies out of eight ECG studies reported unadjusted HRs on the relation between presence of ≥1 PAC vs. no PACs on baseline 12-lead ECG and AF. Since these two studies represented the results of three separate cohorts, we were able to perform meta-analysis. The studies differed in their methods in using either a 10-s 12-lead ECG[30] or a 15-s 12-lead ECG[33] as the baseline measuring device. Overall meta-analysis of the studies resulted in a statistically significant unadjusted summary HR 3.24; 95% CI 1.57–6.71. Because of high statistical heterogeneity (I2 = 97%), we additionally calculated the 95% PI 0.14–75.90 (Figure 2). We concluded that there is insufficient evidence for an association between presence of ≥1 PAC on ECG and future AF detection.

Association between dichotomized premature atrial contraction-count on Holter and atrial fibrillation

Of the 25 included Holter studies, 15 reported unadjusted HRs on dichotomized baseline PAC-count on Holter for AF.[36–38,41,43–46,48,51–53,55,56,58] Meta-analysis of only those four studies with the most commonly used cut-off for dichotomization (≈100 PACs/24 h) resulted in an unadjusted summary HR 4.86; 95% CI 3.02–7.82.[36,38,46,52] Due to high

Figure 2. *Meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF, grouped according to their recording device and respective statistical approach to PAC-count as a predictor for the outcome.*



ARIC, Atherosclerosis Risk In Communities; CHS, Cardiovascular Health Study; CI, confidence interval; IV, inverse variance; PACs, premature atrial contractions. Totals under 'frequent PACs' and 'infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study. Totals under 'frequent PACs' in subgroup 'Holter, continuous' represent total cohort size, since no dichotomization, or ordinal comparison was applied in these studies.

heterogeneity (I2 = 48%), we calculated the 95% PI 1.31–17.97 (Supplementary material, Figure S1). Subsequent overall meta-analysis of all 15 studies that reported unadjusted HRs on dichotomized baseline PAC-count on Holter for AF with any cut-off for dichotomization, remained statistically significant with unadjusted summary HR

of 2.96; 95% CI 2.33–3.76; I2 67%, and 95% PI 1.35–6.51 for having 'frequent PACs' at baseline (Figure 2 and Supplementary material, Figure S1). A funnel plot for the 15 Holter studies that reported unadjusted HRs for AF based on dichotomized PAC-count did not indicate reporting bias (Supplementary material, Figure S2).

Association between premature atrial contraction-count as an ordinal variable on Holter and atrial fibrillation

Five out of 25 included Holter studies reported unadjusted HRs on ordinal baseline PAC-count on Holter for AF.[29,44,47,53,58] Four studies categorized baseline PAC-count as quartiles,[29,44,47,58] resulting in both a significant unadjusted summary HR of 4.68; 95% CI 3.35–6.54; I2 37% and 95% PI 2.03–10.80 (Supplementary material, Figure S6). An overall analysis of all five studies that applied any categorization of baseline PAC-count as an ordinal variable to predict AF again resulted in an unadjusted summary HR of 3.93; 95% CI 2.53–6.09; I2 73% and 95% PI of 1.07–14.46 (Figure 2 and Supplementary material, Figure S6).

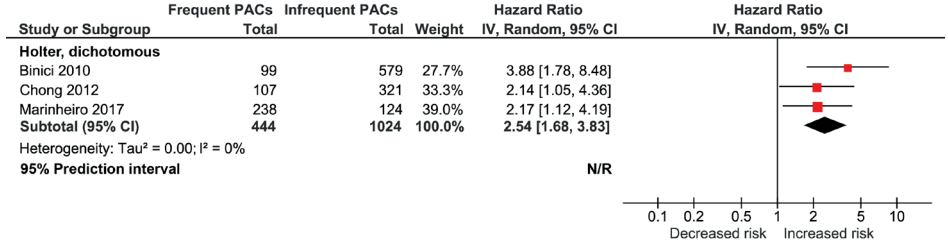
Association between premature atrial contraction-count as a continuous variable on Holter and atrial fibrillation

Six out of 25 included Holter studies reported unadjusted HRs on the relationship between continuous PAC-count on Holter and AF.[29,45,48,53,55,58] Five studies applied a base-10 log-transformed scale to baseline PAC-count.[29,45,48,53,58] Meta-analysis of these five studies resulted in unadjusted summary HR of 1.57; 95% CI 1.39–1.76; I2 45% and 95% PI 1.14–2.17 (Supplementary material, Figure S9). In overall meta-analysis of all six studies that presented unadjusted HRs for continuous PAC-count on any scale for AF, both the unadjusted summary HR and 95% PI remained significant, at 1.57; 95% CI 1.42–1.74; I2 32% and 1.37–1.79, respectively (Figure 2 and Supplementary material, Figure S9).

Association between premature atrial contractions and ischaemic stroke and/or transient ischaemic attack

Three out of 25 Holter studies reported unadjusted HRs on dichotomized Holter data for the outcome first stroke.[38,53,55] Meta-analysis of the three studies resulted in a summary unadjusted HR of 2.54; 95% CI 1.68–3.83; I2 0% (Figure 3). None of the eight ECG studies reported unadjusted HRs on the outcome ischaemic stroke and/or TIA. Furthermore, we were unable to meta-analyse Holter data with PAC-count as an ordinal or continuous variable for ischaemic stroke and/or TIA, since no three or more studies presented unadjusted HRs for these associations.

Figure 3. *Meta-analysis of studies that reported unadjusted hazard ratios for the outcome first stroke based on dichotomized Holter data.*



CI, confidence interval; IV, inverse variance; N/R, not relevant; PACs, premature atrial contractions. Totals under 'frequent PACs' and 'infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Association between premature atrial contractions and all-cause mortality

Association between dichotomized premature atrial contraction-count on ECG and all-cause mortality

We were unable to perform meta-analysis of ECG data for all-cause mortality since only two out of eight ECG studies reported unadjusted HRs on this association.[32,33]

Association between dichotomized premature atrial contraction-count on Holter and all-cause mortality

Among the 25 Holter studies, six studies reported unadjusted HRs on the relationship between dichotomized PAC-count and all-cause mortality.[38,43,48,53,55,56] All used different cut-offs in their respective definitions of dichotomized PAC-count as the predictor (Supplementary material, Table S3). Overall meta-analysis of the six dichotomized Holter studies for all-cause mortality resulted in a summary unadjusted HR of 2.14; 95% CI 1.94–2.37; I2 0% (Figure 4).

Association between premature atrial contraction-count as an ordinal variable on Holter and all-cause mortality

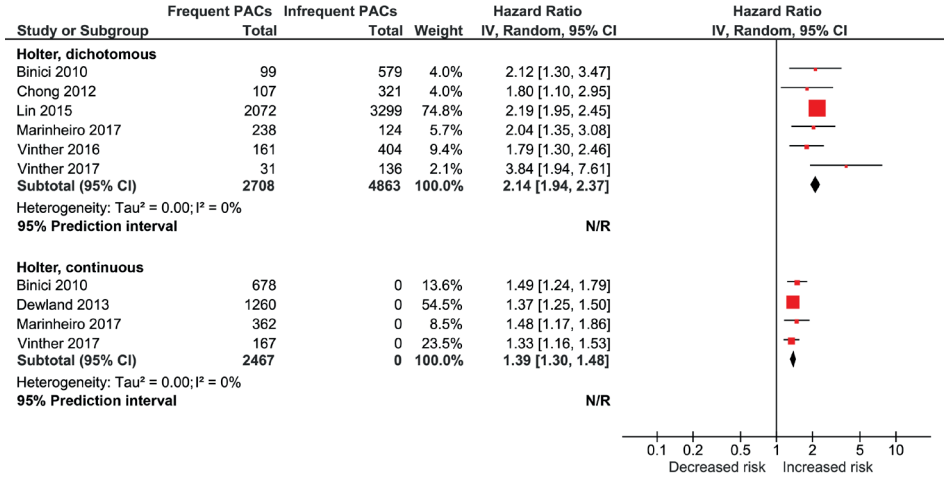
We were unable to meta-analyse ordinal Holter data for all-cause mortality, since only two out of 25 Holter studies presented unadjusted HRs for this association.[29,53]

Association between premature atrial contraction-count as a continuous variable on Holter and all-cause mortality

Four out of 25 Holter studies reported unadjusted HRs on continuous PAC-count for all-cause mortality.[29,48,53,55] Three studies applied a base-10 log-scale to baseline PAC-count,[29,48,53] resulting in a summary unadjusted HR of 1.37; 95% CI 1.28–1.48; I2 0% (Supplementary material, Figure S15). Overall meta-analysis with one other study that

applied a linear scale to baseline PAC-count55 resulted in a summary unadjusted HR of 1.39; 95% CI 1.30–1.48; I2 0% (Figure 4 and Supplementary material, Figure S15).

Figure 4. *Meta-analysis of studies that reported unadjusted hazard ratios for the outcome all-cause mortality, grouped according to their respective statistical approach to PAC-count as a predictor for the outcome.*



CI, confidence interval; IV, inverse variance; N/R, not relevant; PACs, premature atrial contractions. Totals under 'frequent PACs' and 'infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study. Totals under 'frequent PACs' in subgroup 'Holter, continuous' represent total cohort size, since no dichotomization, or ordinal comparison was applied in these studies.

Explorative subgroup analyses

For results of as well as a discussion on each subgroup meta-analysis, we refer to the Supplementary material, Results and Figures.

DISCUSSION

This systematic review with meta-analysis shows a significant association between finding 'frequent PACs' on 24–48 h Holter and future AF detection, first stroke, as well as all-cause mortality in older patients without a known history of AF, where baseline PAC-count on Holter was dichotomized by any cut-off. Moreover, there was a significant association between increasing PAC-count on 24–48 h Holter and future AF detection as well as all-cause mortality in older patients without a known history of AF, where baseline PAC-count was considered on a continuous log-transformed or linear scale. Although there was a strong trend among included studies, there was insufficient evidence to conclude that the presence of ≥1 PACs on ECG is associated with future AF detection.

Strengths and limitations

This systematic review and meta-analysis has a number of strengths. First, our review includes only studies that had made efforts to exclude patients with known AF by performing at least a medical history check and rhythm recording (ECG or Holter) at baseline. This is important as AF patients are known to have higher PAC-count during episodes of sinus rhythm.[59,60] As such, failure to exclude patients with a history of AF would likely have led to an overestimation of the association between baseline PAC-count and incidence of the studied outcomes. Second is our inclusion of studies from various ethnic backgrounds as well as various high- and low-risk populations (i.e. primary care, post-stroke). Third is the presentation of our data, in particular the distinction we apply between ECG and Holter studies, as well as between PACs as a dichotomized, continuous, and ordinal variable in Holter studies.

The primary limitation of our systematic review was the considerable heterogeneity in statistical definitions of the predictor, especially the differences in cut-offs used for dichotomization. We made efforts to account for this limitation by not only calculating the 95% PI in case of high statistical heterogeneity, but also applying a stepwise approach in our selection of studies for meta-analysis. Here, we first selected studies based on the most commonly applied statistical definition of PAC-count as a predictor, followed by an overall analysis if analysis of the homogeneous sample showed a definitive association. The funnel plot for the meta-analysis of dichotomized Holter data for AF showed that reporting bias is likely not a large source of bias in this analysis. The language restriction within our search, which we applied for practical as well as quality-related reasons, has been shown not to lead to significant bias.[61]

Unadjusted hazard ratio as the primary expression of associations

Since our primary aim was to assess the association between baseline PACs and the outcomes of interest, we chose the unadjusted HR as the primary unit of analysis for this systematic review and meta-analysis. Analysis of the unadjusted HRs serves to explore if there is any association between PACs and the studied outcome, while meta-analysis of adjusted models on these associations could later serve to explore to what extent PACs played an independent role within that association (if any). The considerable heterogeneity in variables used within multivariable models composed an extra argument not to present adjusted HRs in the main text.

From relative hazard to absolute risk

In this study, we used the relative measure 'unadjusted HR'. From a clinical perspective the magnitude of this effect can only be translated to an absolute risk when accompanied by knowledge on the baseline risk for a given patient. We attempted to

provide insight into this baseline hazard by displaying the incidences of the studied outcomes (AF, ischaemic stroke and/or TIA, and all-cause mortality) among 'infrequent PACs' and 'frequent PACs' patients in each study (see Supplementary material, Table S3). The difference between the two incidences provides an estimate of the absolute risk difference for the studied outcome in that population. The clinician and the patient could use this information to decide whether, e.g. more stringent (periodic) rhythm monitoring could be a preferred strategy upon detecting 'frequent PACs' in absence of AF.

When looking, for example, at the incidences in the study by Binici et al.,[55] a population-based cohort of patients >55 years with participants randomly assigned to baseline Holter monitoring, the incidences of AF detection during the 6.3-year follow-up were 2.6% or 4.3/1000 person-years among the 'infrequent PACs' group and 7.1% or 12.8/1000 person-years among the 'frequent PACs' group (Supplementary material, Table S3). Knowledge on a patient's 'PACs status' (here: ≥30 PACs/h or any runs of ≥20 PACs on baseline 48-h monitor for 'frequent PACs') in a comparable cohort could thus provide a clinician with valuable information, i.e. whether or not his patient has an 8.5/1000 person-year higher risk of future AF detection.

Previous work

Two recent systematic reviews,[18,19] one of which also included a meta-analysis,[18] provided a synthesis on the composite outcome stroke and death, and recurrent stroke, respectively. While differing in inclusion criteria and methods of data synthesis, both studies reached the same conclusions: baseline PAC-count is associated with an increased risk of the studied outcomes. Our study adds to these reviews by being the first to provide a combined overview of four major cardiovascular outcomes with which PACs are associated, including AF.

Implications for clinical practice and future research

This systematic review and meta-analysis may have implications for clinicians who face the question how to interpret the not uncommon finding of frequent PACs on Holter in the absence of current AF. The data firmly indicate that the notion of frequent PACs as an innocent finding on Holter must be revised. Therefore, these results warrant more research into which patients should be referred for more stringent evaluation for AF upon detecting frequent PACs on Holter. A next step would be to perform an individual patient data (IPD) meta-analysis in order to optimally adjust for confounders, as well as to estimate an optimal cut-off for dichotomization for each of the outcomes. An IPD meta-analysis is also required to research whether the association between PACs and the outcomes, as well as the relative and absolute risks, may be different

between patient groups depending on fundamental patient characteristics such as age and gender. We, therefore, urge authors to cooperate in future IPD meta-analyses on the association between PACs and the outcomes studies in this systematic review and meta-analysis.

We emphasize here that evidence on cost-effectiveness of rhythm evaluation after detection of frequent PACs on Holter in the absence of actual AF is still lacking. We further emphasize that the findings from this systematic review and meta-analysis on the outcomes stroke and all-cause mortality are not sufficient to lead to the recommendation that physicians should start any treatment in those patients with frequent PACs independent of whether AF has been detected. However, the findings of this study do warrant further prospective clinical studies into the predictive value of finding 'frequent PACs' on Holter for the outcomes studied in this systematic review and meta-analysis.

CONCLUSION

This systematic review with meta-analysis shows a significant association between frequent PACs on Holter and the onset of AF, brain ischaemia, and mortality in older patients without a history of AF. These outcomes indicate that the notion of frequent PACs as an innocent finding on Holter must be revised. The findings of this study warrant an IPD meta-analysis in order to optimally adjust for confounders and to estimate optimal cut-offs for each outcome in different populations, as well as further prospective clinical studies into the predictive value of finding 'frequent PACs' on Holter for the outcomes studied in this work.

Acknowledgements: We thank Dr Linda Johnson from Lund University, Sweden, Dr Li-Yung Lui from The Osteoporotic Fractures in Men Study, United States, Dr Rita Marinheiro and Dr Leonor Parreira from Centro Hospitalar de Setubal, Portugal, Dr João Pinho from Hospital de Braga, Portugal, and Dr Kenji Yodogawa from Nippon Medical School, Japan, for providing us with additional data. We thank Dr Wim Busschers and Dr Joris de Groot from Academic Medical Center Amsterdam, The Netherlands, for statistical advice, and advice in interpretation of individual study results, respectively.

Funding: This work was supported by the Netherlands Organisation for Health Research and Development (ZonMw) [80-83910-98-13046]. The authors had full autonomy in design, conduct, and reporting of the article.

Conflict of interest: None declared.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY METHODS

Modified QUIPS-tool: modifications made to interpretation of the QUIPS domains

For risk of bias assessment, we modified the Quality In Prognosis Studies (QUIPS) tool for assessing risk of bias in prognosis studies to fit the purposes of our systematic review and meta-analysis. The most important modification was our choice not to downgrade when the study did not explicitly define the predictor (Domain 3; Prognostic Factor Measurement) or the outcome AF (Domain 4; Outcome Measurement). We did so because we considered premature atrial contractions (PACs) and atrial fibrillation (AF) to be common electrocardiographic (ECG) findings that any skilled ECG analyst should be able to interpret similarly. Consequently, failure by a study to explicitly define what they considered a PAC or AF on an ECG or Holter would not lead to downgrading of this Domain, as long as the authors defined their statistical definition of PACs as a predictor for the studied outcome.

Furthermore, we decided to assess Domain 4 for outcome AF as having moderate bias when no systematic follow-up monitoring was performed. This decision was based on the assumption that patients with higher baseline PAC-count may be more likely to have cardiac symptoms more frequently. Patients with more frequent cardiac symptoms will likely receive more frequent follow-up monitoring. A higher baseline PAC-count may therefore have led to an increased likelihood of AF detection within studies where patients received follow-up monitoring at the physician's discretion, as opposed to a regular, predefined scheme of follow-up ECGs for all patients regardless of cardiac symptoms. The means of outcome ascertainment of the included studies are depicted in Supplementary material, Table S2.

Overall risk of bias assessment as applied in this systematic review and meta-analysis

As recommended by the QUIPS authors we defined two domains as most important for this purpose: Domain 4 (Outcome Measurement) and Domain 6 (Statistical Analysis and Reporting). In accordance with Hayden et al. we defined low overall risk of bias as having low risk of bias on all six domains. We defined moderate risk of bias as having moderate risk of bias on any of domains 1-6, except when both domains 4 and 6 had moderate risk of bias. We defined high overall risk of bias as having one or more domains as high risk of bias, or having both domains 4 and 6 as moderate risk of bias.

SUPPLEMENTARY TABLES

Supplementary Table S1. Search strategy

PubMed

Filters activated: English, Dutch, French, German, Italian, Spanish
((((("Atrial Premature Complexes"[Mesh] OR premature atrial[tiab] OR atrial premature[tiab] OR premature supraventricular[tiab] OR supraventricular premature[tiab] OR atrial ectop*[tiab] OR ectopic atrial[tiab] OR ectopic supraventricular[tiab] OR supraventricular ectop*[tiab] OR atrial extrasystole*[tiab] OR supraventricular extrasystole*[tiab])) **AND** ("Atrial Fibrillation"[Mesh] OR "Stroke"[Mesh] OR "Thromboembolism"[Mesh] OR "Embolism"[Mesh] OR "Brain Ischemia"[Mesh] OR "Ischemic Attack, Transient"[Mesh] OR "Mortality"[Mesh] OR "Death"[Mesh] OR "mortality" [Subheading] OR atrial fibrillation*[tiab] OR stroke*[tiab] OR thromboemboli*[tiab] OR thrombo-emboli*[tiab] OR emboli*[tiab] OR CVA[tiab] OR CVAs[tiab] OR cerebrovascular accident*[tiab] OR transient Ischemic attack*[tiab] OR transient ischaemic attack*[tiab] OR brain ischemi*[tiab] OR brain ischaemi*[tiab] OR atrial fibrillat*[tiab] OR TIA[tiab] OR TIAs[tiab] OR reversible ischemic neurological deficit*[tiab] OR reversible ischaemic neurological deficit*[tiab] OR reversible ischemic neurologic deficit*[tiab] OR reversible ischaemic neurologic deficit*[tiab] OR mortalit*[tiab] OR death*[tiab])))) **NOT** (("Animals"[Mesh] NOT "Humans"[Mesh]) NOT ("Editorial" [Publication Type] OR "Letter" [Publication Type] OR "News" [Publication Type] OR "Comment" [Publication Type] OR "Case Reports" [Publication Type] OR letter*[ti] OR comment*[ti] OR abstracts[ti]))

Embase

Database(s):**Embase Classic+Embase**

Search Strategy:

#	Searches
1	supraventricular premature beat/ or (premature atrial or atrial premature or premature supraventricular or supraventricular premature or atrial ectop* or ectopic atrial or ectopic supraventricular or supraventricular ectop* or atrial extrasystole* or supraventricular extrasystole*).ti,ab,kw.
2	exp atrial fibrillation/ or exp cerebrovascular accident/ or exp thromboembolism/ or exp brain ischemia/ or transient ischemic attack/ or exp mortality/ or exp death/ or mo.fs. or (atrial fibrillat* or stroke* or thromboemboli* or thrombo-emboli* or emboli* or CVA or CVAs or cerebrovascular accident* or cerebro-vascular accident* or transient isch?emic attack* or brain isch?emi* or TIA or TIAs or reversible isch?emic neurological deficit* or reversible isch?emic neurologic deficit* or mortalit* or death*).ti,ab,kw.
3	animal/ not human/ not (editorial/ or letter/ or literature/ or case report/ or (letter* or comment* or abstracts).ti.)
4	(1 and 2) not 3
5	limit 4 to (dutch or english or french or german or italian or spanish)

Cochrane Library

ID	Search
#1	MeSH descriptor: [Atrial Premature Complexes] explode all trees
#2	premature atrial or atrial premature or premature supraventricular or supraventricular premature or atrial ectop* or ectopic atrial or ectopic supraventricular or supraventricular ectop* or atrial extrasystole* or supraventricular extrasystole*:ti,ab,kw (Word variations have been searched)
#3	#1 or #2
#4	MeSH descriptor: [Atrial Fibrillation] explode all trees

#5	MeSH descriptor: [Stroke] explode all trees
#6	MeSH descriptor: [Thromboembolism] explode all trees
#7	MeSH descriptor: [Embolism] explode all trees
#8	MeSH descriptor: [Brain Ischemia] explode all trees
#9	MeSH descriptor: [Ischemic Attack, Transient] explode all trees
#10	MeSH descriptor: [Mortality] explode all trees
#11	MeSH descriptor: [Death] explode all trees
#12	atrial fibrillat* or stroke* or thromboemboli* or thrombo-emboli* or emboli* or CVA or CVAs or cerebrovascular accident* or cerebro-vascular accident* or transient ischemic attack* or transient ischaemic attack* or brain ischemi* or brain ischaemi* or TIA or TIAs or reversible ischemic neurological deficit* or reversible ischaemic neurological deficit* or reversible ischemic neurologic deficit* or reversible ischaemic neurologic deficit* or mortalit* or death*:ti,ab,kw (Word variations have been searched)
#13	#4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12
#14	#3 and #13 in Cochrane Reviews (Reviews and Protocols), Other Reviews and Trials

Supplementary Table S2. Characteristics of the included studies.

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Acharya 2015	US	1,357	Veterans who underwent 24h Holter for any indication between 2000 and 2010. - Mean age total cohort: 64.0 years - Mean age 'frequent PACs' group: 71.4 ± 11.8 years - Mean age 'infrequent PACs' group: 60.3 ± 13.4 years - Men: 92.9%	R	7.5	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Binici 2010	DK	678	Random subset of men aged 55 years and all men and women aged 60, 65, 70, and 75 years from general population included between 1998 and 1999. - Mean age total cohort: 64.5 ± 6.8 years - Mean age 'frequent PACs' group: 67.6 ± 6.3 years - Mean age 'infrequent PACs' group: 63.9 ± 6.7 years - Men: 58.6%	P	6.3	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Blanch Gracia 2013	ES	183	Hypertensive outpatients who underwent ECG for any indication between 2001 and 2008. - Mean age total cohort: 63.5 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 40.6%	R	4.0	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Cabrera 2016	ES	299	Consecutive patients referred from primary care physicians or the cardiology department to investigate symptoms, ECG abnormalities or structural heart disease between March 2011 and October 2011. - Mean age total cohort: 62.5 ± 17.9 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 53.5%	R	3.3	Regular follow-up ECGs for all patients (mean 6.72 ± 3.51 ECGs per patient)

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Chong 2012	HK	428	Consecutive patients undergoing 24h Holter for palpitations, dizziness, or syncope, between 2002 and 2003. - Mean age total cohort: 66 ± 10.2 years - Mean age 'frequent PACs' group: 71.3 ± 9.8 years - Mean age 'infrequent PACs' group: 65.1 ± 9.9 years - Men: 43.7%	P	6.1	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Chun 2016	KR	684	Consecutive general hospital patients who received ≥2 instances of 24h Holter monitoring and who qualified as having 'frequent' (>100/24h) PACs on the first 24h Holter, between 1999 and 2008. - Mean age total cohort: 61.8 ± 15.0 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 49.0%	R	4.9	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Dewland 2013	US	1,260	Sample from a community-based cohort of Medicare beneficiaries ≥65 years, recruited between 1989 and 1990, with an oversampling of black residents. - Mean age total cohort: 71 (range 68-75) years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 45.2%	P	13.0	Annual study ECG; AF and mortality diagnoses from reviewing medical records and death certificates
Engstrom 2000	SE	388	Population-based cohort "Men Born in 1914" with baseline examination at 68 years of age, followed up until death or December 1996. - Mean age total cohort: 68 years - Mean age 'frequent PACs' group: 68 years - Mean age 'infrequent PACs' group: 68 years - Men: 100.0%.	P	11.1	Mortality data obtained from national register

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Folkeringa 2006	NL	258	Patients with left ventricular hypertrophy (interventricular septal diameter > 11 mm) and ejection fraction > 55% who underwent an exercise test between 1994 – 2004. - Mean age total cohort: 65 ± 11 years - Mean age 'frequent PACs' group: NA - Mean age 'infrequent PACs' group: NA - Men: 60%	R	3.7	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Gladstone 2015	CA	237	Consecutive patients aged ≥55 years admitted for acute cryptogenic stroke or transient ischemic attack enrolled in the intervention arm of the EMBRACE trial. - Mean age total cohort: 72.2 ± 8.6 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 53.6%	P	2.0	Automated event recorder ≤30 days upon study entry; follow-up ECG at physician's discretion thereafter
Inohara 2013	JP	7,692	Healthy participants from randomly selected districts in Japan followed between 1990 and 2005 - Mean age total cohort: 52.5 ± 13.7 years - Mean age 'frequent PACs' group: 67.3 years - Mean age 'infrequent PACs' group: 52.4years - Men: 41.5%	P	14.0	Mortality ascertained from national register
Johnson 2015	SE	383	Randomly selected individuals from a community-based cohort, examined between 1991 and 1996, AF history excluded from analysis randomly selected for 24-hour ECG monitoring - Mean age total cohort: 64.6 ± 5.9 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 44.9%	P	10.3	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Kochhauser 2014	DE	70	Consecutive patients admitted for acute cryptogenic stroke and subsequently implanted with ILR between 2010 and 2012. - Mean age total cohort: 58.8 ± 13.4 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 61.4%	P	1.5	All patients received ILR on admission
Lin 2015	TW	5,371	Consecutive patients who underwent clinically indicated 24-hour Holter monitoring between 2002 and 2004. Indications included palpitations, syncope, suspected arrhythmia. - Mean age total cohort: 61.8 ± 18.6 years - Mean age 'frequent PACs' group: 70.06 ± 15.79 years - Mean age 'infrequent PACs' group: 56.6 ± 18.3 years - Men: 60.0%	R	10.0	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Marinheiro 2017	PT	362*	Consecutive patients referred for elective 24-hour Holter monitoring between 2005 and 2010. - Mean age total cohort: 71.3 years - Mean age 'frequent PACs' group: 71.2 years - Mean age 'infrequent PACs' group: 71.4 years - Men: 56.4%	P	7.1	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Murakoshi 2015	JP	63,197	Individuals who participated in annual community-based health check-ups in 1993 and were followed until 2008. - Mean age total cohort: 58.8 ± 9.9 years - Mean age 'frequent PACs' group: 64.8 ± 8.5 years - Mean age 'infrequent PACs' group: 58.4 ± 9.9 years - Men: 32.4%	P	AF: 5.8 ACM: 14.3	Annual follow-up ECG

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Nguyen 2017	US	ARIC: 15,132 CHS: 5,577	ARIC cohort: Probability sample from a community-based cohort of patients aged 45-64 years from 4 communities, recruited between 1987 and 1998. CHS cohort: Sample from a community-based cohort of Medicare beneficiaries ≥65 years, recruited between 1989 and 1990, with an oversampling of black residents. - Mean age total ARIC cohort: 54.1 years - Mean age total CHS cohort: 71.2 years - Median age 'frequent PACs' group ARIC cohort: 58 years (IQR 51-62) - Median age 'frequent PACs' group CHS cohort: 75 years (IQR 71-80) - Median age 'infrequent PACs' group ARIC cohort: 54 years (IQR 49-59) - Median age 'infrequent PACs' group CHS cohort: 71 years (IQR 68-76) - Men in ARIC cohort: 44.7% - Men in CHS cohort: 41.7%	P	ARIC: 10.0 CHS: 12.0	ARIC cohort: annual telephonic follow-up with repeated clinic visits + ECG every 3 years CHS cohort: Annual study ECG; AF and mortality diagnoses from reviewing medical records and death certificates
Nortamo 2017	FI	1,710	Patients >18 or <85 years with angiographically documented coronary artery disease, visiting a university hospital between 2007 and 2011. - Mean age total cohort: 66.5 years - Mean age 'frequent PACs' group: 69 ± 8 years - Mean age 'infrequent PACs' group: 66 ± 9 years - Men: N/A	P	5.6	Contacted regularly by telephone and/or mail to check if they had experienced cardiac events, follow-up ECG at physician's discretion
O'Neal 2016	US	22,975	Participants from general population, with an over sampling of blacks and residents of the stroke belt (North Carolina, South Carolina, Georgia, Alabama, Mississippi, Tennessee, Arkansas and Louisiana), between 2003 and 2007 - Mean age total cohort: 64.4 years - Mean age 'frequent PACs' group: 69 ± 9.6 years - Mean age 'infrequent PACs' group: 64 ± 9.1 years - Men: 43.7%	P	7.1	6-monthly telephonic follow-up, after which local investigator confirmed self-reported stroke episodes from medical records

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
O'Neal 2017	US	13,840	Participants from general population, with an over sampling of blacks and residents of the stroke belt (North Carolina, South Carolina, Georgia, Alabama, Mississippi, Tennessee, Arkansas and Louisiana), between 2003 and 2007. Study limited to those who completed a follow-up visit at ±10 years follow-up - Mean age total cohort: 63 ± 8.4 years - Mean age 'frequent PACs' group: 67 ± 8.9 years - Mean age 'infrequent PACs' group: 63 ± 8.4 years - Men: 44%	P	9.4	6-monthly telephonic follow-up, after which local investigator confirmed self-reported stroke episodes from medical records; all completed a follow-up visit with ECG at ±10 years follow-up
Perez 2009	US	42,751	Veterans undergoing ECG for usual clinical indications between March 1987 and July 2000 - Mean age total cohort: 56.1 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 89.7%	R	5.3	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Pinho 2015	PT	205	Consecutive patients admitted for acute ischemic stroke or TIA of undetermined etiology between 2005 and 2012. - Mean age total cohort: 55.2 ± 15.1 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 52.2%	R	2.3	Follow-up ECG at physician's discretion; AF, stroke, and TIA diagnosed from reviewing medical records
Qureshi 2014	US	7,394	A representative sample of community dwelling individuals ≥20 years old enrolled between 1988 and 1994. - Mean age total cohort: 59.2 years - Mean age 'frequent PACs' group: 72 ± 11 years - Mean age 'infrequent PACs' group: 59 ± 13 years - Men: 53.1%	P	13.0	Mortality data obtained from national register

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Raman 2017	US	2,350	Community-dwelling men aged 65 and older able to ambulate without assistance, and without history of bilateral hip replacement without baseline AF or PM, enrolled between 2000 and 2002 at 6 centers in the US. - Mean age total cohort: 75.8 ± 5.3 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 100.0%	P	8.0	Follow-up ECG at physician's discretion; queried every 4 months about cardiovascular events requiring hospitalization or emergency department visit by mailed questionnaire and/or telephone contact
Suzuki 2013	JP	2,589	All new patients referred to a cardiac care hospital for any indication between 2004 and 2011. - Mean age total cohort: 54.2 ± 15.5 years - Mean age 'frequent PACs' group: 63.7 ± 13.3 years - Mean age 'infrequent PACs' group: 51.0 ± 14.8 years - Men: 55.5%	R	1.6	Follow-up ECG at physician's discretion; AF diagnosed from reviewing medical records
Thijs 2016	CA, EU, US	221	Consecutive patients ≥40 years admitted for acute cryptogenic stroke or TIA, randomized to the intervention arm of the CRYSTAL-AF trial between 2009 and 2012. - Mean age total cohort: 61.6 ± 11.4 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 64.3%	P	1.7	All patients received ILR on admission for acute cryptogenic stroke; scheduled follow-up visits at 1 month, 6 months, 12 months, and every 6 months thereafter

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Vinther 2016	DK	565	Consecutive patients admitted for acute ischemic stroke between 2008 and 2011. (retrospective) - Mean age total cohort: 71.5 years - Mean age 'frequent PACs' group: 75.4 ± 9.9 years - Mean age 'infrequent PACs' group: 69.9 ± 13.5 years - Men: 55.4%	R	4.0	Follow-up ECG at physician's discretion; AF and stroke diagnosis from reviewing medical records; mortality from national registry
Vinther 2017	DK	167	Consecutive patients admitted for acute ischemic stroke between 2012 and 2014. - Mean age total cohort: 69.9 years - Mean age 'frequent PACs' group: 78 ± 8.3 years - Mean age 'infrequent PACs' group: 68 ± 12.7 years - Men: 60.5%	P	2.7	Follow-up 48h Holter at 6 and 12 months; AF, stroke and all-cause mortality further diagnosed from reviewing medical records and national register
Wallmann 2003	CH	99	Consecutive patients who suffered an acute ischaemic stroke, inclusion period not reported - Mean age total cohort: 63.2 years - Mean age 'frequent PACs' group: 70 ± 12 years - Mean age 'infrequent PACs' group: 61 ± 14 years - Men: 57.6%	P	1.9	Follow-up ECG at primary care physician's discretion; AF diagnosed from reviewing medical records
Wallmann 2007	CH	127	Consecutive patients who suffered an acute ischaemic stroke, inclusion period not reported - Mean age total cohort: 61.5 years - Mean age 'frequent PACs' group: 70 ± 9 years - Mean age 'infrequent PACs' group: 56 ± 14 years - Men: 54.7%	P	0.5	All patients received follow-up 7d Holter at 0, 3, and 6 months

Supplementary Table S2. Characteristics of the included studies. (*continued*)

Study	Country	Number of participants	Study population	Study design	Mean follow-up (years)	Outcome ascertainment
Weber-Krüger 2017	DE	184	Consecutive patients with acute cerebral ischemia between 2009 and 2010. - Mean age total cohort: 71.7 years - Mean age 'frequent PACs' group: 74 years - Mean age 'infrequent PACs' group 66 years - Men: 61.4%	P	3.7	Follow-up ECG after 3 years or telephonic follow-up
Yamada 2000	JP	75	Consecutive outpatients with stable CHF screened in heart failure unit of a general hospital between 1995 and 1997. - Mean age total cohort: 65 ± 11 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 81.3 %	P	1.8	Follow-up ECG every 2 weeks during initial 8 weeks, every 4 weeks thereafter; 24h Holter and echocardiography at 2 and 6 months after study entry, every 6 months thereafter
Yodogawa 2013	JP	68	Consecutive patients admitted for acute ischemic stroke between 2006 and 2011. - Mean age total cohort: 69.9 ± 9.6 years - Mean age 'frequent PACs' group: N/A - Mean age 'infrequent PACs' group: N/A - Men: 54.4%	P	0.9	All patients examined with 12-lead ECG every 4 weeks; 24-h Holter every 3 months thereafter

AF = atrial fibrillation; ARIC = Atherosclerosis Risk In Communities; CA = Canada; CH = Switzerland; CHF = congestive heart failure; CHS = Cardiovascular Health Study; CRYSTAL-AF = Cryptogenic Stroke and Underlying Atrial Fibrillation; DE = Germany; DK = Denmark; ECG = electrocardiogram; EMBRACE = Event Monitoring Belt for Recording Atrial Fibrillation After a Cerebral Ischemic Event; ES = Spain; EU = European Union; FI = Finland; h = hour; HK = Hong Kong; ILR = internal loop recorder; JP = Japan; KR = South Korea; N/A = not available; NL = The Netherlands; P = prospective; PAC = premature atrial contraction; PM = pacemaker; PT = Portugal; R = retrospective; SE = Sweden; TIA = transient ischemic attack; TW = Taiwan; US = United States. * Marinheiro 2017 divided their cohort (n = 2,480) into those with upper 5th percentile of PACs/h on 24h baseline Holter (>97 PACs/h, n = 124) and a propensity score matched group with (≤97 PACs/h, n = 238)

Supplementary Table S3. Results of the included studies.

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Acharya 2015	H-24	1) ≥100 PACs/24h (Dic)	- AF	1 → AF: UV HR 4.34 (3.09-6.08); MV HR 2.97 (1.85-4.80)	ABCDE	- AF in total cohort: 155/1357 (11.4%) - AF in 'frequent PACs' group: 106/486 (21.8%) - AF in 'infrequent PACs' group: 49/871 (5.6%)
Binici 2010	H-48	1) ≥30 PACs/h or any runs of ≥20 PACs (Dic) 2) PACs/h (Con, linear for each increment of 10 PACs/h) 3) Length of runs of PACs (Con, linear for lengthening of run by every 4 PACs)	- AF - FS - ACM	1) → AF: UV HR 3.19 (1.30-7.86); MV HR 2.73 (1.07-6.96) 1) → FS: UV HR 3.88 (1.78-8.48); MV HR 2.37 (1.02-5.50) 1) → ACM: UV HR 2.12 (1.30-3.47); MV HR 1.40 (0.83-2.36) 2) → AF: UV HR 1.57 (1.12-2.21); MV 1.49 (1.02-2.17) 2) → FS: UV HR N/A; MV HR 0.83 (0.58-1.98) 2) → ACM: UV HR 1.49 (1.24-1.79); MV HR 1.27 (1.04-1.55) 3) → AF: UV HR 1.30 (1.15-1.47), MV HR 1.29 (1.14-1.47) 3) → FS: UV HR 1.12 (0.97-1.29) 3) → ACM: UV HR 1.12 (1.03-1.21), MV HR 1.06 (0.97-1.15)	ABFGH	- AF in total cohort: 22/678 (3.2%); 5.5/1,000 PYs - AF in 'frequent PACs' group: 7/99 (7.1%); 12.8/1,000 PYs - AF in 'infrequent PACs' group: 15/579 (2.6%); 4.3/1,000 PYs - FS in total cohort: 27/678 (4.0%); 6.7/1,000 PYs - FS in 'frequent PACs' group: 10/99 (10.1%) 18.8/1,000 PYs - FS in 'infrequent PACs' group: 17/579 (2.9%) 4.9/1,000 PYs - ACM in total cohort: 87/678 (12.8%) 21.4/1,000 PYs - ACM in 'frequent PACs' group: 21/99 (21.2%); 37.2/1,000 PYs - ACM in 'infrequent PACs' group: 66/579 (11.4%); 18.9/1,000 PYs
Blanch Gracia 2013	E-10	1) ≥1 PACs (Dic)	- AF	1) → AF: UV HR N/A; MV HR N/A	ABCDFIJK	- AF in total cohort: 23/183 (12,6%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Cabrera 2016	H-24	1) ≥0.2% PACs/total number of beats/24h (Dic)	- AF	1) → AF: UV HR 3.64 (1.78-7.4); MV HR 2.7 (1.2-5.8)	ABCGKLMNO	- AF in total cohort: 31/299 (10.4%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A
Chong 2012	H-24	1) ≥100 PACs/24h (Dic)	- AF - FS - ACM	1) → AF: UV HR 3.9 (3.2-11.1) ; MV HR 3.22 (1.9-5.5) 1) → FS: UV HR 2.1 (1.1-4.8); MV HR N/A 1) → ACM: UV HR 1.8 (1.1-3.6); MV HR N/A	ALP NB: L for outcome AF only	- AF in total cohort: 60/428 (14.0%) - AF in 'frequent PACs' group: 31/107 AF (29.0%); 48.3/1,000 PYs - AF in 'infrequent PACs' group: 29/321 (9.0%); 14.3/1,000 PYs - FRS in total cohort: 41/428 (9.6%) - FRS in 'frequent PACs' group: 16/107 (15.0%) - FRS in 'infrequent PACs' group: 25/321 (7.8%) - ACM in total cohort: 69/428 (16.1%) - ACM in 'frequent PACs' group: 24/107 (22.4%) - ACM in 'infrequent PACs' group: 45/321 (14.0%)
Chun 2016	H-24	1) >5409PACs/24h (Dic)	- AF	1) → AF: UV HR: N/A; MV HR 0.559 (0.275-1.137)	ACK	- AF in total cohort: 64/684 (9.4%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Dewland 2013	H-24	1) PACs /h (Con, LT †) 2) ≥9.5 PACs/h (Ord, Q4 [≥9.5 PACs/h] vs Q1 [<0.8PACs/h])	- AF ‡ - ACM	1) → AF: UV HR 1.73 (1.55-1.94); MV HR 1.68 (1.48-1.91) 1) → ACM: UV HR; MV HR 2) → AF: UV HR 5.01 (3.50-7.17) ‡; MV HR 4.92 (3.39-7.16) ‡ 2) → ACM: UV HR 1.93 (1.59-2.34); MV HR 1.35 (1.10-1.66)	ABCFGILNZ	- AF in total cohort: 343/1260 (27.2%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A - ACM in total cohort: 837/1260 (66.4%); 573/1260 (45.5%) died without known AF - ACM in groups 'frequent PACs' and 'infrequent PACs': N/A
Engstrom 2000	H-24	1) ≥218 PACs/24h (Dic)	- ACM	1) → ACM: UV HR N/A; MV HR 1.3 (0.88-1.8)	FGLP	- ACM in total cohort: 170/388 (43.8%) - ACM in 'frequent PACs' group: 40/77 (51.9%); 48.3/1,000 PYs - ACM in 'infrequent PACs' group: 130/311 (41.8%); 37.1/1,000 PYs
Folkeringa 2006	ETE	1) PACs/min during ETE (Con, LN) 2) PACs/min during recovery from ETE (Con, LN)	- AF	1) → AF: Unadjusted Pearson's r: 0.17, no 95% CI reported, p-value reported (= 0.410); adjusted Pearson's r: N/A 2) → AF: Unadjusted Pearson's r: 0.171, no 95% CI reported, p-value reported (= 0.006); adjusted Pearson's r: no point estimate reported, 95% CI: 0.14-0.98, p-value reported (= 0.010)	CK	- AF in total cohort: 17% - AF in groups 'frequent PACs' and 'infrequent PACs': N/A

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Gladstone 2015	H-24	1) PACs/24h (Con, LN) 2) Number of runs ≥4 PACs/24h (Con, linear)	- AF	1) → AF at 2 years: UV logistic regression N/A; MV logistic regression: only p-value for odds ratio reported (= 0.0027)	ACK	- AF in total cohort: 38/237 (16.0%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A
Inohara 2013	E-10	1) ≥1 PAC (Dic)	- ACM	1) → ACM: UV HR 3.98 (2.77-5.71); MV HR 1.55 (1.07-2.24)	ABCFGHIMPR	- ACM in total cohort: 1,211/7,692 (15.7%) - ACM in 'frequent PACs' group: 30/64 (46.9%) - ACM in 'infrequent PACs' group: 1,181/7,628 (15.5%)
Johnson 2015	H-24	1) ≥30 PACs/h or any runs of ≥20 PACs (Dic) 2) PACs/h (Con, LT) 3) Number of runs of ≥3 PACs/h (Con, LN)	- AF/Afl ‡	1) → AF: UV HR 3.91 (2.12-7.21) ||; MV HR 2.66 (1.38-5.12) ‡ 2) → AF: UV HR N/A; MV HR 1.39 (1.16-1.68) ‡ 3) → AF: UV HR N/A; MV HR 1.95 (1.21–3.13)	ABFGPRST	- AF in total cohort: 45/383 (11.7%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A
Kochhauser 2014	H-24	1) >14.1 PACs/h (Dic) 2) Run of >3 PACs (Dic)	- AF	1) → AF: RR 4.0 (1.1-14.6) 2) → AF: RR 6.9 (1.8-26.7)	ABCGK	- AF in total cohort: 12/70 (17.1%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Lin 2015	H-24	1) >76 PACs/24h (Dic)	- AF - ACM	1) → AF: UV HR 2.31 (1.90-2.80); MV HR 1.76 (1.43-2.16) 1) → ACM: UV HR 2.19 (1.95-2.45); MV HR 1.38 (1.23-1.59)	ABDGLN	- AF in total cohort: 418/5,371 (7.8%) - AF in 'frequent PACs' group: 242/2,072 (11.7%) - AF in 'infrequent PACs' group: 176/3,299 (5.3%) - ACM in total cohort: 1,209/5,371 (22.5%) - ACM in 'frequent PACs' group: 671/2072 (32.4%) - ACM in 'infrequent PACs' group: 538/3299 (16.3%)
Marinheiro 2017	H-24	1) >97 PACs/h (Dic) 2) PACs/h (Con, LT) 3) >97 PACs/h (Ord, T3 [>97/h] vs Q1 [<30/h])	- AF - FS - ACM	1) → AF: UV HR 1.87 (1.24-2.81); MV HR 1.76 (1.17-2.66) || 1) → FS: UV HR 2.17 (1.12-4.19); MV HR 2.01 (1.03-3.93) || 1) → ACM: UV HR 2.04 (1.35-3.08); MV HR 1.84 (1.21-2.81) || 2) → AF: UV HR 1.44 (1.14-1.81); MV HR 1.40 (1.15-1.72) || 2) → FS: UV HR 1.66 (1.18-2.35); MV HR 1.71 (1.21-2.43) || 2) → ACM: UV HR 1.38 (1.11-1.71); MV HR 1.38 (1.11-171) || 3) → AF: UV HR 2.05 (1.31-3.20); MV HR 2.05 (1.31–3.23) 3) → FS: UV HR 2.71 (1.60-4.61); MV HR 2.83 (1.65–4.84) 3) → ACM: UV HR 2.40 (1.54-3.73); MV HR 2.17 (1.48–3.28)	ABFGHILMNP	- AF in total cohort: 39/362 (10.8%) - AF in 'frequent PACs' group: 67.2/1,000 PYs - AF in 'infrequent PACs' group: 33.3/1,000 PYs - FS in total cohort: 54/362 (14.9%) - FS in 'frequent PACs' group: 34.9/1,000 PYs - FS in 'infrequent PACs' group: 11.5/1,000 PYs - ACM in total cohort: 129/362 (35.6%) - ACM in 'frequent PACs' group: 77.8/1,000 PYs - ACM in 'infrequent PACs' group: 33.3/1,000 PYs

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Murakoshi 2015	E-15	1) ≥1 PAC (Dic)	- AF - ACM	1) → AF: UV HR 6.40 (5.12-7.98); MV HR 4.44 (3.53-5.60) ** 1) → ACM: UV HR 1.93 (1.70-2.20); MV HR 1.14 (1.02-1.27) **	ABCDFGHIJMPR	- AF in total cohort: 386/63,197 (0.6%); 1.05/1,000 person-years - AF in 'frequent PACs' group: 112/3,858 (2.9%) - AF in 'infrequent PACs' group: 274/59,339 (0.4%) - ACM in total cohort (at 10 years follow-up): 4,178/63,197 (6.6%) - ACM in 'frequent PACs' group (at 10 years follow-up): 483/3,858 (12.5%) - ACM in 'infrequent PACs' group (at 10 years follow-up): 3,695/59,339 (6.2%)
Nguyen 2017	E-10	1) ≥1 PAC (Dic)	- AF	1) → AF in ARIC cohort: UV HR 2.8 (2.2-3.6); MV HR 2.4 (1.9-3.1) 1) → AF in CHS cohort: UV HR 1.9 (1.5-2.4); MV HR 1.6 (1.3-2.0)	ABFGILNUZ	- AF in total ARIC cohort: 2,059/15,132 (13.6%) - AF in total CHS cohort: 1,534/5,577 (27.5%) - AF in groups 'frequent PACs' and 'infrequent PACs' in ARIC & CHS cohorts: N/A
Nortamo 2017	H-24	1) Run of ≥4 PACs <30 sec (Dic) 2) ≥1,427 PACs/24h (Ord, Q4 [≥1,427] vs Q1 [≤507])	- AF	1) → AF: UV HR 2.597 (1.828-3.689); MV HR 2.529 (1.763-3.628) 2) → AF: UV HR 8.253 (4.133-16.479); MV HR 8.139 (3.967-16.696)	AFGILKP	- AF in total cohort: 143/1,710 (8.4%) - AF in 'frequent PACs' group: 47/275 (17%) - AF in 'infrequent PACs' group: 96/1,435 (7%)

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
O'Neal 2016	E-10	1) ≥1 PAC (Dic)	- FS - RS §	1) → FS: UV HR N/A; MV HR 1.35 (1.05-1.75) 1) → RS: UV HR N/A; MV HR 1.52 (0.94-2.45) §	ABCDFGHLPSUVZ	- FRS in total cohort: 549/22,975 (2.4%) - FRS in 'frequent PACs' group: 68/1,687 (4.0%); 6.0 (4.7-7.6)/1,000 PYs - FRS in 'infrequent PACs' group: 481/21,288 (2.3%); 3.2 /1,000 PYs
O'Neal 2017	E-10	1) ≥1 PAC (Dic)	- AF	1) → AF: UV OR N/A; MV OR 1.92 (1.57, 2.35)	ABDFGHILNPSUVZ	- AF in total cohort: 1,015/12,840 (7.3%) - AF in 'frequent PACs' group: 139/950 (15%) - AF in 'infrequent PACs' group: 876/12,890 (6.8%)
Perez 2009	E-120	1) ≥1 PAC (Dic)	- AF	1) → AF: UV HR N/A; MV HR 2.1 (1.6-2.7)	ABC	- AF in total cohort: 1,050/42,751 (2.5%) - AF in 'frequent PACs' group: 80/872 (9.2%) - AF in 'infrequent PACs' group: 970/41,871 (2.3%)

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Pinho 2015	H-24	1) >30 PACs/h (Dic) 2) PACs/h (Con, LT)	- AF/Afl - RST # - RS - RT	1) → AF: UV HR 4.04 (1.43 – 11.40); HR MV: 2.28 (0.69 - 7.55) || 1) → RST: - Among all n = 205: UV HR 4.33 (1.67-11.22); MV HR 2.19 (0.80-6.02) || - Among n = 184 that did not develop AF or Aflutter during follow-up: UV HR 5.33 (1.86-15.33); MV HR 3.34 (1.05-10.63) # 1) → RS: UV HR 3.11 (0.85 – 11.41); MV HR 1.16 (0.30 – 4.53) || 1) → RT: UV HR 6.58 (1.89 – 22.90); MV HR 2.68 (0.67 – 10.71) || 2) → AF: UV HR 1.92 (1.29 – 2.87); MV HR 1.44 (0.92 – 2.25) || 2) →RST: UV HR 1.69 (1.18-2.43); MV HR 1.21 (0.81-1.81)|| 2) → RS: UV HR 1.34 (0.85-2.12); MV HR 0.89 (0.56-1.42) || 2) → RT: UV HR 2.50 (1.47-4.24); MV HR 1.82 (0.97-3.41) ||	ADFGHP	- AF or Aflutter in total cohort: 21/205 (10.2%) - AF in groups 'frequent PACs' and 'infrequent PACs': N/A - RST in total cohort: 22/184 (12.0%) - RST in groups 'frequent PACs' and 'infrequent PACs': annual recurrence rate at <10 PACs/h = 2.9%, at 10-30 PACs/h = 11.0%, at >30 PACs/h = 22.6%

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Qureshi 2014	E-10	1) ≥1 PAC (Dic)	- ACM	1) → ACM: UV HR N/A; MV HR 1.41 (1.08-1.80)	ABCDFGHIPWXZ	- ACM in total cohort: 2,458/7,394 (33.2%) - ACM in 'frequent PACs' group: 72/89 (80.9%); 0.858/1000 PYs - ACM in 'infrequent PACs' group: 2,386/7,305 (32.7%); 0.254/1000 PYs
Raman 2017	PSG	1) ≥21.15 PACs/h (Dic) 2) ≥21.15 PACs/h (Ord, Q4 [≥21.15] vs Q1 [<2.19]) 3) PACs/h (Con, LT) 4) ≥5 PACs/h (Dic)	- AF	1) → AF: UV HR 1.77 (1.38-2.28); MV HR 1.53 (1.18-1.98) || 2) → AF: UV HR 3.56 (2.35-5.39); MV HR 2.99 (1.94-4.62) 3) → AF: UV HR 1.58 (1.35-1.84); MV HR 1.44 (1.23-1.70) || 4) → AF: UV HR N/A; MV HR N/A	ADGLZ in 1) and 3) Additionally FHINRUW in 2) and 4)	- AF in total cohort: 269/2,350 (11.4%) - AF in group '≥5 PACs/h': 185/1278 (14.5%) - AF in group '<5 PACs/h': 84/1072 (7.8%)
Suzuki 2013	H-24	1) ≥102 PACs/24h (Dic)	- AF	1) → AF: UV HR 9.44 (4.85-18.37) ; MV HR 6.89 (3.45-13.74) ††	ABCDKMP	- AF in total cohort: 38/2589 (1.5%) - AF in 'frequent PACs' group: 27/647 (3.4%); 23.8/1,000 PYs - AF in 'infrequent PACs' group: 11/1942 (0.6%); 15.5/1,000 PYs
Thijs 2016	H-24	1) >123 PACs/24h (Ord, Q4 [>123/24h] vs Q1 [0/24h])	- AF	1) → AF at 12 months: UV HR 3.94 (1.30-11.97); MV HR N/A 1) → AF at 36 months: UV HR 3.47 (1.38-8.70); MV HR N/A	ACFO	- AF at 36 months in total cohort: 42/221 (19.0%) - AF at 36 months in groups 'frequent PACs' and 'infrequent PACs': N/A

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Vinther 2016	H-48	1) Runs of ≥3 PACs (Dic)	- AF - RST - ACM	1) → AF: UV HR 1.44 (0.61-3.44); MV HR 1.24 (0.50-3.09) 1) → RST: UV HR 1.31 (0.81-2.12); MV HR 1.43 (0.88-2.33) 1) → ACM: UV HR 1.79 (1.30-2.47); MV HR 1.39 (0.99-1.96)	ABGO NB: G for outcome AF only; O for outcome stroke only	- AF in total cohort: 22/565 (3.9%) - AF in 'frequent PACs' group: 8/161 (5.0%) - AF in 'infrequent PACs' group: 14/404 (3.5%) - RST in total cohort: 73/565 (12.9%) - RST in groups 'frequent PACs' and 'infrequent PACs': N/A - ACM in total cohort: 158/565 (28.0%) - ACM in groups 'frequent PACs' and 'infrequent PACs': N/A

Vinther 2017	H-24	1) >14 PACs/h and ≥3 runs of ≥3 consecutive PACs/24h (Dic)
		2) PACs/24h (Con, LT)
		3) Runs of ≥3 PACs (Dic)
			- AF ‡
			- ACM
				1) → AF: UV HR 3.55 (0.98-12.8) ‡; MV HR 3.05 (0. 70-13.3) ‡
				1) → ACM: UV HR 3.84 (1.94-7.61) // MV 2.06 (0.97-4.35)
				2) → AF: UV HR 1.29 (1.04-1.60); MV HR 1.24 (0.89-1.73) ‡
				2) → ACM: UV HR 1.33 (1.16-1.53) // MV 1.14 (0.96-1.34)
				3) → AF: UV HR 2.18 (0.57-8.36); MV HR 1.24 (0.27-5.63)
				3) → ACM: UV HR 2.81 (1.34-5.87); MV HR 1.85 (0.87-3.92)
					ABKY
					NB: K for outcome AF only; L & Y for outcome all-cause Mortality only
						- AF in total cohort: 9/167 (5.4%)
						- AF in groups 'frequent PACs' and 'infrequent PACs': N/A
						- ACM in total cohort: 34/167 (20.4%)
						- ACM in groups 'frequent PACs' and 'infrequent PACs': N/A

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Wallmann 2003	H-24	1) ≥70 PACs/24h (Dic)	- AF	1) → AF: OR 8.9 (2.4-33.1)	AKL	- AF in total cohort: 12/99 (12.1%) - AF in 'frequent PACs' group: 8/24 (33.3%) - AF in 'infrequent PACs' group: 4/75 (5.3%)
Wallmann 2007	H-24	1) ≥70 PACs/24h (Dic)	- AF	1) → AF: OR 6.6 (1.6-28.2)	ABGK	- AF in total cohort: 18/127 (14.2%) - AF in 'frequent PACs' group: 13/50 (26.0%) - AF in 'infrequent PACs' group: 5/77 (6.5%)
Weber-Krüger 2017	H-24	1) Run ≥ 5 PACs < 30 sec (Dic)	- AF - RS - ACM	1) → AF: UV HR N/A; MV HR only p-value reported (= 0.09) 1) → RS: UV HR N/A; MV HR 2.742 (1.044-7.199) 1) → ACM: UV HR N/A; MV HR N/A	N/A	- AF in total cohort: 14/184 (7.6%) - AF in 'frequent PACs' group: 8/67 (11.9%) - AF in 'infrequent PACs' group: 6/117 (5.1%) - RS in total cohort: 18/139 (12.9%) - RS in 'frequent PACs' group: 10/48 (20.1%) - RS in 'infrequent PACs' group: 8/91 (8.8%) - ACM in total cohort: 29/184 (15.8%) - ACM in 'frequent PACs' group: 14/67 (20.9%) - ACM in 'infrequent PACs' group: 15/117 (12.8%)

Supplementary Table S3. Results of the included studies. (*continued*)

Study	Baseline recording device	Statistical definition(s) applied to PAC-count as the predictor (statistical approach)	Studied outcome(s)	Effect measure (95% CI) for relationship between prognostic factor (number) & studied outcome	Variables used in multi-variable model	Incidence rate of the studied outcome(s); absolute frequency (%) and per 1,000 PYs
Yamada 2000	H-24	1) >100PACs/24h or ≥1 runs of ≥2 PACs (Dic)	- AF	1) → AF: UV HR 3.7 (1.0-13.2) , MV 1.1 (0.2-6.3)	ACEK	- AF in total cohort: 10/75 (13.3%) - AF in 'frequent PACs' group: 6/23 (26.1%) - AF in 'infrequent PACs' group: 4/52 (7.7%)
Yodogawa 2013	H-24	1) >100 PACS/24h (Dic)	- AF - FS	1) → AF: only p-value for log-rank test for Kaplan-Meier curve reported (< 0.001) 1) → FS: only p-value for log-rank test for Kaplan-Meier curve reported (< 0.001)	ACEK	- AF in total cohort: 17/68 (25.0%); - AF in groups 'frequent PACs' and 'infrequent PACs': N/A

A = age; ACM = all-cause mortality; AF = atrial fibrillation; Aflutter = atrial flutter; ARIC = Atherosclerosis Risk In Communities; AV = atrioventricular; B = sex; C = electrocardiographic parameters; CHS = Cardiovascular Health Study; Con = continuous; D = medication; Dic = dichotomous; E = laboratory parameters; ECG = electrocardiogram; E-10 = 10-second 12-lead ECG; E-15 = 15-second 12-lead ECG; E-120 = 2-minute ECG strip; ETE = exercise test continuous ECG; F = diabetes mellitus; FRS = first or recurrent stroke; FS = first stroke; G = hypertension or systolic blood pressure; H = cholesterol levels; h = hour; H-24 = 24-hour Holter; H-48 = 48-hour Holter; I = body mass index; J = ischemic stroke; K = echocardiographic parameters; L = ischemic heart disease; M = glomerular filtration rate or serum creatinine; MV = multivariable; N = congestive heart failure; N/A = not available; O = CHADS2 or CHA_2DS_2-VASc atrial fibrillation risk score; OR = odds ratio; Ord = ordinal; P = smoking; PAC = premature atrial contraction; PSG = polysomnography continuous ECG; PYs = person-years; Q = quartile; R = alcohol use; RS = recurrent stroke; RST = recurrent stroke or transient ischemic attack (TIA); S = education; T = physical activity; U = study center; UV = univariable; V = income; W = chronic obstructive pulmonary disease; X = malignancy; Y = stroke severity; Z = race; † We derived the 10-log-transformed continuous PAC-count for Dewland 2013 by dividing the originally reported base-2 log(HR) and corresponding log(95% CI start) and log(95% CI end) by 10log(2); ‡ Death as competing risk; § Among n = 1468 who self-reported prior stroke at baseline and were initially excluded from the O'Neal cohort; || Previously unpublished data provided by the authors; # Pinho 2015's original report included only those patients who did not develop AF or Aflutter during follow-up, excluding n = 21 from analysis on continuous PAC-count for the composite outcome 'recurrent stroke or TIA'. The authors later provided us with data on the composite outcome based on all n = 205 of the originally included participants for both dichotomized and continuous PAC-count; ** The presented adjusted HRs and 95% CIs for Murakoshi 2015 are the result of meta-analysis of two separate results presented by Murakoshi et al., since the study reports only separate analyses for men (n = 20,476) and women (n = 42,721); †† The presented adjusted HRs and 95% CIs for Suzuki 2013 are the result of meta-analysis of two separate results presented by Suzuki et al., since study reports only separate analyses for group CHADS2<2 (n = 2361) and group CHADS2≥2 (n = 231)

SUPPLEMENTARY RESULTS

Subgroup analysis on the association between dichotomized PAC-count on ECG and AF

We performed no subgroup analysis according to overall risk of bias, population or follow-up duration since subdivision would result in less than the required 3 studies for meta-analysis.

Subgroup analysis on the association between dichotomized PAC-count on Holter and AF

Subgroup analysis according to overall risk of bias showed similar results as the overall analysis for group 'moderate overall risk of bias', but could not replicate the findings of the overall analysis in group 'low overall risk of bias' due to a non-significant 95% PI (Supplementary Figure 3). In subgroup analysis according to population, the only subgroup not to result in similar findings as the overall analysis was group 'general population, not necessarily assigned to baseline Holter for cardiac symptoms' (Supplementary Figure 4). In subgroup analysis according to follow-up duration, subgroup '≥ 5 years follow-up' resulted in similar findings as the overall analysis, whereas in subgroup '< 5 years follow-up' the association resulted in a non-significant 95% PI (Supplementary Figure 5).

Subgroup analysis on the association between PAC-count as an ordinal variable on Holter and AF

Subgroup analyses according to overall risk of bias as well as population resulted in similar findings as the overall analysis for all meta-analyzable subgroups (Supplementary Figures 7 & 8). In subgroup analysis according to follow-up duration the results of the overall analysis could not be replicated among 4 studies with follow-up duration ≥ 5 years due to a non-significant 95% PI (Supplementary Figure 8).

Subgroup analysis on the association between PAC-count as a continuous variable on Holter and AF

In subgroup analyses according to overall risk of bias, population, and follow-up duration all meta-analyzable subgroups showed similar results as in the overall analysis (Supplementary Figures 10-12).

Subgroup analysis on the association between PAC-count on Holter and ischemic stroke and/or TIA

We performed no subgroup analysis according to population or follow-up duration since subdivision would result in less than the required 3 studies for meta-analysis.

Subgroup analysis on the association between dichotomized PAC-count on Holter and all-cause mortality

Subgroup analysis according to risk of bias was not possible as all 6 studies were of moderate overall risk of bias. In subgroup analyses according to population as well as follow-up duration all meta-analyzable subgroups resulted in similar results as the overall analysis (Supplementary Figures 13 & 14).

Subgroup analysis on the association between PAC-count as a continuous variable on Holter and all-cause mortality

In subgroup analyses according to overall risk of bias as well as follow-up duration all meta-analyzable groups resulted in similar conclusion as in the overall analysis (Supplementary Figures 16 & 17). Subgroup analysis according to population was not possible since none of the subgroups included ≥3 studies.

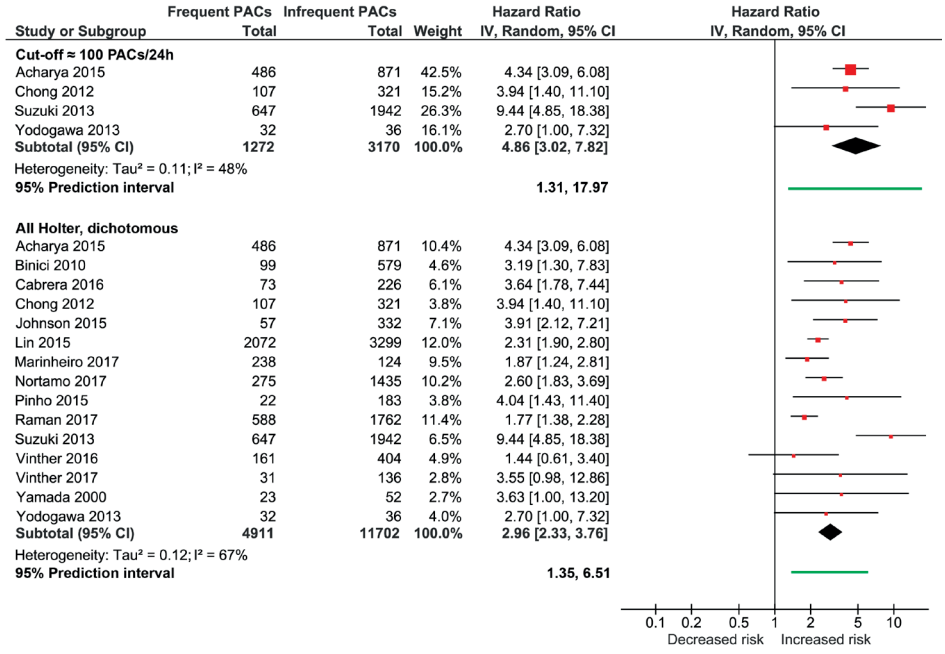
Discussion on the subgroup analyses

In our subgroup analyses arguably the most interesting finding is the difference in conclusions between subgroup 'general population, not necessarily assigned to baseline Holter for cardiac symptoms' and its respective overall meta-analysis. In this subgroup, we found a significantly positive 95% CI but a non-significant 95% PI, whereas in subgroups 'general population, assigned to baseline Holter for cardiac symptoms' as well as 'post-stroke patients' the findings remained significantly positive (Supplementary Figure 4). The subgroup analysis leads us to conclude that there is currently insufficient evidence for an association between dichotomized baseline PAC-count and future AF detection among asymptomatic community-dwelling patients. Indeed it may be the case that PAC-count can be more valuable in terms of outcome prediction in some populations than in others. However, we argue that the high statistical heterogeneity, and therefore wide 95% PI among these 3 studies which led us to revising our conclusions for this subgroup, may largely be due to the differences in applied cut-offs for dichotomization between the studies. Binici et al. as well as Johnson et al. both apply the cut-off '≥30 PACs/h or any runs of ≥20 PACs' on baseline 48-hour and 24-hour Holter, respectively. Raman et al. use the cut-off '≥21.15 PACs/h' on baseline polysomnography continuous ECG (Supplementary material, Table S2). As stated by Gladstone et al., the statistical definition of PAC-count as the predictor has previously been suggested to play a major role in PAC-count's ability to accurately predict the outcome. Indeed, meta-analysis of only the studies by Binici and Johnson on dichotomized Holter data for AF (not shown) resulted in 0% statistical heterogeneity, meaning all heterogeneity in meta-analysis of the 3 studies was caused by the data from Raman et al. Moreover, we saw in meta-analysis of continuous Holter data for AF (Supplementary Figure 11) that the data by Binici et al.

and Raman et al. completely overlap when regarded on a statistically more homogeneous scale. A re-analysis of Raman's dichotomized Holter data according to a similar cut-off as applied in Binici et al. and Johnson et al., or – ideally – an individual patient data (IPD) meta-analysis of the 3 studies would be advised for definitively answering whether PAC-count dichotomized by a homogeneous cut-off is associated with future AF detection among asymptomatic community-dwelling patients as well. Until such an analysis is performed, however, we are still to conclude that the current evidence suggests a difference in prognostic value of PAC-count on Holter for AF among different population types.

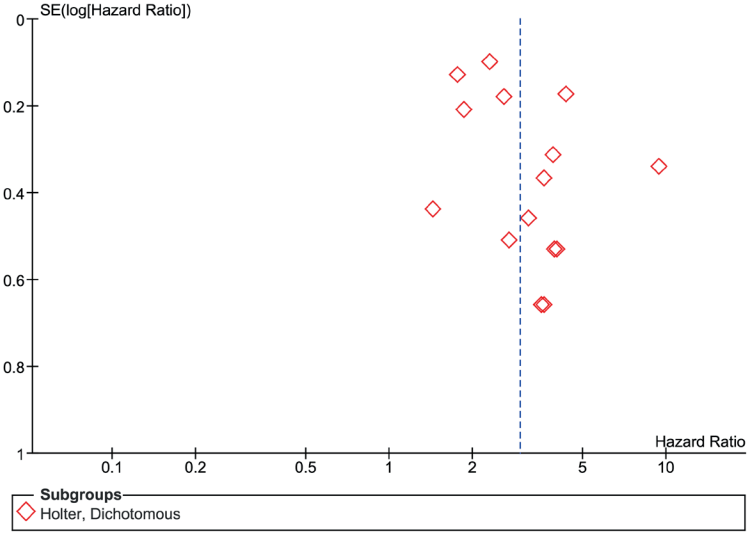
In light of study characteristics predetermined for subgroup analysis, we note that all 3 available cohorts for our analysis of ECG data for AF were performed in asymptomatic community-dwelling patients. We therefore encourage researchers to publish any available data on the association between presence of ≥1 PACs on ECG and future AF detection among populations other than asymptomatic community-dwelling patients. This may help determine the clinically important question whether there still may be a significant prognostic value of finding a PAC on ECG for AF among e.g. patients with cardiac symptoms or post-stroke patients.

Supplementary Figure 1. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on dichotomized PAC-count on baseline Holter, grouped according to cut-off for dichotomization.



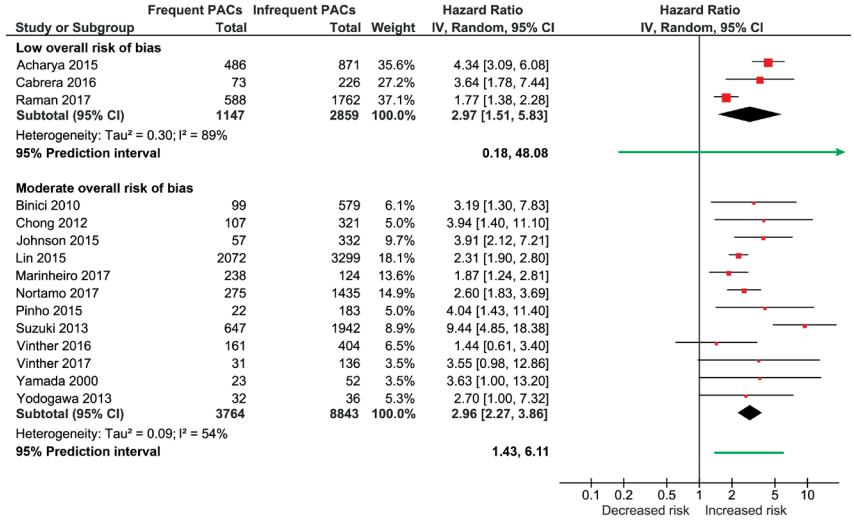
CI = confidence interval; IV = inverse variance; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 2. Funnel plot of studies that reported unadjusted hazard ratios for the outcome AF based on dichotomized PAC-count on Holter.



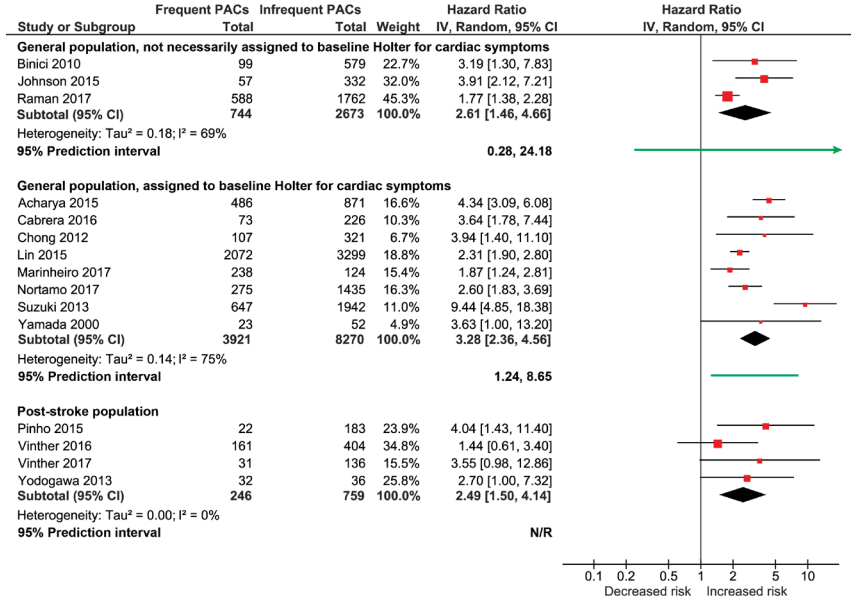
SE = standard error.

Supplementary Figure 3. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on dichotomized PAC-count on baseline Holter, grouped according to overall risk of bias.



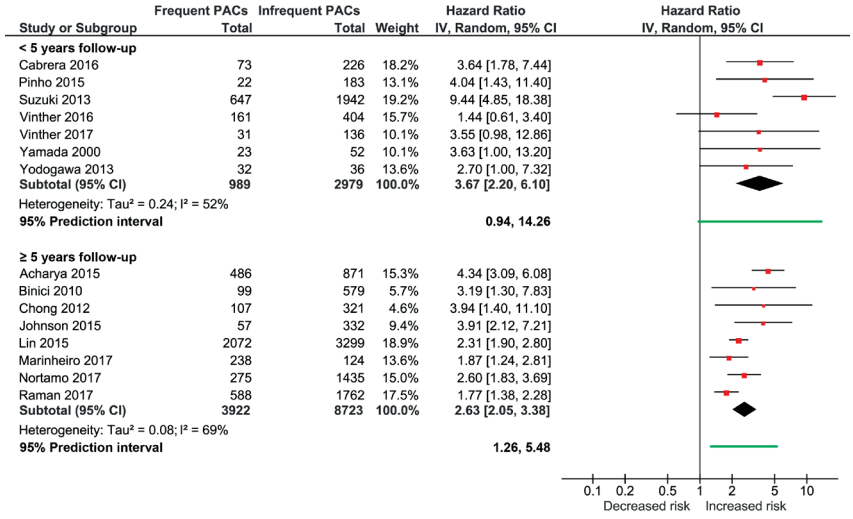
CI = confidence interval; IV = inverse variance; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 4. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on dichotomized PAC-count on baseline Holter, grouped according to population.



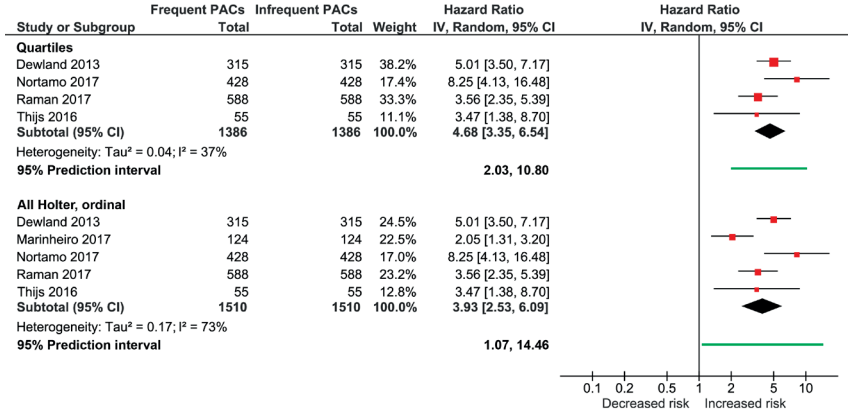
CI = confidence interval; IV = inverse variance; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 5. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on dichotomized PAC-count on baseline Holter, grouped according to follow-up duration.



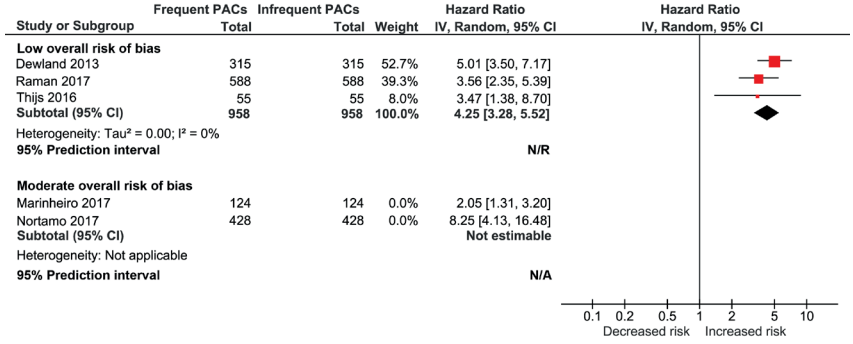
CI = confidence interval; IV = inverse variance; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 6. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on ordinal PAC-count on baseline Holter, grouped according to categorization of the ordinal variable.



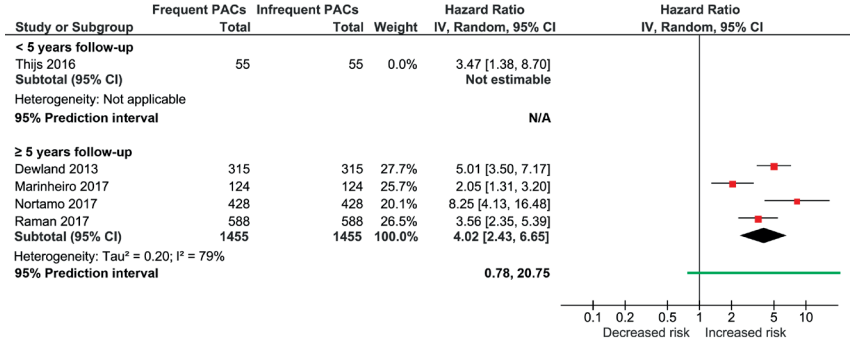
CI = confidence interval; IV = inverse variance; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 7. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on ordinal PAC-count on baseline Holter, grouped according to overall risk of bias.



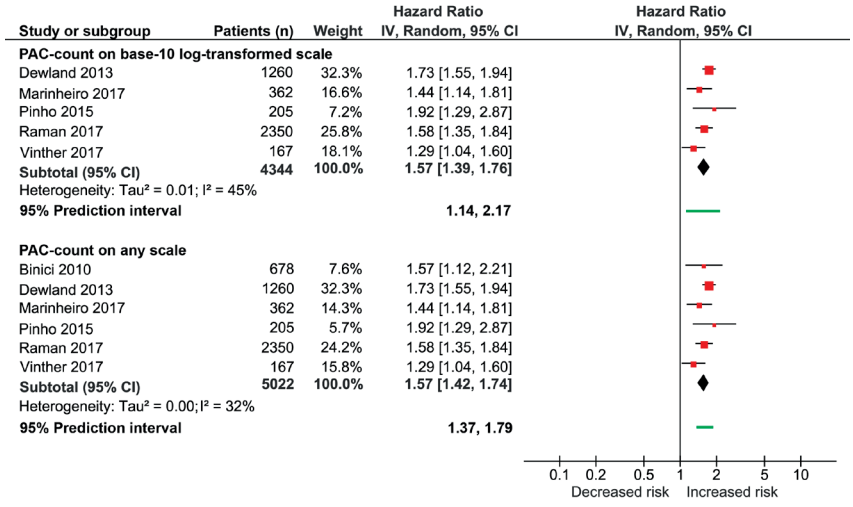
CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 8. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on ordinal PAC-count on baseline Holter, grouped according to follow-up duration.



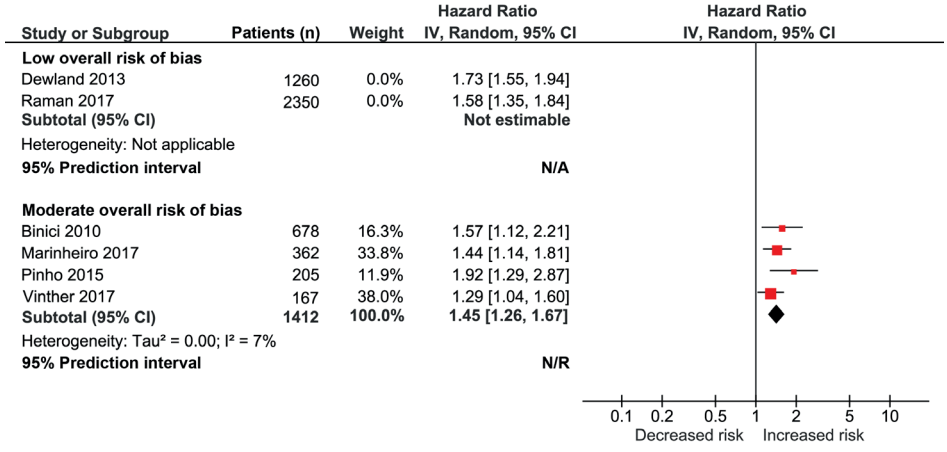
CI = confidence interval; IV = inverse variance; N/A = not applicable; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 9. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on continuous PAC-count on baseline Holter, grouped according to scale applied to continuous PAC-count.



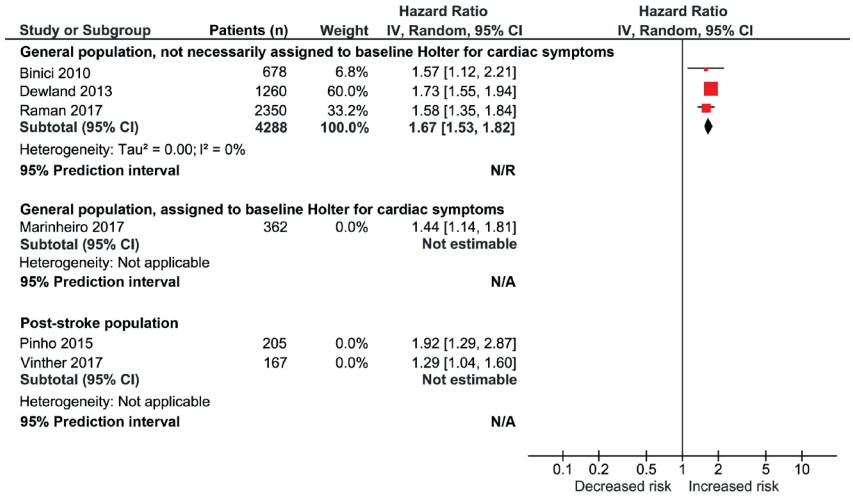
CI = confidence interval; IV = inverse variance; PACs = premature atrial contractions. Totals under 'Patients (n)' represent total cohort size.

Supplementary Figure 10. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on continuous PAC-count on baseline Holter, grouped according to overall risk of bias.



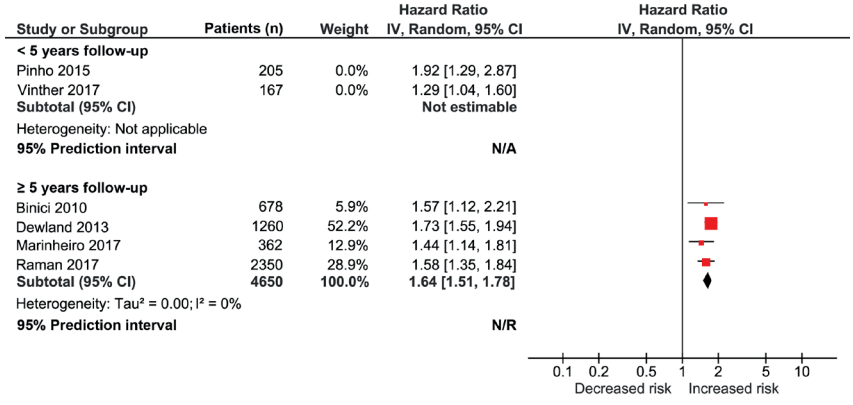
CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Patients (n)' represent total cohort size.

Supplementary Figure 11. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on continuous PAC-count on baseline Holter, grouped according to population.



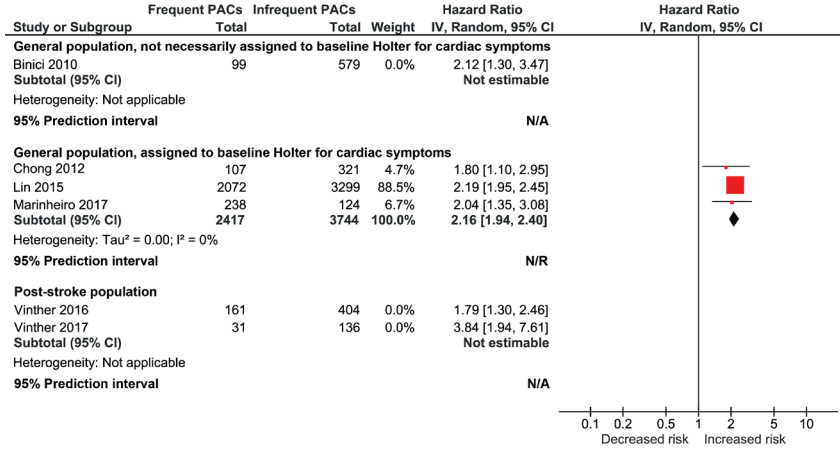
CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Patients (n)' represent total cohort size.

Supplementary Figure 12. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome AF based on continuous PAC-count on baseline Holter, grouped according to follow-up duration.



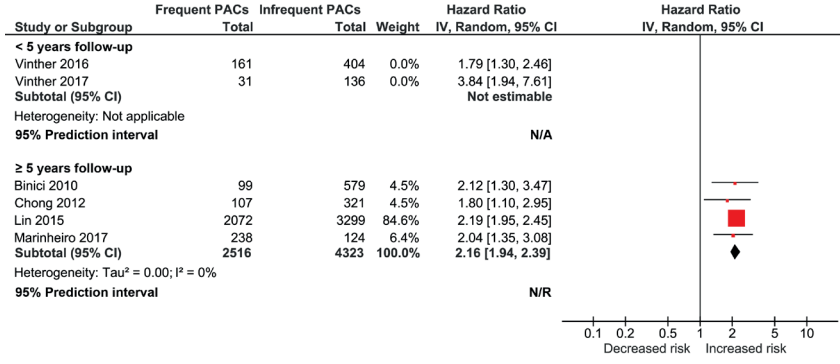
CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Patients (n)' represent total cohort size.

Supplementary Figure 13. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome all-cause mortality based on dichotomized PAC-count on baseline Holter, grouped according to population.



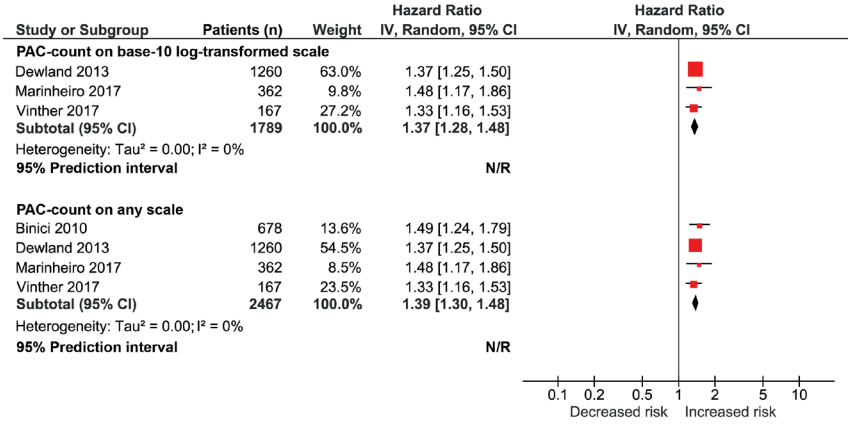
CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 14. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome all-cause mortality based on dichotomized PAC-count on baseline Holter, grouped according to follow-up duration.



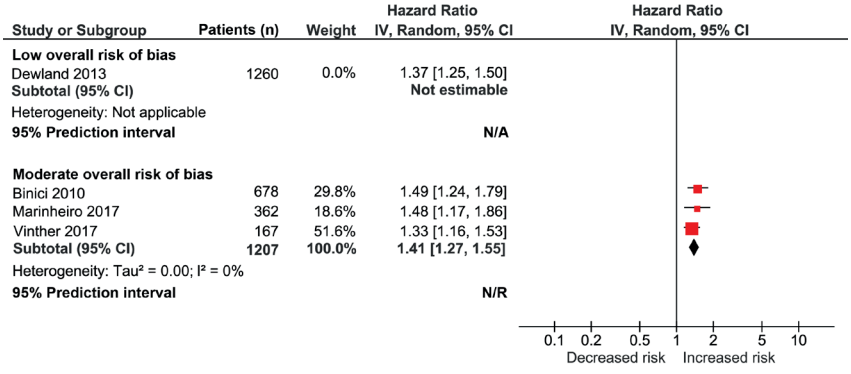
CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Frequent PACs' and 'Infrequent PACs' represent the number of participants that were grouped according the applied cut-off for 'frequent PACs' and 'infrequent PACs' in their respective study.

Supplementary Figure 15. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome all-cause mortality based on continuous PAC-count on baseline Holter, grouped according to scale applied to continuous PAC-count.



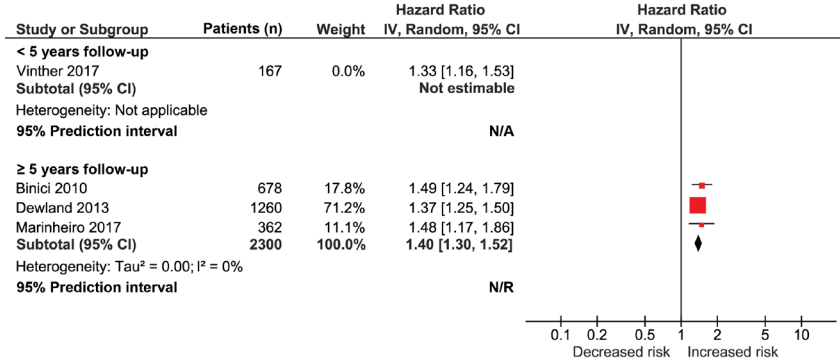
CI = confidence interval; IV = inverse variance; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Patients (n)' represent total cohort size.

Supplementary Figure 16. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome all-cause mortality based on continuous PAC-count on baseline Holter, grouped according to overall risk of bias.



CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Patients (n)' represent total cohort size.

Supplementary Figure 17. Subgroup meta-analysis of studies that reported unadjusted hazard ratios for the outcome all-cause mortality based on continuous PAC-count on baseline Holter, grouped according to follow-up duration.



CI = confidence interval; IV = inverse variance; N/A = not applicable; N/R = not relevant; PACs = premature atrial contractions. Totals under 'Patients (n)' represent total cohort size.

REFERENCES

1 Conen D, Adam M, Roche F, Barthelemy JC, Felber Dietrich D, Imboden M. Premature atrial contractions in the general population: frequency and risk factors. Circulation 2012;126:2302–8.

2 Marcus GM, Dewland TA. Premature atrial contractions: a wolf in sheep's clothing? J Am Coll Cardiol 2015;66:242–4.

3 Gaita F, Castagno D. Do supraventricular premature beats identify patients at high risk for atrial fibrillation? J Cardiovasc Med 2017;18:e117–20.

4 Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial fibrillation and mechanisms of stroke: time for a new model. Stroke 2016;47:895–900.

5 Glotzer TV, Ziegler PD. Cryptogenic stroke: is silent atrial fibrillation the culprit? Heart Rhythm 2015;12:234–41.

6 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). Europace 2010;12:1360–420.

7 January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JCJr et al. AHA/ACC/HRS Guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol 2014;64:e1–76. 2014;

8 Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS Med 2013;10:e1001380.

9 Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009;62:1006–12.

10 Hayden JA, Tougas ME, Riley R, Iles R, Pincus T. Individual recovery expectations and prognosis of outcomes in nonspecific low back pain: prognostic factor exemplar review. Cochrane Database of Systematic Reviews 2014, Issue 9. Art. No.: CD011284. doi:10.1002/14651858.CD011284.

11 PROSPERO International Prospective Register of Systematic Reviews: National Institute for Health Research. https://www.crd.york.ac.uk/prospero/ (27 July 2018, date last accessed).

12 Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158:280–6.

13 Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011;342:d549.

14 Partlett C, Riley RD. Random effects meta-analysis: coverage performance of 95% confidence and prediction intervals following REML estimation. Stat Med 2017;36:301–17.

15 Higgins JPT, Green S (eds). Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0 [updated March 2011]. London: The Cochrane Collaboration, 2011. http://handbook.cochrane.org (27 July 2018, date last accessed).

16 IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 2016;6:e010247.

17 Gladstone DJ, Dorian P, Spring M, Panzov V, Mamdani M, Healey JS et al. Atrial premature beats predict atrial fibrillation in cryptogenic stroke: results from the EMBRACE trial. Stroke 2015;46:936–41.

18 Huang BT, Huang FY, Peng Y, Liao YB, Chen F, Xia TL et al. Relation of premature atrial complexes with stroke and death: Systematic review and meta-analysis. Clin Cardiol 2017;40:962–9.

19 Sejr MH, Riahi S, Larsen TB, Nielsen JC, Nielsen PB. Premature atrial complexes in an ischemic stroke population and risk of recurrent stroke: a systematic review. Expert Rev Cardiovasc Ther 2017;15:447–55.

20 Kumarathurai P, Mouridsen MR, Mattsson N, Larsen BS, Nielsen O, Gerds TA et al. Atrial ectopy and N-terminal pro-B-type natriuretic peptide as predictors of atrial fibrillation: a population-based cohort study. Europace 2017;19:364–70.

21 Larsen BS, Kumarathurai P, Falkenberg J, Nielsen OW, Sajadieh A. Excessive atrial ectopy and short atrial runs increase the risk of stroke beyond incident atrial fibrillation. J Am Coll Cardiol 2015;66:232–41.

22 Larsen BS, Kumarathurai P, Nielsen OW, Sajadieh A. The circadian variation of premature atrial contractions. Europace 2016;18:1573–80.

23 Mattsson N, Kumarathurai P, Larsen BS, Nielsen OW, Sajadieh A. Mild hypokalemia and supraventricular ectopy increases the risk of stroke in community-dwelling subjects. Stroke 2017;48:537–43.

24 Chun KJ, Hwang JK, Park SJ, On YK, Kim JS, Park KM. Electrical PR interval variation predicts new occurrence of atrial fibrillation in patients with frequent premature atrial contractions. Medicine 2016;95:e3249.

25 Nguyen KT, Vittinghoff E, Dewland TA, Mandyam MC, Stein PK, Soliman EZ et al. Electrocardiographic predictors of incident atrial fibrillation. Am J Cardiol 2016;118:714–9.

26 Yamada S, Lin CY, Chang SL, Chao TF, Lin YJ, Lo LW et al. Risk of stroke in patients with short-run atrial tachyarrhythmia. Stroke 2017;48:3232–8.

27 O''Neal WT, Kamel H, Kleindorfer D, Judd SE, Howard G, Howard VJ et al. Premature atrial contractions on the screening electrocardiogram and risk of ischemic stroke: the reasons for geographic and racial differences in stroke study. Neuroepidemiology 2016;47:53–8.

28 O'Neal WT, Kamel H, Judd SE, Safford MM, Vaccarino V, Howard VJ et al. Usefulness of atrial premature complexes on routine electrocardiogram to determine the risk of atrial fibrillation (from the REGARDS Study). Am J Cardiol 2017;120:782–5.

29 Dewland TA, Vittinghoff E, Mandyam MC, Heckbert SR, Siscovick DS, Stein PK et al. Atrial ectopy as a predictor of incident atrial fibrillation: a cohort study. Ann Intern Med 2013;159:721–8.

30 Nguyen KT, Vittinghoff E, Dewland TA, Dukes JW, Soliman EZ, Stein PK et al. Ectopy on a single 12-lead ECG, incident cardiac myopathy, and death in the community. J Am Heart Assoc 2017;6:e006028.

31 Blanch Gracia P, Freixa Pamias R, Codinach Huix P, Martin Baranera M, Armario Garcia P. Predictores electrocardiográficos y ecocardiográficos de fibrilación auricular en pacientes hipertensos. Hipertension y Riesgo Vascular 2013;30:12–7.

32 Inohara T, Kohsaka S, Okamura T, Watanabe M, Nakamura Y, Higashiyama A et al. Long-term outcome of healthy participants with atrial premature complex: a 15-year follow-up of the NIPPON DATA 90 cohort. PLoS One 2013;8:e80853.

33 Murakoshi N, Xu D, Sairenchi T, Igarashi M, Irie F, Tomizawa T et al. Prognostic impact of supraventricular premature complexes in community-based health checkups: the Ibaraki Prefectural Health Study. Eur Heart J 2015;36:170–8.

34 Qureshi W, Shah AJ, Salahuddin T, Soliman EZ. Long-term mortality risk in individuals with atrial or ventricular premature complexes (results from the Third National Health and Nutrition Examination Survey). Am J Cardiol 2014;114:59–64.
35 Perez MV, Dewey FE, Marcus R, Ashley EA, Al-Ahmad AA, Wang PJ et al. Electrocardiographic predictors of atrial fibrillation. Am Heart J 2009;158:622–8.
36 Acharya T, Tringali S, Bhullar M, Nalbandyan M, Ilineni VK, Carbajal E et al. Frequent atrial premature complexes and their association with risk of atrial fibrillation. Am J Cardiol 2015;116:1852–7.
37 Cabrera S, Valles E, Benito B, Alcalde O, Jimenez J, Fan R et al. Simple predictors for new onset atrial fibrillation. Int J Cardiol 2016;221:515–20.
38 Chong BH, Pong V, Lam KF, Liu S, Zuo ML, Lau YF et al. Frequent premature atrial complexes predict new occurrence of atrial fibrillation and adverse cardiovascular events. Europace 2012;14:942–7.
39 Chun KJ, Hwang JK, Choi SR, Park SJ, On YK, Kim JS et al. Electrocardiogram PR interval is a surrogate marker to predict new occurrence of atrial fibrillation in patients with frequent premature atrial contractions. J Korean Med Sci 2016;31:519–24.
Google ScholarCrossref Find in my library PubMedWorldCat
increased risk in men with high frequency of atrial ectopic beats. Stroke 2000;31:2925–9.
41 Johnson LSB, Juhlin T, Juul-Möller S, Hedblad B, Nilsson PM, Engström G. A prospective study of supraventricular activity and incidence of atrial fibrillation. Heart Rhythm 2015;12:1898–904.
42 Kochhäuser S, Dechering DG, Dittrich R, Reinke F, Ritter MA, Ramtin S et al. Supraventricular premature beats and short atrial runs predict atrial fibrillation in continuously monitored patients with cryptogenic stroke. Stroke 2014;45:884–6.
43 Lin CY, Lin YJ, Chen YY, Chang SL, Lo LW, Chao TF et al. Prognostic significance of premature atrial complexes burden in prediction of long-term outcome. J Am Heart Assoc 2015;4:e002192.
44 Nortamo S, Kentta TV, Ukkola O, Huikuri HV, Perkiomaki JS. Supraventricular premature beats and risk of new-onset atrial fibrillation in coronary artery disease. J Cardiovasc Electrophysiol 2017;28:1269–74.
45 Pinho J, Braga CG, Rocha S, Santos AF, Gomes A, Cabreiro A et al. Atrial ectopic activity in cryptogenic ischemic stroke and TIA: a risk factor for recurrence. J Stroke Cerebrovasc Dis 2015;24:507–10.
46 Suzuki S, Sagara K, Otsuka T, Kano H, Matsuno S, Takai H et al. Usefulness of frequent supraventricular extrasystoles and a high CHADS2 score to predict first-time appearance of atrial fibrillation. Am J Cardiol 2013;111:1602–7.
47 Thijs VN, Brachmann J, Morillo CA, Passman RS, Sanna T, Bernstein RA et al. Predictors for atrial fibrillation detection after cryptogenic stroke: results from CRYSTAL AF. Neurology 2016;86:261–9.
8 Vinther KH, Tveskov C, Moller S, Auscher S, Osmanagic A, Egstrup K. Excessive premature atrial complexes and the risk of recurrent stroke or death in an ischemic stroke population. J Stroke Cerebrovasc Dis 2017;26:1163–70.
49 Wallmann D, Tuller D, Kucher N, Fuhrer J, Arnold M, Delacretaz E. Frequent atrial premature contractions as a surrogate marker for paroxysmal atrial fibrillation in patients with acute ischaemic stroke. Heart 2003;89:1247–8.

50 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.

51 Yamada T, Fukunami M, Shimonagata T, Kumagai K, Ogita H, Asano Y et al. Prediction of paroxysmal atrial fibrillation in patients with congestive heart failure: a prospective study. J Am Coll Cardiol 2000;35:405–13.

52 Yodogawa K, Seino Y, Ohara T, Hayashi M, Miyauchi Y, Katoh T et al. Prediction of atrial fibrillation after ischemic stroke using P-wave signal averaged electrocardiography. J Cardiol 2013;61:49–52.

53 Marinheiro R, Parreira L, Amador P, Sa C, Duarte T, Caria R. Excessive atrial ectopic activity as an independent risk factor for ischemic stroke. Int J Cardiol 2017;249:226–30.

54 Weber-Krüger M, Lutz C, Zapf A, Stahrenberg R, Seegers J, Witzenhausen J et al. Relevance of supraventricular runs detected after cerebral ischemia. Neurology 2017;89:1545–52.

55 Binici Z, Intzilakis T, Nielsen OW, Kober L, Sajadieh A. Excessive supraventricular ectopic activity and increased risk of atrial fibrillation and stroke. Circulation 2010;121:1904–11.

56 Vinther KH, Tveskov C, Moller S, Rosen T, Auscher S, Osmanagic A et al. Prevalence and prognostic significance of runs of premature atrial complexes in ischemic stroke patients. J Stroke Cerebrovasc Dis 2016;25:2338–43.

57 Folkeringa RJ, Hartgers J, Tieleman RG, Gorgels AP, Dassen WR, Crijns HJ. Atrial extrasystoles after exercise predict atrial fibrillation in patients with left ventricular hypertrophy. Heart 2006;92:545–6.

58 Raman D, Kaffashi F, Lui LY, Sauer WH, Redline S, Stone P et al. Polysomnographic heart rate variability indices and atrial ectopy associated with incident atrial fibrillation risk in older community-dwelling men. JACC Clin Electrophysiol 2017;3:451–60.

59 Chen SA, Hsieh MH, Tai CT, Tsai CF, Prakash VS, Yu WC et al. Initiation of atrial fibrillation by ectopic beats originating from the pulmonary veins: electrophysiological characteristics, pharmacological responses, and effects of radiofrequency ablation. Circulation 1999;100:1879–86.

60 Vincenti A, Brambilla R, Fumagalli MG, Merola R, Pedretti S. Onset mechanism of paroxysmal atrial fibrillation detected by ambulatory Holter monitoring. Europace 2006;8:204–10.

61 Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M et al. The effect of english-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care 2012;28:138–44.

5

Longitudinal association of premature atrial contractions with atrial fibrillation and brain ischemia in people with type 2 diabetes: The Hoorn Diabetes Care System cohort

Peter P. Harms*, Jelle C.L. Himmelreich*, Marieke T. Blom, Joline W.J. Beulens, Giel Nijpels, Petra Elders, Wim A.M. Lucassen

* Shared first author, these authors contributed equally to this work

Submitted

ABSTRACT

Background: Premature atrial contractions (PACs) on electrocardiogram (ECG) are potential markers for imminent onset of both atrial fibrillation (AF) and brain ischemia. We investigated the association of PACs with incident AF and brain ischemia separately, and of incident AF with brain ischemia in people with type 2 diabetes without pre-existing AF or cerebrovascular disease.

Methods: A prospective longitudinal study of 12,242 people with type 2 diabetes without known AF or cerebrovascular disease from the Hoorn Diabetes Care System cohort. Annually repeated measurements (1998-2018) included cardiovascular risk factors, over 85,000 ECGs, and self-reported cardiovascular events. We assessed the association of PACs with incident AF and brain ischemia events (transient ischemic attack and ischemic stroke) and of incident AF with brain ischemia events using time-dependent Cox-regression models for repeated measurements, adjusted for time-varying cardiovascular risk factors and medication use (Hazard Ratios [HR] with 95% confidence interval [CI]).

Results: At baseline, mean age of the study population was 62.2 ± 11.9 years. During a median follow-up of 7.0 (interquartile range [IQR]: 3.4-11.0) years, 1,031 (8.4%) of the participants had PACs at any study ECG, and 566 (4.6%) had incident AF at any of the median 6 (IQR: 3-10) annual ECG recordings. Brain ischemia events occurred in 517 (4.2%) people, of which 304 were transient ischemic attacks, and 213 ischemic strokes. After adjustment, PACs on any previous study ECG were associated with incident AF (HR: 1.96; 95% CI: 1.53-2.50), but not with overall brain ischemia events (HR: 1.09; 95% CI: 0.76-1.56), transient ischemic attack (HR: 0.91; 95% CI: 0.57-1.46) or ischemic stroke (HR: 1.50; 95% CI: 0.88-2.54). AF was not associated with brain ischemia events (HR: 0.95; 95% CI: 0.55-1.63).

Conclusions: In people with T2D without a history of AF or brain ischemia events, PACs (prevalent or incident) are associated with a two-fold increased risk of incident AF, and might warrant targeted screening for AF.

BACKGROUND

People with type 2 diabetes (T2D) have an approximately 35% higher risk of developing atrial fibrillation (AF) than people without T2D.[1] Whereas T2D confers an up to two-fold higher risk of stroke because the involved metabolic alterations catalyse vascular arteriosclerosis and thrombogenesis,[2] AF further increases the risk of brain ischemia events.[3,4] When AF is diagnosed, prescription of anti-coagulant medication reduces the excess risk of brain ischemia events by a fifth to two-thirds.[5] Therefore, early detection of AF in people with T2D is likely to decrease the burden of stroke in this high-risk group.[3]

Periodic electrocardiographic (ECG) screening for cardiovascular risk assessment in people with T2D is currently recommended only for those with concomitant hypertension or suspected cardiovascular disease, and screening for AF is only recommended in people aged 65 years and older.[6] However, ECG markers might enable early recognition of people at increased risk of developing AF and facilitate targeted screening.[7] Potential ECG markers for AF are premature atrial contractions (PACs). Until recently, these ectopic beats were regarded as benign and clinically insignificant findings.[8] Recent prospective studies in the general population report that PACs are associated with an up to five-fold higher risk of AF, and a one-and-a-half-fold higher risk of ischemic stroke.[9-14] Additionally, a meta-analysis demonstrated that frequent PACs are associated with incident AF, brain ischemia or all-cause mortality,[15] and the association was more pronounced in sub-populations at higher risk of cardiovascular disease. The association of PACs with brain ischemia could be both independent from and (partially) mediated by AF.

In people with T2D, PACs and AF are common findings,[16] and the association of several other ECG abnormalities with cardiovascular events did not differ across subgroups by age, hypertension or estimated cardiovascular risk.[17] Moreover, studies that analysed repeated ECG recordings during follow-up reported stronger associations with the outcomes, than studies that only analysed baseline ECGs,[18-20] indicating that repeated ECG recordings could provide additional insight into the value of PACs in assessing risk of AF. However, no studies investigated the association of PACs with AF or brain ischemia events in people with T2D, or have considered incident PACs after the baseline measurement.

Therefore, we aimed to investigate the association of PACs with incident AF and brain ischemia separately and of incident AF with brain ischemia in people with T2D without pre-existing AF or cerebrovascular disease.

METHODS

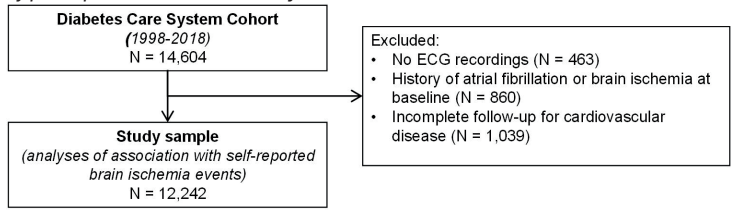
Design and population

The Hoorn Diabetes Care System (DCS) cohort consists of people with T2D from the West-Friesland region in The Netherlands. Details of the cohort have been described previously.[21] Initiated in 1998 as a prospective dynamic cohort of people with T2D, General Practitioners (GPs) could refer their T2D patients to the DCS center for annual follow-up measurements and treatment. From 2010, all people with newly diagnosed T2D in the West-Friesland region were referred to the DCS center. In 2018, the DCS cohort consisted of 14,604 people with T2D, approximately 95% of all people with T2D from the catchment region. At the DCS center, trained research personnel annually examined participants according to standard operating procedures, including anthropometrics, blood pressure, blood samples, an ECG recording, and documentation of medication use and self-reported cardiovascular events.

Study sample

We used the annual examination data from the period 1998-2018. Of the 14,604 people in the DCS cohort with at least one annual examination, we excluded 463 because they did not have any ECG recorded. We also excluded 860 people with a history of AF or ischemic cerebrovascular disease at baseline, defined as AF on ECG, or self-reported transient ischemic attack (TIA) or ischemic stroke at the participant's first ECG recording after entry into the DCS cohort. We excluded a further 1,039 people because they had incomplete follow-up for cardiovascular events, including cerebrovascular events. The remaining 12,242 (83.8%) participants were included in the analyses (Figure 1).

Figure 1. *Study participant inclusion/exclusion flowchart.*



ECG, electrocardiogram

Premature atrial contractions and atrial fibrillation

During the annual examinations, trained personnel recorded a standard 10-second 12-lead resting ECG. One trained examiner subsequently evaluated and coded all ECGs according to the Minnesota Classification (MC) system.[22] In a random sample (n=60), the coding was compared with the coding of two independent cardiologist, showing a

specific agreement for atrial arrhythmic abnormalities of between 0.87 (95% CI, 0.77 to 0.93) and 0.97 (95% CI, 0.95 to 0.99).[23] In 2020, a consistency over time analysis in which the examiner blindly recoded random ECGs from both 2002 (n=60) and 2016 (n=60), showed a specific agreement for atrial arrhythmic abnormalities between 0.97 (95% CI, 0.92 to 0.99) and 0.99 (95% CI, 0.97 to 1.00) (unpublished results).

We defined PACs as MC codes 8-1-1 and 8-4-2 (atrial or junctional premature beats in 10% or more of recorded complexes; generally translating to ≥1 PACs on the 10-second 12-lead ECG), and AF MC codes 8-3-1 to 8-3-4 (persistent or intermittent atrial fibrillation or atrial flutter). The exact MC descriptions are given in Supplementary Table S1.

Brain ischemia events

At the annual DCS examinations, cardiovascular morbidity was assessed through self-report and was classified as TIA, stroke, myocardial infarction, angina pectoris, peripheral artery disease, heart failure, cardiac arrest and arrhythmia. The self-reported cardiovascular events were validated against the electronic patient registration of the regional hospital in a random sample of 453 participants. The sensitivity and specificity were 86% and 90%, and positive and negative predictive values were 90% and 87%, respectively.[21]

We defined brain ischemia as TIA or ischemic stroke and used the self-reported TIA and stroke events to record brain ischemia events. Stroke was considered ischemic if subsequently anti-coagulants were prescribed, defined as (new) use of anti-thrombotic medication (ATC codes B01) reported at the first follow-up examination within one and a half year after the stroke date.

Deceased participants were registered every six months via the national population registry. The cause of death was determined from GP and regional hospital records and coded according to the International Classification of Diseases, Injuries and Causes of Death, ninth revision (ICD-9). For death due to TIA and ischemic stroke, we used ICD-9 codes 435 and 434, respectively.

Covariables

We recorded sex, date of birth, date of T2D diagnosis and educational level at entry into the DCS cohort through self-report. Highest achieved education was classified as either low (primary), middle (secondary), or high (tertiary). All other variables were assessed annually.

Smoking behaviour was classified as: never, former smoker, and current smoker. Body mass index (BMI) was calculated by dividing body weight by height squared. Systolic and diastolic blood pressure was measured using an automatic oscillometric digital blood pressure device (Welch Allyn ProBP 3400, Skaneateles Falls, New York, USA). We determined fasting glucose level, glycated haemoglobin (HbA1c), total cholesterol, high-density lipoprotein (HDL) cholesterol, triglyceride, creatinine level, and urinary albumin and creatinine from overnight fasting blood or urine samples (Cobas c501 analyzer, Roche Diagnostics, Mannheim, Germany). Cholesterol ratio was determined by dividing the total cholesterol by HDL-cholesterol. We calculated low-density lipoprotein (LDL) cholesterol using the Friedewald formula.[24] The estimated glomerular filtration rate (eGFR) was calculated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation,[25] and the urinary albumin creatinine ratio (UACR) by dividing albumin in mg/l by creatinine in mmol/l.

We obtained information on medication use by inspecting dispensing labels, registering the name of the drug, prescribed quantity, dosage, and the Anatomical Therapeutic Chemical classification code. The use of glucose-lowering medication (A10A or A10B codes) was classified as: no medication, oral medication only, insulin use only, and combined oral and insulin use. Anti-thrombotic (B01 codes), anti-hypertensive (C02, C03A, C03B, C03E, C03DA, C07, C08, or C09 codes) and lipid-modifying (C10 codes) medication use were classified as: no or yes.

Hypertension was defined as elevated blood pressure (systolic>140 mmHg or diastolic>90 mmHg) and/or anti-hypertensive medication use, dyslipidaemia as elevated LDL-cholesterol (>2.4 mmol/l) and/or lipid lowering medication use. We categorized eGFR as: normal or high (>90 ml/min), mildly decreased (60-90 ml/min), moderately decreased (30-60 ml/min) or severely decreased (<30 ml/min). We categorized albuminuria as: normal to mild (<3 mg/mmol), moderate (3-30 mg/mmol) or severe (>30 mg/mmol).

Statistical analyses

Baseline was defined as the first study visit with available 12-lead ECG. We calculated baseline characteristics for the total study sample and for participants with and without prevalent PACs at baseline and with and without PACs during any study visit, reporting means with standard deviations (SD), medians with interquartile range (IQR) or percentages as appropriate. Differences in baseline characteristics were compared using an unpaired t-test, Wilcoxon rank-sum (Mann-Whitney) test, or Chi-square test for normally distributed, skewed, and categorical variables.

We calculated the prevalence and incidence of PACs at baseline and during the entire follow-up, respectively, and the incidence of AF stratified by prevalent or incident PACs at baseline and any examination and reported the median follow-up time, the median number of ECG recordings, and the number of observed TIA and/or ischemic stroke events. In addition, we calculated incidence rates and plotted incidence curves for AF and brain ischemia events stratified by PACs or AF at any study visit.

During follow-up, 3,193 (3.6%) ECG measurements were missing. In case of missing ECG data, we imputed PAC and AF status with the last value carried forward method until the next non-missing value. The missing measurements for the covariables were calculated stratified by consecutive annual examination (Supplementary Figure S1). The total proportion of missing values for all covariables over the 20 year follow-up period was only 1.7% (34,385/2,041,710), therefore we excluded the missing values pair-wise.

We used time-dependent Cox-regression models for repeated measurements to evaluate the association of PACs with incident AF and brain ischemia events (TIA and ischemic stroke) and of incident AF with brain ischemia events, computing hazard ratios (HRs) with 95% confidence intervals (95% CI). To adjust for confounding, we built models using a step-wise approach: model 1, unadjusted; model 2, adjusted for age and sex; model 3, additionally adjusted for smoking behavior, BMI, systolic blood pressure, HbA1c, and TC/HDL–C ratio; model 4, additionally adjusted for education, eGFR, UACR, glucose-lowering medication use, lipid-modifying medication use, and antihypertensive medication use.

PACs and AF were modelled as irreversible, meaning that a participant's status remained positive once a PAC or AF was detected on an annual ECG, even when PAC or AF was not detected at subsequent annual visit ECGs. Covariables were modelled as time-varying (age, smoking behavior, BMI, systolic blood pressure, HbA1c, TC/HDL–C ratio, eGFR, UACR, antihypertensive medication, glucose-lowering medication and lipid-modifying medication), or as time-constant (sex, and education).

Follow-up duration was defined as the time from baseline to first AF or brain ischemia event (depending on the outcome of the analysis), last contact date (in case of loss to follow-up) or date of death (if participants died of a cause other than brain ischemia), whichever occurred first.

In all analyses, significance was assessed at the $p<0.05$ (two-sided) or 95% CI level. Analyses were performed using SPSS version 26 (IBM corporation, New York, USA),[26]

and R (studio) version 4.0.3 (R foundation for statistical computing, Vienna, Austria)[27] with the R packages haven (2.3.1),[28] epiR (2.0.19),[29] survival (3.2-7),[30] survminer (0.4.8),[31] and ggplot2 (3.3.2).[32]

RESULTS

Baseline characteristics

At baseline, the mean age of the study population was 62.2 ±11.9 years, median T2D duration was 0.6 (IQR: 0.2-3.2) years, and 53.1% was male (Table 1). Compared to participants without PACs, participants with PACs were older, more frequently male, former smoker and lower educated, had higher blood pressure, lower eGFR, higher UACR, and used more insulin, antithrombotic, antihypertensive and lipid-lowering medication. These differences were more prominent in people with PACs at baseline compared to people with PACs at any examination during follow-up. Additionally, participants with PACs at baseline and/or during follow-up, had higher incidence of AF during follow-up.

Follow-up for PACs, AF, and brain ischemia events

During a median follow-up of 7.0 (IQR: 3.4-11.0) years, 1,031 (8.4%) of the participants had PACs at any study ECG, and 566 (4.6%) had incident AF at any of the median 6 (IQR: 3-10) annual ECG recordings. Brain ischemia events occurred in 517 (4.2%) people, of which 304 were TIA, and 213 ischemic strokes.

The crude incidence rate of AF per 1,000 person-years was more than three-fold higher in participants with PACs (20.4; 95% CI: 16.8-24.5) on any previous study ECG, compared to participants without PACs (6.1; 95% CI: 5.6-6.6)) (Supplementary Table S2). The crude incidence rate of brain ischemia events (TIA or ischemic stroke) per 1,000 person-years was a marginally significant one-and-a-half-fold higher in participants with PACs (8.0; 95% CI: 5.9-10.7)), compared to participants without PACs (5.3; 95% CI: 4.8-5.8)), and a statistically insignificant one-and-a-half-fold higher in participants with AF (7.9; 95% CI: 4.9-12.0)), compared to participants without AF (5.4; 95% CI: 4.9-5.9)). The cumulative incidence curves provided similar results for the difference between participants with and without PACs, albeit somewhat attenuated. There was no significant difference in cumulative incidence of brain ischemia events between participants with and without AF (Figure 2).

Table 1. *Characteristics of the study population at baseline, stratified by prevalence and incidence of PACs at baseline, or at any examination.*

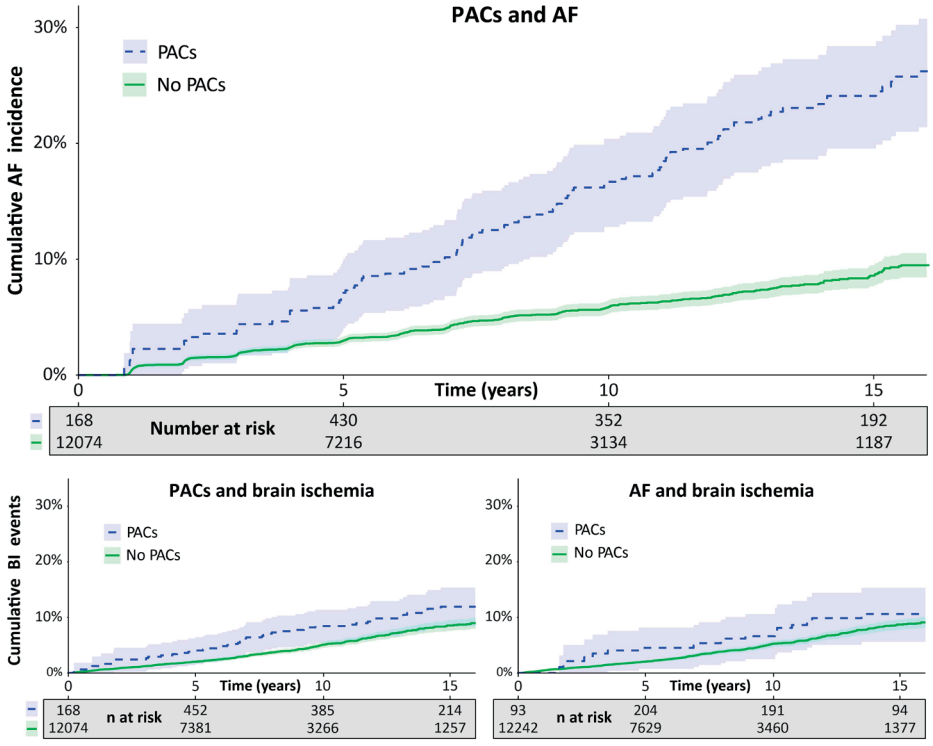
Characteristic	Total	Prevalence at baseline		At any examination	
		no PACs	PACs	no PACs	PACs
Number	12,242[†]	12,074	168	11,211	1,031
Incident AF during follow-up	4.6%	4.6%	9.5%	4.0%	11.3%
Age (years)	62.2 ±11.9	62.1 ±11.9	70.6 ±10.3	61.9 ±12.0	66.1 ±9.8
Men (%)	53.1%	53.0%	60.7%	52.9%	56.1%
T2D duration (years)	0.6 (0.2-3.2)	0.6 (0.2-3.2)	0.9 (0.2-3.3)	0.6 (0.2-3.2)	0.6 (0.2-3.1)
Educational level (%)					
low	43.6%	43.5%	50.0%	42.9%	50.7%
middle	40.7%	40.7%	43.0%	41.3%	34.7%
high	15.7%	15.8%	7.0%	15.8%	14.6%
Smoking behavior (%)					
never	39.3%	39.3%	34.4%	39.3%	38.6%
former	39.1%	39.0%	47.5%	38.6%	44.0%
current	21.7%	21.7%	18.1%	22.0%	17.4%
BMI (kg/m^2)	30.3 ±5.5	30.3 ±5.5	29.6 ±5.0	30.3 ±5.5	29.8 ±5.0
SBP (mmHg)	142.9 ±20.9	142.8 ±20.9	149.7 ±23.4	142.5 ±20.9	147.0 ±20.6
DBP (mmHg)	80.9 ±10.0	80.9 ±10.0	77.7 ±10.5	80.9 ±10.0	81.0 ±10.6
HbA1c (%)	6.7 (6.2-7.7)	6.7 (6.2-7.7)	6.6 (6.1-7.5)	6.7 (6.2-7.7)	6.7 (6.2-7.6)
HbA1c (mmol/mol)	50.0 (44.0-60.7)	50.0 (44.0-60.7)	48.6 (43.2-58.2)	50.0 (44.0-61.0)	49.7 (44.0-59.6)
Fasting glucose (mmol/l)	7.9 (7.0-9.2)	7.9 (7.0-9.2)	7.5 (6.8-8.4)	7.9 (7.0-9.2)	7.8 (6.8-9.0)
Total cholesterol (mmol/l)	5.1 ±1.2	5.1 ±1.2	4.9 ±1.1	5.1 ±1.2	5.1 ±1.1
LDL cholesterol (mmol/l)	3.0 ±1.0	3.0 ±1.0	2.8 ±0.9	3.0 ±1.0	3.1 ±1.0
HDL cholesterol(mmol/l)	1.2 ±0.3	1.2 ±0.3	1.3 ±0.4	1.2 ±0.3	1.3 ±0.4
TC/HDL ratio	4.2 (3.4-5.2)	4.2 (3.4-5.2)	3.9 (3.2-4.7)	4.2 (3.4-5.2)	4.1 (3.3-5.1)

Table 1. *Characteristics of the study population at baseline, stratified by prevalence and incidence of PACs at baseline, or at any examination. (continued)*

Characteristic	Total	Prevalence at baseline		At any examination	
		no PACs	PACs	no PACs	PACs
Triglycerides (mmol/l)	1.6 (1.2-2.3)	1.6 (1.2-2.3)	1.4 (1.1-1.9)	1.6 (1.2-2.3)	1.6 (1.1-2.1)
eGFR (ml/min)	80.2 ±18.6	80.4 ±18.6	69.9 ±18.9	80.6 ±18.7	76.1 ±16.9
UACR (mg/mmol)	0.6 (0.4-1.4)	0.6 (0.4-1.4)	0.7 (0.4-2.1)	0.6 (0.4-1.4)	0.7 (0.4-1.5)
Glucose-lowering medication use (%)					
no medication	30.7%	30.7%	31.0%	30.8%	30.4%
oral only	59.8%	59.9%	54.2%	59.7%	60.9%
insulin only	4.8%	4.7%	8.9%	4.8%	4.2%
oral & insulin	4.7%	4.7%	6.0%	4.7%	4.6%
Anti-thrombotic medication (%)	22.5%	22.3%	36.9%	22.2%	25.6%
Anti-hypertensive medication (%)	55.6%	55.4%	69.6%	55.3%	58.7%
Lipid-lowering medication (%)	39.7%	39.5%	49.4%	39.8%	38.5%
Hypertension (%)	74.2%	74.1%	85.1%	73.6%	80.9%
Dyslipidemia (%)	91.9%	91.8%	92.6%	91.8%	92.8%
eGFR categories (%)					
normal or high (>90)	31.7%	32.0%	14.1%	32.8%	19.7%
mildly decreased (60-90)	54.8%	54.7%	58.9%	53.8%	65.2%
moderately decreased (30-60)	13.0%	12.8%	25.8%	12.8%	14.5%
severely decreased (<30)	0.5%	0.5%	1.2%	0.5%	0.6%
Albuminuria					
normal to mild (UACR <3)	86.2%	86.3%	80.5%	86.2%	86.4%
moderate (UACR 3-30)	11.8%	11.7%	15.1%	11.8%	11.3%
severe (UACR >30)	2.1%	2.0%	4.4%	2.0%	2.3%

AF, atrial fibrillation; BMI, body mass index; CVD, cardiovascular disease; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; HbA1c, hemoglobin A1c; HDL, high density lipoprotein; LDL, low density lipoprotein; PACs, premature atrial contractions; SBP, systolic blood pressure; TC, total cholesterol; UACR, urinary albumin creatinine ratio. Data are presented as mean ± SD, median (interquartile range), or percentage. Baseline was defined as the first annual examination with an ECG recording after entry into the DCS cohort. [†] Approximately 90% of the study population had European ancestry.

Figure 2. *Cumulative incidence of AF and brain ischemia, stratified by presence of PACs or AF on study ECG with confidence intervals (shaded areas).*

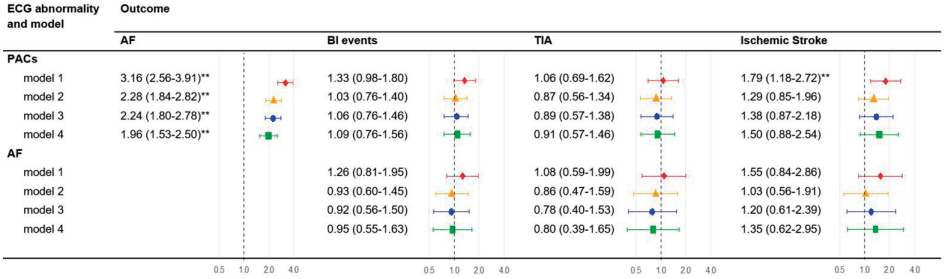


AF, atrial fibrillation; BI, brain ischemia; PACs, premature atrial contractions.
Note: The curves are based on instantaneous hazard functions that change over time, and therefore in time-dependent analyses depict the cumulative events for (hypothetical) participants with an ECG abnormality status that is time-constant over the whole follow-up period.

Association between PACs, AF and brain ischemia

After adjustment for all covariables in model 4, PACs on any previous study ECG were associated with incident AF (HR: 1.96; 95%CI: 1.53-2.50), but not with brain ischemia events (HR: 1.09; 95%CI: 0.76-1.56), TIA (HR: 0.91; 95%CI: 0.57-1.46) or ischemic stroke (HR: 1.50; 95%CI: 0.88-2.54) (Figure 3). However, the HRs for TIA of just below one were consistently of the opposite direction and lower in magnitude compared to the HRs for ischemic stroke of approximately one-and-a-half. Lastly, AF on ECG after baseline was not associated with brain ischemia events (HR: 0.95; 95%CI: 0.55-1.63), TIA (HR: 0.80; 95%CI: 0.39-1.65) or ischemic stroke (HR: 1.35; 95%CI: 0.62-2.95).

Figure 3. *Hazard ratios with 95% confidence intervals for atrial fibrillation and brain ischemia events from time-dependent Cox-regression models.*



AF, atrial fibrillation; BI, brain ischemia; PACs, premature atrial contractions; TIA, transient ischemic attack.
Model 1, unadjusted.
Model 2, adjusted for age and sex.
Model 3, adjusted for age, sex, smoking behavior, BMI, systolic blood pressure, HbA1c, and TC/HDL–C ratio.
Model 4, adjusted for age, sex, smoking behavior, BMI, systolic blood pressure, HbA1c, and TC/HDL–C ratio, education, eGFR, UACR, glucose-lowering medication, lipid-modifying medication, and antihypertensive medication use.
* significant at the p<0.05 level (two-sided).
** significant at the p<0.01 level (two-sided).

DISCUSSION

Principal findings

This study showed an approximately two-fold increased incidence of AF in people with T2D with PACs, compared to people with T2D without PACs. However, we observed no increased risk for brain ischemia events, TIA or ischemic stroke in people with T2D with PACs or AF.

Comparison to previous work

Kamel and colleagues (2016) proposed that both PACs and AF should be seen as signs of atrial cardiomyopathy, and that PACs may be independently related to clinical outcomes, while AF is an already clinically established (thrombogenic) variant.[33] In line with this novel framework, previous studies conducted in general populations reported associations of PACs with AF: a three-fold to five-fold increased risk of AF (over a 14-year follow-up period) in the Japanese IPHS cohort,[10] an almost two-fold increased risk of AF in the North-American REGARDS cohort,[12] and a roughly one-and-a-half-fold increased risk of AF in the North-American CHS and ARIC cohorts.[13] These risks are similar to the two-fold increased incidence of AF in people with T2D in our current study. This finding suggests that the association between PACs and AF is not modified by T2D, and somewhat contradicts the more pronounced association in other high-risk populations observed by a recent meta-analysis.[15] The higher risk found in the Japanesse IPHS cohort maybe explained by the different ancestry, the on average

twice as long follow-up or the six times lower overall incidence rate of AF in that study, possibly a result of not counting atrial flutter as AF. Our study is the first to confirm the association of PACs with AF in people with T2D.

Our study does not indisputably confirm previously reported associations of PACs with a roughly one-and-a-half-fold increased risk of ischemic stroke or stroke mortality in the IPHS, the ARIC, and the REGARDS cohorts.[9-11,14] Despite similar point estimates across our analyses models, the association between PACs and ischemic stroke was only significant in the unadjusted model, not in the adjusted models. This difference could result from our time-dependent cox regression analysis that more accurately adjusts for repeated measurements of both PACs and confounders, or because the power of the separate TIA and ischemic stroke analyses was too low. In the latter case, PACs might be associated with ischemic stroke both through AF and subsequent thromboembolism, and independently from AF via arteriosclerotic small vessel disease. We are the first to report on the association of PACs with incident TIA, in people without previous brain ischemia events. In our analysis, there was no association of PACs with TIAs.

We found no association of AF with brain ischemia events. A plausible explanation is that in the DCS cohort, incident AF on one of the annual study ECG recordings was commonly followed by initiation of adequate thrombosis prophylactic medication and (intensified) cardiovascular risk management according to the guidelines of the Dutch College of General Practitioners.[34]

PACs have been called signs of atrial cardiomyopathy that convey a risk of stroke both independent and through its association with AF.[33] Secondary analyses from previous studies hinted that AF could indeed be a mediator of the association between PACs and cardiovascular risk factors, stroke and mortality.[35,36] Therefore, a proper mediation analysis of the association between PACs, AF and brain ischemia events would be of interest. However, we were unable to perform a mediation analysis because methods for mediation analysis with time-dependent survival analysis have not yet been established.

Strengths and limitations

This study has a number of strengths. First, the DCS is a large unselected population-based cohort with real-world data of people with T2D in primary care. Second, the DCS dataset contains measurements of over 85,000 annual study visits from over 15 years of follow-up with detailed information and a high level of completeness. Third, time-varying analyses closely resemble clinical practice in which individuals' risk

profiles change over time, compared to classical time-to-event analyses that assess exposures and confounders only at baseline. Finally, PACs and AF were assessed with annual study ECGs, enabling an assessment of the potential for incident AF detection with periodic screening ECG in people with T2D.

A limitation of this study is the use of study ECGs for the detection of AF, which might have resulted in missing (paroxysmal) AF that was not present at the annual check-ups. This could potentially have led to underestimation of the association between PACs and AF if people with missed (paroxysmal) AF previously had PACs on a study ECG. By extension, this could also have led to underestimation of the association between PACs and brain ischemia, if antithrombotic medication was initiated in people with missed (paroxysmal) AF that had PACs on a study ECG. Another limitation is the use of mostly self-reported events. We could not distinguish ischemic from hemorrhagic stroke solely based and the self-reported cardiovascular morbidity. However, the self-reported cardiovascular morbidity registration was validated against hospital records in a sub-sample. In addition, we included only stroke events followed by antithrombotic medication use. Finally, the power of the separate TIA and ischemic stroke analyses is on the low side due to the limited number of events. This is reflected in the confidence intervals and increases the risk of a type II error. Therefore, the separate TIA and ischemic stroke analyses should be interpreted with caution.

Clinical relevance

There is a need for low-cost markers that help physicians decide which people with T2D would benefit from more stringent follow-up or targeted screening for risk of cardiovascular events. This need will not likely decrease in the foreseeable future because the prevalence of T2D is increasing worldwide.[37] Our findings indicate that PACs could constitute such a marker, and that encountering PACs on an ECG in people with T2D might warrant targeted or intensified screening for AF. Moreover, our findings support further research into the added value of ECG-aided clinical screening models for people with T2D.

CONCLUSIONS

In people with T2D without a history of AF or brain ischemia events, PACs (prevalent or incident) are associated with a two-fold increased risk of incident AF, and might warrant targeted screening for AF.

Acknowledgements: This study was been made possible by collaboration with the Diabetes Care System West-Friesland. The authors thank participants of this study as well as research staff of the Diabetes Care System West-Friesland.

Ethics declaration: This research was conducted with approval by the local medical ethical review board, and in accordance with the Helsinki Declaration on ethical principles for medical research involving human subjects.

Data availability: The data that support the findings of this study are available from the Diabetes Care System (DCS) cohort steering committee, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission of the DCS cohort steering committee.

Funding: This work was supported by the Netherlands Organization for Health Research and Development (ZonMw) [80-83910-98-13046], the Dutch Heart Foundation grant CVON2017-15 RESCUED, the EFSD's European Pilot Research Grants for Innovative Measurement of Diabetes Outcomes, a grant from the Stichting Stoffels-Hoornstra charity foundation, and Amsterdam University Medical Centers.

The study funders were not involved in the design of the study; the collection, analysis, and interpretation of data or writing the report.

Competing interests: The authors had complete autonomy in the design, conduct, and reporting of the manuscript. The authors declare no conflicts of interest.

Author contributions: Jelle C.L. Himmelreich, Peter P. Harms, Marieke T. Blom, Joline W.J. Beulens, Giel Nijpels, Wim A.M. Lucassen, Petra Elders contributed to conception, design, data acquisition and analyses, interpretation of the results, review, and editing. Jelle C.L. Himmelreich and Peter P. Harms drafted the first manuscript and all authors reviewed the subsequent versions. All authors agreed to be accountable for aspects pertaining integrity or accuracy, and approved the final version. Jelle C.L. Himmelreich and Peter P. Harms are the guarantors of this work, and as such, had full access to all the data in the study and take responsibility for the integrity and the accuracy of the data analysis.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLES

Supplementary Table S1. *The ECG abnormality categories used in this study based on aggregated Minnesota Classification codes.*

ECG abnormality category	Minnesota Classification	
	code	definition
Premature Atrial Contractions	8-1-1	Presence of frequent atrial or junctional premature beats (10% or more of recorded complexes).
	8-4-2	Supraventricular tachycardia intermittent. Three consecutive atrial or junctional premature beats occurring at a rate ≥ 100.
Atrial Fibrillation	8-3-1	Atrial fibrillation.
	8-3-2	Atrial flutter.
	8-3-3	Intermittent atrial fibrillation (code if 3 or more clear-cut, consecutive sinus beats are present in any lead).
	8-3-4	Intermittent atrial flutter (code of 3 or more clear-cut, consecutive sinus beats are present in any lead).

Note: During coding, no distinction was made between a few codes for two reasons:
1. Differentiation between them was deemed clinically irrelevant:
Atrial fibrillation or flutter (codes 8-3-1 and 8-3-2)).
2. Since distinction between persistent and intermittent varieties of an ECG abnormality is generally impossible with a 10 second ECG recording:
Atrial fibrillation (codes 8-3-1 and 8-3-3);
Atrial flutter (codes 8-3-2 and 8-3-4)

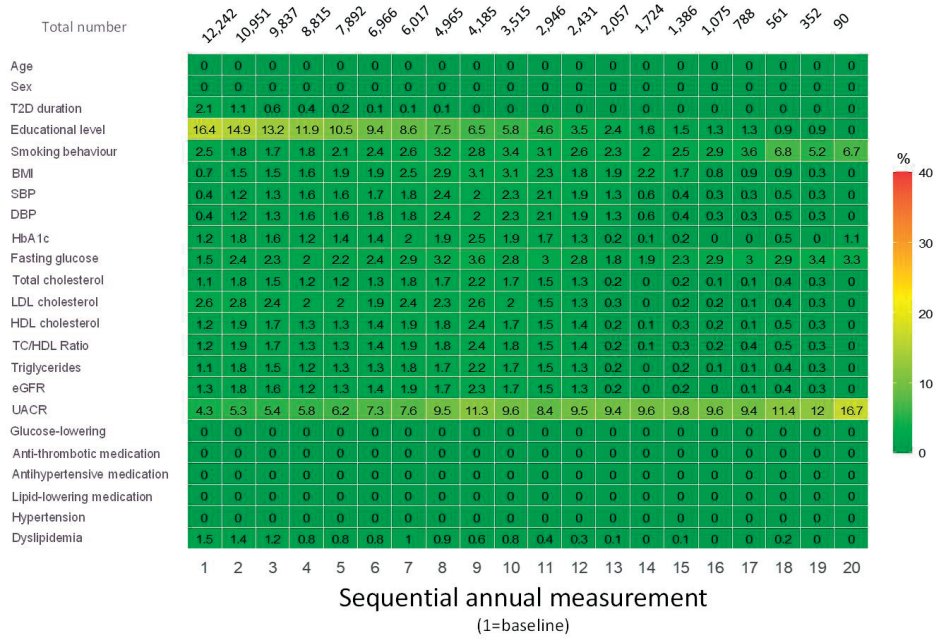
Supplementary Table S2. *Incidence rates with 95% confidence intervals for atrial fibrillation and brain ischemia events by ECG abnormality.*

ECG abnormality category	AF	BI events	Time (person years)	Incidence rate (per 1000 person years)
PAC				
no	528		86788.4	6.1 (5.6-6.6)
yes	111		5447.6	20.4 (16.8-24.5)
PAC				
no		470	88823.9	5.3 (4.8-5.8)
yes		47	5865.4	8.0 (5.9-10.7)
AF				
no		496	92014.6	5.4 (4.9-5.9)
yes		21	2674.7	7.9 (4.9-12.0)

AF, atrial fibrillation; BI, brain ischemia; PACs, premature atrial contractions; TIA, transient ischemic attack.

SUPPLEMENTARY FIGURES

Supplementary Figure S1. *Missing values of covariables at baseline and at subsequent follow-up measurements.*



BMI, body mass index; CVD, cardiovascular disease; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; HbA1c, haemoglobin A1c; HDL, high density lipoprotein; LDL, low density lipoprotein; SBP, systolic blood pressure; T2D, type 2 diabetes; TC, total cholesterol; UACR, urinary albumin creatinine ratio.
Data are presented as percentage. Baseline was defined as the first annual examination with an ECG recording after entry into the DCS cohort.

REFERENCES

1. Seyed Ahmadi S, Svensson AM, Pivodic A, Rosengren A, Lind M. Risk of atrial fibrillation in persons with type 2 diabetes and the excess risk in relation to glycaemic control and renal function: a Swedish cohort study. Cardiovasc Diabetol. 2020;19(1):9. https://doi.org/10.1186/s12933-019-0983-1.
2. Gregg EW, Cheng YJ, Srinivasan M, Lin J, Geiss LS, Albright AL, et al. Trends in cause-specific mortality among adults with and without diagnosed diabetes in the USA: an epidemiological analysis of linked national survey and vital statistics data. Lancet. 2018;391(10138):2430-40. https://doi.org/10.1016/S0140-6736(18)30314-3.
3. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). Eur Heart J. 2020. https://doi.org/10.1093/eurheartj/ehaa612.
4. 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(2):263-72. https://doi.org/10.1378/chest.09-1584.
5. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med. 2007;146(12):857-67. https://doi.org/10.7326/0003-4819-146-12-200706190-00007.
6. Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, Delgado V, et al. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J. 2020;41(2):255-323. https://doi.org/10.1093/eurheartj/ehz486.
7. Aizawa Y, Watanabe H, Okumura K. Electrocardiogram (ECG) for the Prediction of Incident Atrial Fibrillation: An Overview. J Atr Fibrillation. 2017;10(4):1724. https://doi.org/10.4022/jafib.1724.
8. Marcus GM, Dewland TA. Premature Atrial Contractions: A Wolf in Sheep's Clothing? J Am Coll Cardiol. 2015;66(3):242-4. https://doi.org/10.1016/j.jacc.2015.04.069.
9. Ofoma U, He F, Shaffer ML, Naccarelli GV, Liao D. Premature cardiac contractions and risk of incident ischemic stroke. J Am Heart Assoc. 2012;1(5):e002519. https://doi.org/10.1161/JAHA.112.002519.
10. Murakoshi N, Xu D, Sairenchi T, Igarashi M, Irie F, Tomizawa T, et al. Prognostic impact of supraventricular premature complexes in community-based health checkups: the Ibaraki Prefectural Health Study. Eur Heart J. 2015;36(3):170-8. https://doi.org/10.1093/eurheartj/ehu407.
11. O'Neal WT, Kamel H, Kleindorfer D, Judd SE, Howard G, Howard VJ, et al. Premature Atrial Contractions on the Screening Electrocardiogram and Risk of Ischemic Stroke: The Reasons for Geographic and Racial Differences in Stroke Study. Neuroepidemiology. 2016;47(1):53-8. https://doi.org/10.1159/000448619.
12. O'Neal WT, Kamel H, Judd SE, Safford MM, Vaccarino V, Howard VJ, et al. Usefulness of Atrial Premature Complexes on Routine Electrocardiogram to Determine the Risk of Atrial Fibrillation (from the REGARDS Study). Am J Cardiol. 2017;120(5):782-5. https://doi.org/10.1016/j.amjcard.2017.06.007.

13. Nguyen KT, Vittinghoff E, Dewland TA, Dukes JW, Soliman EZ, Stein PK, et al. Ectopy on a Single 12-Lead ECG, Incident Cardiac Myopathy, and Death in the Community. J Am Heart Assoc. 2017;6(8). https://doi.org/10.1161/JAHA.117.006028.
14. Huang BT, Huang FY, Peng Y, Liao YB, Chen F, Xia TL, et al. Relation of premature atrial complexes with stroke and death: Systematic review and meta-analysis. Clin Cardiol. 2017;40(11):962-9. https://doi.org/10.1002/clc.22780.
15. Himmelreich JCL, Lucassen WAM, Heugen M, Bossuyt PMM, Tan HL, Harskamp RE, et al. Frequent premature atrial contractions are associated with atrial fibrillation, brain ischaemia, and mortality: a systematic review and meta-analysis. Europace. 2019;21(5):698-707. https://doi.org/10.1093/europace/euy276.
16. Harms PP, van der Heijden AA, Rutters F, Tan HL, Beulens JWJ, Nijpels G, et al. Prevalence of ECG abnormalities in people with type 2 diabetes: The Hoorn Diabetes Care System cohort. J Diabetes Complications. 2021;35(2):107810. https://doi.org/10.1016/j.jdiacomp.2020.107810.
17. Harms PP, Elders PPJM, Rutters F, Lissenberg-Witte BI, Tan HL, Beulens JWJ, et al. Longitudinal association of ECG abnormalities with major adverse cardiac events in people with type 2 diabetes: The Hoorn Diabetes Care System cohort. manuscript submitted for publication 2022.
18. Auer R, Bauer DC, Marques-Vidal P, Butler J, Min LJ, Cornuz J, et al. Association of major and minor ECG abnormalities with coronary heart disease events. JAMA. 2012;307(14):1497-505. https://doi.org/10.1001/jama.2012.434.
19. Sawai T, Imano H, Muraki I, Hayama-Terada M, Shimizu Y, Cui R, et al. Changes in ischaemic ECG abnormalities and subsequent risk of cardiovascular disease. Heart Asia. 2017;9(1):36-43. https://doi.org/10.1136/heartasia-2016-010846.
20. Strom Moller C, Zethelius B, Sundstrom J, Lind L. Persistent ischaemic ECG abnormalities on repeated ECG examination have important prognostic value for cardiovascular disease beyond established risk factors: a population-based study in middle-aged men with up to 32 years of follow-up. Heart. 2007;93(9):1104-10. https://doi.org/10.1136/hrt.2006.109116.
21. van der Heijden AA, Rauh SP, Dekker JM, Beulens JW, Elders P, t Hart LM, et al. The Hoorn Diabetes Care System (DCS) cohort. A prospective cohort of persons with type 2 diabetes treated in primary care in the Netherlands. BMJ Open. 2017;7(5):e015599. https://doi.org/10.1136/bmjopen-2016-015599.
22. Prineas RJ, Crow RS, Blackburn HW. The Minnesota code manual of electrocardiographic findings : standards and procedures for measurement and classification. Boston, Mass.: J. Wright; 1982.
23. Nijpels G, van der Heijden AAWA, Elders P, Beulens JWJ, de Vet HCW. The interobserver agreement of ECG abnormalities using Minnesota codes in people with type 2 diabetes. PLOS ONE. 2021;16(8):e0255466. https://doi.org/10.1371/journal.pone.0255466.
24. Martin SS, Blaha MJ, Elshazly MB, Toth PP, Kwiterovich PO, Blumenthal RS, et al. Comparison of a novel method vs the Friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile. JAMA. 2013;310(19):2061-8. https://doi.org/10.1001/jama.2013.280532.
25. Levey AS, Stevens LA. Estimating GFR using the CKD Epidemiology Collaboration (CKD-EPI) creatinine equation: more accurate GFR estimates, lower CKD prevalence estimates,

and better risk predictions. Am J Kidney Dis. 2010;55(4):622-7. https://doi.org/10.1053/j.ajkd.2010.02.337.
26. IBM Team. SPSS statistics. 26 ed. New York, USA: IBM corporation; 2020.
27. R Core Team. R: A Language and Environment for Statistical Computing. 4.0.3 ed. Vienna, Austria: R Foundation for Statistical Computing; 2020.
28. Wickham H, Miller E. haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. 2.3.1 ed2020.
29. Stevenson M, Sergeant E, Nunes T, Heuer C, Marshall J, Sanchez J, et al. epiR: Tools for the Analysis of Epidemiological Data. 2.0.19 ed2021.
30. Therneau TM. survival: Survival Analysis. 3.2-7 ed2020.
31. Kassambara A, Kosinski M, Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. 0.4.8 ed2020.
32. Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, et al. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. 3.3.2 ed2020.
33. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model. Stroke. 2016;47(3):895-900. https://doi.org/10.1161/STROKEAHA.115.012004.
34. Konings K. Revision Dutch Guideline Cardiovascular Disease Prevention 2019. Nederlands Tijdschrift Voor Geneeskunde. 2019;163.
35. Christensen MA, Nguyen KT, Stein PK, Fohtung RB, Soliman EZ, Dewland TA, et al. Atrial ectopy as a mediator of the association between race and atrial fibrillation. Heart Rhythm. 2017;14(12):1856-61. https://doi.org/10.1016/j.hrthm.2017.09.034.
36. Nguyen KT, Vittinghoff E, Dewland TA, Dukes JW, Soliman EZ, Stein PK, et al. Ectopy on a Single 12-Lead ECG, Incident Cardiac Myopathy, and Death in the Community. Journal of the American Heart Association. 2017;6(8). https://doi.org/ARTN e006028 10.1161/JAHA.117.006028.
37. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes Res Clin Pract. 2010;87(1):4-14. https://doi.org/10.1016/j.diabres.2009.10.007.

6

Diagnostic accuracy of a smartphone-operated, single-lead electrocardiography device for detection of rhythm and conduction abnormalities in primary care

Jelle C.L. Himmelreich*, Evert P.M. Karregat*, Wim A.M. Lucassen, Henk C.P.M. van Weert, Joris R. de Groot, M. Louis Handoko, Robin Nijveldt, Ralf E. Harskamp

* Shared first author, these authors contributed equally to this work

Annals of Family Medicine 2019, Volume 17, Issue 5,

September/October 2019, Pages 403-11,

https://doi.org/10.1370/afm.2438

ABSTRACT

Purpose: To validate a smartphone-operated, single-lead electrocardiography (1L-ECG) device (AliveCor KardiaMobile) with an integrated algorithm for atrial fibrillation (AF) against 12-lead ECG (12L-ECG) in a primary care population.

Methods: We recruited consecutive patients who underwent 12L-ECG for any non-acute indication. Patients held a smartphone with connected 1L-ECG while local personnel simultaneously performed 12L-ECG. All 1L-ECG recordings were assessed by blinded cardiologists as well as by the smartphone-integrated algorithm. The study cardiologists also assessed all 12L-recordings in random order as the reference standard. We determined the diagnostic accuracy of the 1L-ECG in detecting AF or atrial flutter (AFL) as well as any rhythm abnormality and any conduction abnormality with the simultaneously performed 12L-ECG as the reference standard.

Results: We included 214 patients from 10 Dutch general practices. Mean ± SD age was 64.1±14.7 years, and 53.7% of the patients were male. The 12L-ECG diagnosed AF/AFL, any rhythm abnormality, and any conduction abnormality in 23, 44, and 28 patients, respectively. The 1L-ECG as assessed by cardiologists had a sensitivity and specificity for AF/AFL of 100% (95% CI, 85.2%-100%) and 100% (95% CI, 98.1%-100%). The AF detection algorithm had a sensitivity and specificity of 87.0% (95% CI, 66.4%-97.2%) and 97.9% (95% CI, 94.7%-99.4%).

The 1L-ECG as assessed by cardiologists had a sensitivity and specificity for any rhythm abnormality of 90.9% (95% CI, 78.3%-97.5%) and 93.5% (95% CI, 88.7%-96.7%) and for any conduction abnormality of 46.4% (95% CI, 27.5%-66.1%) and 100% (95% CI, 98.0%-100%).

Conclusions: In a primary care population, a smartphone-operated, 1L-ECG device showed excellent diagnostic accuracy for AF/AFL and good diagnostic accuracy for other rhythm abnormalities. The 1L-ECG device was less sensitive for conduction abnormalities.

PURPOSE

Patients frequently visit their primary care physician with symptoms that may be due to cardiac arrhythmias.[1] Manifestations include palpitations, light-headedness, and (near) fainting and account for 0.8% to 16% of symptoms that prompt patients to visit their primary care physician.[1,2] Some heart rhythm abnormalities, such as ectopic beats, are common electrocardiography (ECG) findings that generally do not require action.[3] Others, such as atrial fibrillation (AF) or atrial flutter (AFL), are present in approximately 2% to 3% of the population and warrant further work-up and management to reduce associated risks of stroke and heart failure.[4-6] When a cardiac arrhythmia is suspected in a symptomatic patient, resting 12-lead ECG (12L-ECG) should always be performed.[7] Unfortunately, in primary care, performing 12L-ECG can be cumbersome, particularly during house visits, and it is not available at every practice. As a result, only in approximately one-third of cases is ECG performed during a symptomatic period.[3]

The availability of an unobtrusive, handheld ECG device is likely to lower the logistical threshold for performing ECG and may therefore improve detection of relevant arrhythmias in primary care.[8] One such device, the KardiaMobile, is a smartphone-connected, single-lead ECG (1L-ECG) device.[9,10] Smartphone-operated ECG has been studied for screening purposes and has shown great promise.[11] A recent report issued by the United Kingdom's National Health Service expects the device to be highly cost saving in the context of primary care.[12]

To our knowledge, the KardiaMobile has not yet been validated against simultaneously performed 12LECG in a primary care population. We hypothesized that the information obtained with smartphone-operated 1L-ECG can be used to accurately detect AF/AFL and common ectopic beats. We therefore performed a multicenter validation study in primary care to assess the validity of 1L-ECG as an office/bedside tool for the detection of arrhythmias as well as rhythm and conduction abnormalities compared with simultaneously performed 12L-ECG as assessed by blinded cardiologists as the reference standard.

METHODS

We reported this diagnostic accuracy study in accordance with the Standards for Reporting of Diagnostic Accuracy Studies (STARD) 2015 statement.[13] The study protocol was approved by our institution's Medical Ethical Review Committee. All participants provided written informed consent.

Study Design

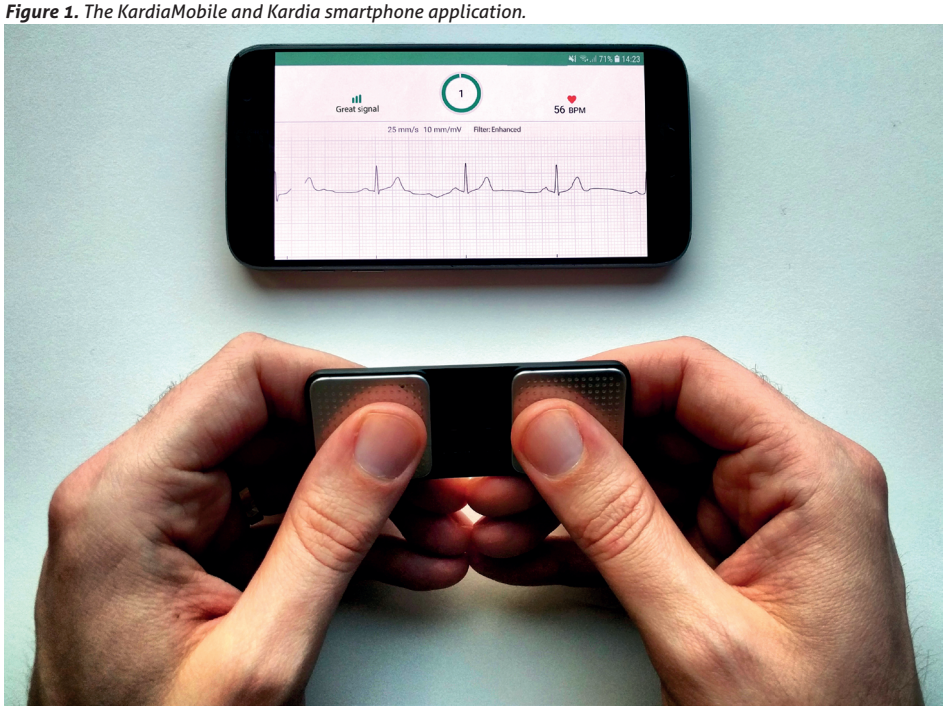
We enrolled consecutive patients as part of the Validation of a mobile bedside ECG Screening and diagnostic Tool for Arrhythmias in general practice (VESTA) study. Eligible patients were aged 18 years or older who were assigned to 12L-ECG for any nonacute indication as ordered by the local primary care physician in 1 of 10 participating general practices across the Netherlands. The practices were in possession of a 12L-ECG device and had qualified and skilled personnel to perform 12L-ECG. Exclusion criteria were a clinically acute indication for ECG as defined by the local primary care physician (eg, suspicion of acute coronary syndrome) and presence of a pacemaker rhythm on 12L-ECG. We categorized patients according to indication for 12L-ECG either because of presentation with new symptoms (symptom-driven ECG) or as an integral part of protocolized care for primary or secondary prevention of cardiovascular disease (protocol-driven ECG). For each participant, the study design involved 3 different readings as follows: (1) the 1L-ECG read by the AF detection algorithm of the smartphone application, (2) the 1L-ECG read by cardiologists, and (3) the standard 12L-ECG read by cardiologists.

Index Test

The KardiaMobile (AliveCor, Inc) is a smartphone-connected, 1L-ECG device that displays ECG recordings in real time (30 seconds) via a smartphone application with a built-in AF detection algorithm (Figure 1). The 1L-ECG recordings were assessed in 2 ways as follows:

1. The AF detection algorithm assessed all 1L-ECG recordings. It classified recordings as either possible AF, normal, or unreadable, or provided no classification. We marked all recordings classified as possible AF as positive for AF. We marked all other algorithm classifications, or when no classification was provided, as negative for AF. The algorithm did not provide a classification for when a 1L-ECG recording was truncated (<30 seconds).
2. Cardiologists (M.L.H., R.N., J.R.dG.) assessed all 1L-ECG recordings in randomized order. The evaluation consisted of scoring each recording for the presence of arrhythmias, ectopic beats, and conduction abnormalities according to a scoring template designed for this study (see Supplementary Methods for exact definitions).

Figure 1. *The KardiaMobile and Kardia smartphone application.*



Photograph by Jelle Himmelreich.

Reference Standard

All 12L-ECG recordings were independently evaluated by 2 cardiologists, and in case of disagreement, by a third cardiologist (M.L.H., R.N., J.R.dG.). We presented 12L-ECG recordings to the cardiologists in randomized order. The evaluation consisted of scoring each recording for the presence of arrhythmias, ectopic beats, and conduction abnormalities according to a scoring template designed for this study (see Supplementary Methods for exact definitions).

Rhythm Measurement

Personnel instructed each patient to commence the KardiaMobile recording by holding the device loosely with both hands (corresponding with lead I for 12L-ECG). We advised patients who used hand lotion or had sweaty hands to wash their hands with soap or to use alcohol wipes on the fingertips to optimize electrical conduction quality. When a steady 1L-ECG signal was visible on the smartphone, the local investigator started a 10-second 12L-ECG recording. We thereby obtained 10 seconds of simultaneous recording. We excluded patients for whom 1 or both ECG types were not available or when there was no 10-second overlap between recording types. The 1L-ECG recordings were not used for clinical decision making.

Data Collection

Three investigators (J.C.L.H., E.P.M.K., R.E.H.) visited participating practices to collect the 12L-ECG recordings (as PDF file or photocopy of paper original) as well as patient data at the time of index ECG from the practice's electronic health records. We collected the corresponding 1L-ECG recordings (PDF files) from the secure online platform that is part of the KardiaMobile software package. Baseline data included sex, age, indication for undergoing 12L-ECG, use of relevant antiarrhythmic drugs, and relevant medical history. We pseudonymized all data before storing it in a secured electronic case report form (Castor EDC).

Statistical Analysis

We expressed diagnostic accuracy for all analyses as sensitivity, specificity, positive and negative likelihood ratios, and positive and negative predictive values, with 95% CI. The primary analyses of this study were (1) the diagnostic accuracy of 1L-ECG as assessed by cardiologists in detecting AF or AFL with 12L-ECG as reference and (2) the diagnostic accuracy of the AF detection algorithm for AF/AFL with 12L-ECG as reference. Secondary analyses were (1) the diagnostic accuracy of 1L-ECG as assessed by cardiologists in detecting any rhythm abnormality, defined as any nonsinus rhythm including AF/AFL and/or presence of any ectopic beat, with 12L-ECG as reference and (2) the diagnostic accuracy of 1L-ECG as assessed by cardiologists in detecting any conduction abnormality, defined as presence of atrioventricular (AV) block, bundle branch block (BBB), and/or left axis deviation and/or left anterior fascicular block, with 12L-ECG as reference. We counted the cardiologists' generic assessment of BBB on 1L-ECG as true positive even if specification of subtype of BBB (left BBB or right BBB) was provided by the corresponding 12L-ECG.

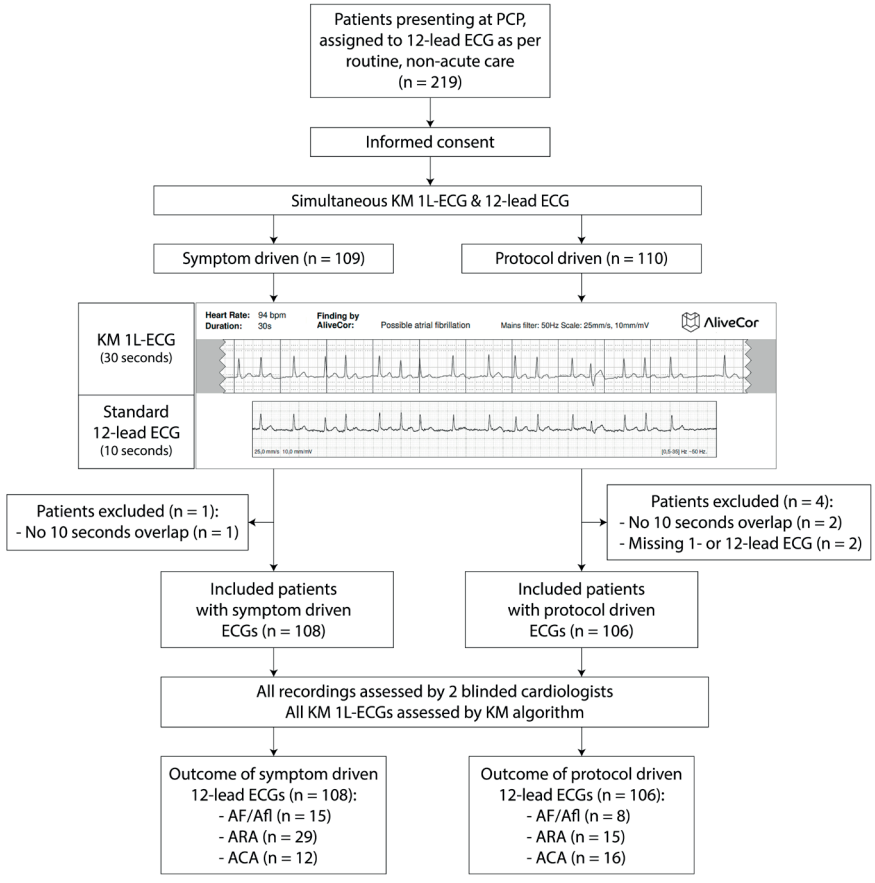
We performed an exploratory analysis of the primary and secondary outcomes stratified by whether ECG was performed based on symptoms or as part of protocol-driven care. We performed a sensitivity analysis on the comparison of the AF detection algorithm vs 12L-ECG for the outcome AF/AFL, in which we excluded all patients with a truncated 1L-ECG recording.

We presented discrete variables as number and percentage and continuous variables as mean ± standard deviation. We compared continuous variables using the Student t test and proportions using the Fisher exact test or Pearson χ^2 test and used 2-tailed tests. We evaluated statistical significance in all analyses at the .05 level and performed analyses using IBM SPSS Statistics version 24.0 (IBM Corp) and MedCalc version 18.10.2 (MedCalc Software).

RESULTS

We included 219 patients during the period April 2017 to July 2018. After excluding 2 patients for missing 1L- and/or 12L-ECG recordings and 3 patients for nonoverlapping recordings, the remaining 214 patients comprised the study population. No adverse device effects were reported. Baseline characteristics of the included patients are listed in Table 1. Mean age was 64.1 ± 14.7 years, and 53.7% of the patients were male. As shown in Figure 2, the indication for performing 12L-ECG was symptom driven for one-half of the patients (n = 108). Among those presenting with new symptoms, most (44.4%) reported palpitations as the primary symptom (Table 2). The 12L-ECG recordings revealed that AF/AFL, any rhythm abnormality, and any conduction abnormality were present in 23, 44, and 28 patients, respectively (Table 3).

Figure 2. *Study flow diagram*



ACA, any conduction abnormality; AF, atrial fibrillation; AFL, atrial flutter; ARA, any rhythm abnormality; bpm, beats per minute; ECG, electrocardiography; PCP, primary care physician; 1L, single-lead; 12L, 12-lead.

Table 1. *Baseline characteristics of the study population*

Variable	Overall (n=214)	Symptom driven ECG patients (n=108)	Protocol driven ECG patients (n=106)
Demographics			
Age (years)	64.1 ± 14.7	59.1 ± 16.3	69.3 ± 10.7*
Female	99 (46.3)	55 (50.9)	44 (41.5)
History			
Obesity (BMI > 30 kg/m^2)	41 (19.2)	12 (11.1)	29 (27.4)*
Smoking			
- Current smoker	36 (16.8)	17 (15.7)	19 (17.9)
- Past history of smoking	72 (33.6)	23 (21.3)	49 (46.2)*
- No history of smoking	72 (33.6)	42 (38.9)	30 (28.3)
- Unknown	34 (15.9)	26 (24.1)	8 (7.5)*
Alcohol abuse	10 (4.7)	5 (4.6)	5 (4.7)
Hypertension	87 (40.7)	31 (28.7)	56 (52.8)*
Diabetes	66 (30.8)	10 (9.3)	56 (52.8)*
Hypercholesterolemia	54 (25.2)	20 (18.5)	34 (32.1)*
Atrial fibrillation or flutter	23 (10.7)	13 (12.0)	10 (9.4)
Other arrhythmia	12 (5.6)	6 (5.6)	6 (5.7)
Coronary heart disease	21 (9.8)	4 (3.7)	17 (16.0)*
TIA or ischemic stroke	13 (6.1)	6 (5.6)	7 (6.6)
Valvular heart disease	9 (4.2)	6 (5.6)	3 (2.8)
Heart failure	8 (3.7)	5 (4.6)	3 (2.8)
Chronic obstructive pulmonary disease	21 (9.8)	9 (8.3)	12 (11.3)
Peripheral vascular disease	19 (8.9)	4 (3.7)	15 (14.2)*
Chronic renal failure	26 (12.1)	9 (8.3)	17 (16.0)
eGFR of those with chronic renal failure (ml/min/1.73m^2)	50.3 ± 6.0	48.9 ± 8.6	51.0 ± 4.5
Medication			
Beta blocker	42 (19.6)	15 (13.9)	27 (25.5)*
Calcium channel blocker	31 (14.5)	11 (10.2)	20 (18.9)
Digoxin	1 (0.5)	0 (0.0)	1 (0.9)
Potassium channel blocker	2 (0.9)	2 (1.9)	0 (0.0)
Sodium channel blocker	1 (0.5)	1 (0.9)	0 (0.0)

BMI, body mass index; eGFR, estimated glomerular filtration rate; m, meter; min, minute; ml, millilitre; SD, standard deviation; TIA, transient ischemic attack.
Data presented as mean and standard deviation for continuous variables and number and percentage between parentheses for categorical variables.
*, $p < 0.05$.

Table 2. *Indications for undergoing 12L-ECG (n = 214)*

Indication	n (%)
Symptom driven ECGs (n = 108)	
Palpitations	48 (44.4)
Other chest symptoms (non-acute)	47 (43.5)
Dyspnea	23 (21.3)
Lightheadedness	16 (14.8)
Fatigue	14 (13.0)
Collapse	3 (2.8)
Other	17 (15.7)
Protocol driven ECGs (n = 106)	
Cardiovascular risk management	34 (32.1)
Known diabetes mellitus	45 (42.5)
Known ischemic heart disease	13 (12.3)
Known heart rhythm disorder	7 (6.6)
Known TIA or ischemic stroke	4 (3.8)
Known heart failure	1 (0.9)
Irregular pulse at examination	1 (0.9)
Follow-up after starting new medication	1 (0.9)

ECG, electrocardiogram; TIA, transient ischemic attack.
We included patients once and reported one reason for ECG per patient. See the Supplementary Methods for how we handled patients with more than one symptom and-or comorbidity.

Diagnostic Accuracy of the 1L-ECG

Data on diagnostic accuracy with calculated 95% CIs are summarized in Table 4. The 2×2 contingency tables with detailed information, including the rhythm diagnoses of all true positives, false negatives, and false positives, can be found in Supplementary Figure S1.

For the primary outcome of AF/AFL, we found that cardiologists were able to correctly classify all 23 cases using 1L-ECG, resulting in a sensitivity and a specificity of 100%. The smartphone-integrated algorithm correctly identified 20 of 23 AF cases and incorrectly classified 4 cases of sinus rhythm as possible AF (sensitivity: 87%; specificity: 97.9%). Interpretation of 1L-ECG was less robust for the secondary endpoints of any rhythm abnormality (sensitivity: 90.9%; specificity: 93.5%) and any conduction abnormality (sensitivity: 46.4%; specificity: 100%). Explicitly for ectopic beats, 1L-ECG correctly classified 20 of 23 cases of known ectopic beats. The false positives for the outcome any rhythm abnormality could all be attributed to misclassified ectopic beats (n = 11).

Table 3. *Outcomes of the 12-lead ECGs (n = 214)*

Outcome	n (%)
Rhythm	
Sinus rhythm	187 (87.4)
Atrial fibrillation	20 (9.3)
Atrial flutter	3 (1.4)
Narrow complex tachycardia	3 (1.4)
Broad complex tachycardia	0 (0.0)
Ectopic atrial rhythm	1 (0.5)
Ectopic beats	
Premature atrial complex	7 (3.3)
Premature ventricular complex	16 (7.5)
Conduction abnormalities	
AV block	7 (3.3)
- 1st degree AV block	7 (100)
- 2nd degree AV block, Wenckebach	0 (0.0)
- 2nd degree AV block, Mobitz II	0 (0.0)
- 3rd degree AV block	0 (0.0)
Bundle branch block	23 (10.7)
- LBBB	5 (21.7)
- RBBB	9 (37.5)
- LAD/LAFB	9 (37.5)
Composite outcomes	
Atrial fibrillation or flutter	23 (10.7)
Any rhythm abnormality*	44 (20.6)
Any conduction abnormality#	28 (13.1)

AV, atrioventricular; LAD/LAFB, left axis deviation and/or left anterior fascicular block; LBBB, left bundle branch block; RBBB, right bundle branch block.
*, n = 6 patients showed 2 rhythm abnormalities on 12-lead ECG (see *Supplementary Figure S1*);
#, n = 1 patients showed 3 conduction abnormalities on 12-lead ECG (see *Supplementary Figure S1*).

Additional Analyses

The stratified analysis according to indication for ECG (symptom or protocol driven) and the sensitivity analysis in which we excluded truncated 1L-ECG recordings (n = 6) rendered similar results (see Supplementary Figure S2, Supplementary Table S1, and Supplementary Table S2, respectively).

Table 4. Diagnostic Accuracy Measures of the Interpretation of 1L-ECG by Cardiologists or the Smartphone Algorithm Using 12L-ECG as Reference Standard

Outcome Assessor	Sensitivity (95%CI)	Specificity (95%CI)	LR+ (95%CI)	LR- (95%CI)	PPV (95%CI)	NPV (95%CI)
Atrial fibrillation or flutter						
Cardiologists	100% (85.2-100)	100% (98.1-100)	∞*	0**	100%#	100%#
Smartphone algorithm	87.0% (66.4-97.2)	97.9% (94.7-99.4)	41.5 (15.5-110.9)	0.13 (0.05-0.38)	83.3% (65.2-93.0)	98.4% (95.6-99.4)
Any rhythm abnormality						
Cardiologists	90.9% (78.3-97.5)	93.5% (88.7-96.7)	14.1 (7.9-25.1)	0.10 (0.04-0.25)	78.4% (67.1-86.7)	97.6% (94.0-99.0)
Any conduction abnormality						
Cardiologists	46.4% (27.5-66.1)	100% (98.0-100)	∞*	0.54 (0.38-0.76)	100%#	92.5% (89.8-94.6)

LR+, positive likelihood ratio; LR-, negative likelihood ratio; N/A, not applicable; NPV, negative predictive value; PPV, positive predicting value; 95%CI, 95% confidence interval. *, LR+ is infinite and 95%CI is not applicable when specificity = 100%[14]; **, LR- is 0 and 95%CI is not applicable when sensitivity = 100%[14]; #, 95%CI is not applicable when PPV or NPV = 100%.[14,15]

DISCUSSION

The diagnostic properties of the KardiaMobile 1L-ECG device as assessed by cardiologists against simultaneously performed 12L-ECG in a primary care population were excellent for AF/AFL. The AF detection algorithm showed high sensitivity and specificity for AF/AFL. Visual assessment of the 1L-ECG recordings by cardiologists resulted in high sensitivity and specificity for rhythm abnormalities and high specificity but low sensitivity for conduction abnormalities. To our knowledge, this is the first study to validate the KardiaMobile device for both AF and common non-AF ECG abnormalities against simultaneously performed 12L-ECG in a primary care population.

Clinical Relevance

Patients who present to their primary care physician with palpitations often no longer have symptoms at the time of consultation or when ECG is performed.[3] When ECG is performed during palpitations, an abnormal heart rhythm is found in approximately one-half of patients, whereas this drops to approximately one-fifth when symptoms are no longer present at the time of ECG.[3] The findings from the present study are therefore highly relevant for primary care physicians because the smartphone-operated ECG device operates as a point- of-care test and allows for immediate rhythm assessment during a symptomatic episode. Moreover, our findings support patients' use of the device at home as a 1L event recorder, provided that the ECG readings

are assessed by a cardiologist. We note that the device is already available on the consumer market for this purpose.

Our stratified analysis by indication for ECG showed that in older patients for whom ECG was not indicated primarily for cardiac symptoms, a negative reading excluded AF with a similarly high degree of certainty as that for symptomatic patients, despite differences in pretest likelihood within our sample. These results may be relevant for primary care physicians because they are encouraged to perform proactive case identification in asymptomatic patients with elevated risk of developing AF (eg, via pulse palpation followed by ECG).[16] Here, the 1L-ECG device could be a valuable point-of-care tool for at-risk patients for whom traveling to the practice for standard 12L-ECG is too cumbersome or for primary care physicians who do not possess a 12L-ECG device.

We added the comparison on any rhythm abnormality because for primary care patients, cardiac symptoms may often be explained by ectopy.[1] We found that the 1L-ECG device can correctly classify instances of ectopic beats, suggesting that it may be useful as a point-of-care diagnostic instrument for this rhythm anomaly.

The 1L-ECG device was less sensitive for conduction abnormalities, which in the present study particularly involved the detection of first-degree AV blocks. For primary care physicians, however, the detection of conduction abnormalities is generally less clinically relevant than the detection of arrhythmias, with the notable exception of decisions regarding the prescribing of QT-prolonging medication.[17] Whereas the QT interval was not scored in the present study, others have reported the KardiaMobile's ability to accurately assess QT intervals.[18]

Strengths and Limitations

Our study had a number of strengths. First, we included consecutive patients who underwent 12L-ECG as part of routine medical practice, resulting in a cohort generalizable to general practice. Second, the study design ensured simultaneous rather than consecutive 1L- and 12L-ECG recordings, as done in prior studies.[10,19,20] This allowed for a comparison on the detection of ectopic beats, which may be a frequent cause for palpitations in primary care.[21] Third, by providing a stratified analysis according to indication for ECG, we were able to show that the 1L-ECG device performed similarly in patients with symptoms vs those who present as part of protocol-driven (secondary) preventive care. Fourth, we ensured standardized interpretation of all recordings by blinded assessment of 1L-and 12L-ECG recordings in random order.

Several limitations should be mentioned. First, this study was not designed to determine to what extent primary care physicians are able to assess the 1L-ECG signal, but rather to describe the test characteristics of the 1L-ECG device in a representative primary care patient sample when analyzed by experts (cardiologists/electrophysiologists). Second, the use of recordings of different durations (10-second 12L-ECG vs 30-second 1L-ECG) may have led us to underestimate the specificity of 1L-ECG in the analysis of any rhythm abnormality, given that ectopic beats might have occurred during the nonoverlapping 20 seconds of the 1L-ECG recording. Third, we presented cardiologists with the PDF file of the 1L-ECG recording instead of having them assess the recording from a smartphone or tablet screen, which is how physicians will often use the device.[20] Fourth, the KardiaMobile application did not provide an automated assessment of conduction intervals in milliseconds, as is done for most 12L-ECGs. This might have affected sensitivity in the analysis on any conduction abnormality. Fifth, the 95% CIs were relatively wide, owing to sample size and prevalence of the studied outcomes among the cohort. Finally, the present study was not designed to study whether the availability of a smartphone 1L-ECG would change ECG use, diagnosis, or patient management.

Previous Work

The good diagnostic properties that we found for the 1L-ECG device for AF/AFL, when assessed by cardiologists or by the smartphone application algorithm, coincide with a number of prior studies[10,11,19,20,22-27] (see Supplementary Table S3 and Supplementary Table S4 for an overview of prior studies that validated the KardiaMobile 1L-ECG device for rhythm and/or conduction abnormalities). Notable exceptions are 2 studies by Chan et al[8,28] and 1 study by Desteghe et al[29] that reported sensitivities of 71.4%, 66.7%, and 65.9%, respectively, for the KardiaMobile algorithm to detect AF. The authors provide no clear explanation for the AF-detection algorithm's low sensitivity in their respective studies, which were all performed with selected elderly patients.

Although a number of studies have assessed the presence of ectopic beats on 1L-ECG recordings, none have validated 1L-ECG for ectopy alone or as part of a composite outcome.[8,10,11,22-24,28,29] One study validated 1L-ECG against 12L-ECG for conduction abnormalities. That study, by Haberman et al,[19] found high specificity but sensitivities of 77.3% and 72.4%, respectively, for AV block and BBB. The results for AV block contrast with those from our present study, in which none of the AV blocks were detected using the 1L-ECG device (Figure 2). We note that Haberman et al[19] determined automated conduction intervals for 1L-ECG before assessment by electrophysiologists, whereas in our present study, automated intervals for 1L-ECG were absent.

Our present work adds to the literature by validating 1L-ECG against 12L-ECG in a primary care setting of consecutive patients and by validating 1L-ECG for a broad spectrum of cardiac arrhythmias and conduction disturbances including ectopic beats.

Future Work

Further study is required to evaluate the safety and efficacy of the 1L-ECG device in the hands of primary care physicians instead of cardiologists, particularly for detecting AF/AFL. Moreover, future studies are warranted to determine whether the availability of 1L-ECG changes the use of 12L-ECG, diagnosis, and/or patient management. Data should be obtained to study the net benefit as well as impact on cost-effectiveness of adding the KardiaMobile algorithm's or cardiologists' assessment to that of primary care physicians for arrhythmia detection. Findings from such studies might determine whether and how the KardiaMobile 1L-ECG device can be safely and effectively implemented in clinical practice as well as used in future screening programs for detecting AF in at-risk general populations.

CONCLUSIONS

A smartphone-operated, 1L-ECG device is a reliable instrument for detecting AF when assessed by the internal detection algorithm, and even more so when assessed by cardiologists. Moreover, the 1L-ECG recording can display atrial and ventricular ectopy with high sensitivity. The 1L-ECG recording was less robust for detecting conduction delays. Our primary care–based study provides important insights for physicians who are in need of a point-of-care ECG device that can lower the logistical threshold for performing ECG to improve diagnostic gain.

Acknowledgments: We thank all participating primary care practices for their cooperation in performing this study. We thank Mrs Lucinda Bertels for her assistance with Figure 1.

Conflicts of interest: Authors report none.

Funding support: This work was supported by the Netherlands Organisation for Health Research and Development (ZonMw) (80-83910-98-13046). Salary support for Dr Harskamp was provided by a Rubicon fellowship of the Netherlands Organization for Scientific Research (NWO). Dr de Groot is supported by a personal VIDI grant from NWO/ZonMW (016.146.310), reports research grants through his institution from Abbott, Atricure, Boston Scientific, and Medtronic, and received consultancy/speakers

fees from Atricure, Bayer, Daiichi Sankyo, Johnson & Johnson, Medtronic, Novartis, and Servier; all outside the scope of this study. All devices and research efforts were paid from university funds. The authors received no funding from the device's producer or local distributor. The authors report no ties to the manufacturer of the investigated device and had full autonomy in the design, conduct, and reporting of this manuscript.

Previous presentations: A scientific poster presentation (hard copy) with preliminary results was presented by Jelle Himmelreich at the HartVaatHAG conference for Dutch primary care practitioners, October 5, 2018, Utrecht, The Netherlands, and at the Amsterdam Public Health Annual Meeting, November 22, 2018, Amsterdam; The Netherlands. Abstracts for these local conferences have not been published either in print or electronically.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY METHODS

Definitions of items scored in all recordings

The cardiologists scored each recording for rhythm (sinus rhythm, atrial fibrillation, atrial flutter, narrow complex (non-sinus) tachycardia, broad complex (non-sinus) tachycardia, ectopic atrial rhythm), presence of ectopic beats (premature atrial or ventricular complexes) and conduction disorders (atrioventricular block defined as PR interval >200ms, bundle branch block defined as QRS duration >120ms, and left axis deviation and/or left anterior fascicular block) according to a scoring template especially designed for this study.

Patients with multiple symptoms and/or comorbidities

In case of multiple symptoms in a symptom driven ECG we used the first reported symptom in the medical record as the index symptom for that patient. When a patient was due to receive a protocol driven ECG, but also reported to have had cardiac symptoms prior to the appointment for ECG, we still counted this ECG as protocol driven since the timing of the ECG was not influenced by the symptoms.

In case of multiple comorbidities in protocol driven ECGs we assessed for which chronic care program the ECG was primarily intended. Since Dutch primary care physicians label all patients who are in the cardiovascular risk management (CVRM) program as 'CVRM patient', we counted the protocol driven ECGs of patients with the CVRM label as such. In case of multiple comorbidities but no CVRM label, we assessed what the stated primary reason was for making the ECG appointment as assessed by documentation of the current and/or previous consultations.

SUPPLEMENTARY TABLES

Supplementary Table S1. *Diagnostic accuracy measures of the interpretation of the single-lead ECG by cardiologists or the smartphone algorithm using 12-lead ECG as reference standard: stratified analysis according to indication for ECG*

Outcome Assessor	Sensitivity (95%CI)	Specificity (95%CI)	LR+ (95%CI)	LR- (95%CI)	PPV (95%CI)	NPV (95%CI)
Symptom driven ECGs (n = 108)						
Atrial fibrillation or flutter						
Cardiologists	100% (78.2-100)	100% (96.1-100)	∞*	0**	100%#	100%#
Smartphone algorithm	86.7% (59.4-98.3)	95.7% (89.4-98.8)	20.2 (7.6-53.6)	0.14 (0.04-0.51)	76.5% (55.0-89.6)	97.8% (92.5-99.4)
Any rhythm abnormality						
Cardiologists	96.6% (82.2-99.9)	94.9% (87.5-98.6)	19.1 (7.3-49.7)	0.04 (0.01-0.25)	87.5% (72.9-94.8)	98.7% (91.6-99.8)
Any conduction abnormality						
Cardiologists	33.3% (9.9-65.1)	100% (96.2-100)	∞*	0.67 (0.45-0.99)	100%#	92.3% (88.9-94.7)
Protocol driven ECGs (n = 106)						
Atrial fibrillation or flutter						
Cardiologists	100% (63.1-100)	100% (96.3-100)	∞*	0**	100%#	100%#
Smartphone algorithm	87.5% (47.4-99.7)	100% (96.3-100)	∞*	0.12 (0.02-0.78)	100%#	99.0% (94.0-99.8)
Any rhythm abnormality						
Cardiologists	80.0% (51.9-95.7)	92.3% (84.8-96.9)	10.4 (4.9-22.1)	0.22 (0.08-0.60)	63.2% (44.6-78.5)	96.6% (91.0-98.7)
Any conduction abnormality						
Cardiologists	56.3% (29.9-80.3)	100% (96.0-100)	∞*	0.44 (0.25-0.76)	100%#	92.8% (88.1-95.7)

LR+, positive likelihood ratio; LR-, negative likelihood ratio; N/A, not applicable; NPV, negative predictive value; PPV, positive predicting value; 95%CI, 95% confidence interval. *, LR+ is infinite and 95%CI is not applicable when specificity = 100%[14]; **, LR- is 0 and 95%CI is not applicable when sensitivity = 100%[14]; #, 95%CI is not applicable when PPV or NPV = 100%.[15]

Supplementary Table S2. *Diagnostic accuracy of the AF detection smartphone algorithm versus 12-lead ECG: sensitivity analysis including only patients with non-truncated 1L-ECG recordings (n = 208)*

Assessor	Sensitivity (95%CI)	Specificity (95%CI)	LR+ (95%CI)	LR- (95%CI)	PPV (95% CI)	NPV (95%CI)
Smartphone algorithm	87.0% (66.4-97.2)	97.8% (94.6-99.4)	40.2 (15.1-107.4)	0.13 (0.05-0.38)	83.3% (65.2-93.0)	98.4% (95.5-99.4)

LR+, positive likelihood ratio; LR-, negative likelihood ratio; N/A, not applicable; NPV, negative predictive value; PPV, positive predicting value; 95%CI, 95% confidence interval.

Supplementary Table S3. *Characteristics of previous studies that reported sensitivity and specificity of the KM 1L-ECG for rhythm and/or conduction abnormalities*

					Assessment of reference by:		
Study	**Population**	**Outcome**	**n**	**Reference standard**	**ALG**	**C/EP**	**PCP**
Brasier 2018[22]	In-house patients with presumed AF and matched controls in SR	AF	408	Visual assessment of the 1L-ECG	x	x	
Chan 2016[28]	Patients with hypertension, DM or age ≥65 years	AF	1,013	Visual assessment of the 1L-ECG	x	x	
Chan 2017[8]	Patients ≥65 years with hypertension or DM attending an outpatient clinic	AF	2,052	Visual assessment of the 1L-ECG	x	x	
Desteghe 2017[29]	Hospitalized patients at cardiology or geriatric wards	AF	378	6- or 12-lead ECG immediately prior to 1L-ECG	x	x	
Haberman 2015[19]	Healthy young adults, elite athletes and cardiology clinic patients	AF/Afl, AVB, BBB	381	12-lead ECG immediately after 1L-ECG	x	x	
Koshy 2018[20]	Patients before and after elective cardioversion	AF/Afl	51	12-lead ECG immediately prior to 1L-ECG	x	x	x
Lau 2013[10]	Known AF and non-AF patients	AF	204	12-lead ECG max 6 hours before 1L-ECG	x	x	
Lowres 2014[11]	All people aged ≥65 years entering a pharmacy	AF	996	Visual assessment of the 1L-ECG	x	x	
Lowres 2016[23]	Patients with postoperative AF following cardiac surgery	AF	42	Visual assessment of the 1L-ECG	x	x	
Orchard 2016[24]	People aged ≥65 years attending flu vaccination	AF	915	Visual assessment of the 1L-ECG	x	x	
Tarakji 2015[25]	Patients with AF undergoing ablation who had iPhones	AF/Afl	55	Simultaneous TTM		x	
William 2018[26]	AF patients who were admitted for antiarrhythmic drug initiation	AF	52	12-lead ECG immediately prior to 1L-ECG	x	x	
Williams 2015[27]	Outpatients known to be in AF or SR	AF	95	Simultaneous 12-lead ECG		x	

AF, atrial fibrillation; Afl, atrial flutter; ALG, smartphone algorithm; AVB, atrioventricular block; BBB, bundle branch block; C/EP, cardiologist and/or electrophysiologist; DM, diabetes mellitus; ECG, electrocardiogram; PCP, primary care physician; SR, sinus rhythm; TTM, transtelephonic monitor; 1L, single-lead.

Supplementary Table S4. *Outcomes of previous studies that reported sensitivity and specificity of the KM 1L-ECG for rhythm and/or conduction abnormalities*

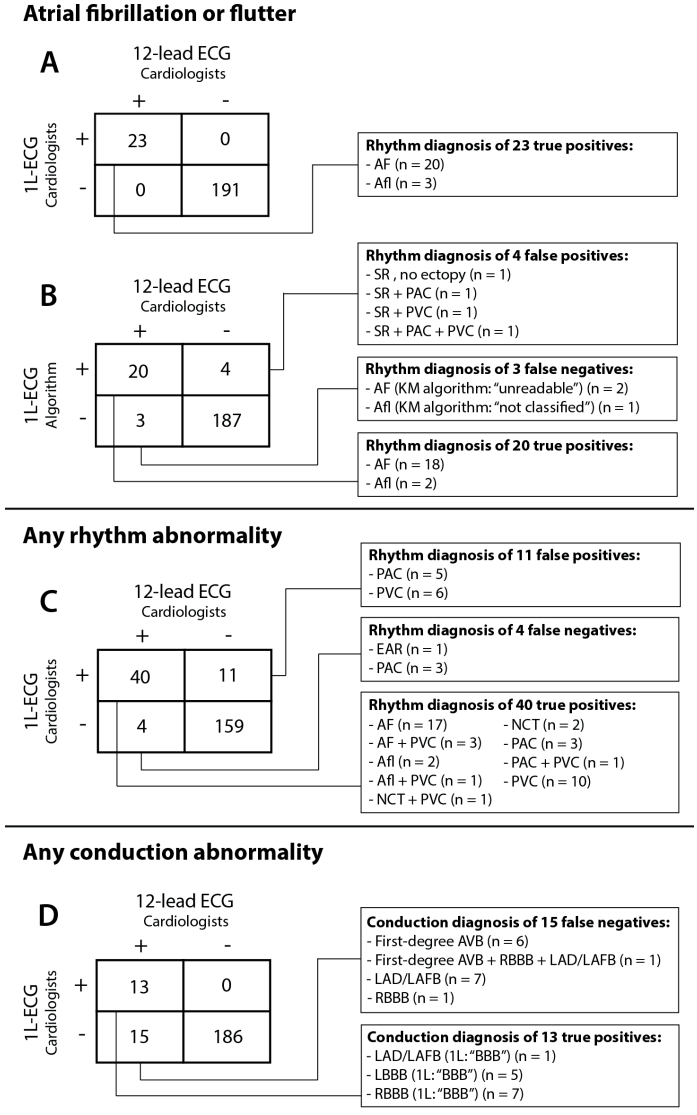
			Outcome					
			AF		AVB		BBB	
Study	Mode of Assessment	Reference standard	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Brasier 2018[22]	Visual	-	-	-	-	-	-	-
	Algorithm	1L-ECG	99.6%	97.8%	-	-	-	-
Chan 2016[28]	Visual	-	-	-	-	-	-	-
	Algorithm	1L-ECG	71.4%	99.4%	-	-	-	-
Chan 2017[8]	Visual	-	-	-	-	-	-	-
	Algorithm	1L-ECG	66.7%	99.5%	-	-	-	-
Desteghe 2017[29]	Visual	12L-ECG	96.2%*	95.6%*	-	-	-	-
	Algorithm	12L-ECG	65.9%	97.6%	-	-	-	-
Haberman 2015[19]	Visual	12L-ECG	94.4%	99.4%	77.3%	96.4%	72.4%	94.9%
	Algorithm	-	-	-	-	-	-	-
Koshy 2018[20]	Visual	12L-ECG	87% (C/EP)* 81% (PCP)*	96% (C/EP)* 90% (PCP)*	-	-	-	-
	Algorithm	12L-ECG	100%#	95%#	-	-	-	-
Lau 2013[10]	Visual	12L-ECG	98%*	92%*	-	-	-	-
	Algorithm	12L-ECG	98%	97%	-	-	-	-
Lowres 2014[11]	Visual	-	-	-	-	-	-	-
	Algorithm	1L-ECG	98.5%	91.4%	-	-	-	-
Lowres 2016[23]	Visual	-	-	-	-	-	-	-
	Algorithm	1L-ECG	94.6%	92.9%	-	-	-	-
Orchard 2016[24]	Visual	-	-	-	-	-	-	-
	Algorithm	1L-ECG	95%	99%	-	-	-	-
Tarakji 2015[25]	Visual	TTM	97%	100%	-	-	-	-
	Algorithm	-	-	-	-	-	-	-
William 2018[26]	Visual	12L-ECG	100%	89%	-	-	-	-
	Algorithm	12L-ECG	96.6%	94%	-	-	-	-
Williams 2015[27]	Visual	12L-ECG	91.4%*	81.1%*	-	-	-	-
	Algorithm	-	-	-	-	-	-	-

AF, atrial fibrillation; Afl, atrial flutter; AVB, atrioventricular block; BBB, bundle branch block; C/EP, cardiologist and/or electrophysiologist; ECG, electrocardiogram; PCP, primary care physician; SR, sinus rhythm; TTM, trans-telephonic monitor; 1L, single-lead; 12L, 12-lead.

* Study reported separate sensitivity and specificity for multiple individual assessors. Values in this table represent the mean sensitivity and specificity for all assessors reported within the original study; # unclassified recordings excluded from analysis by the original study.

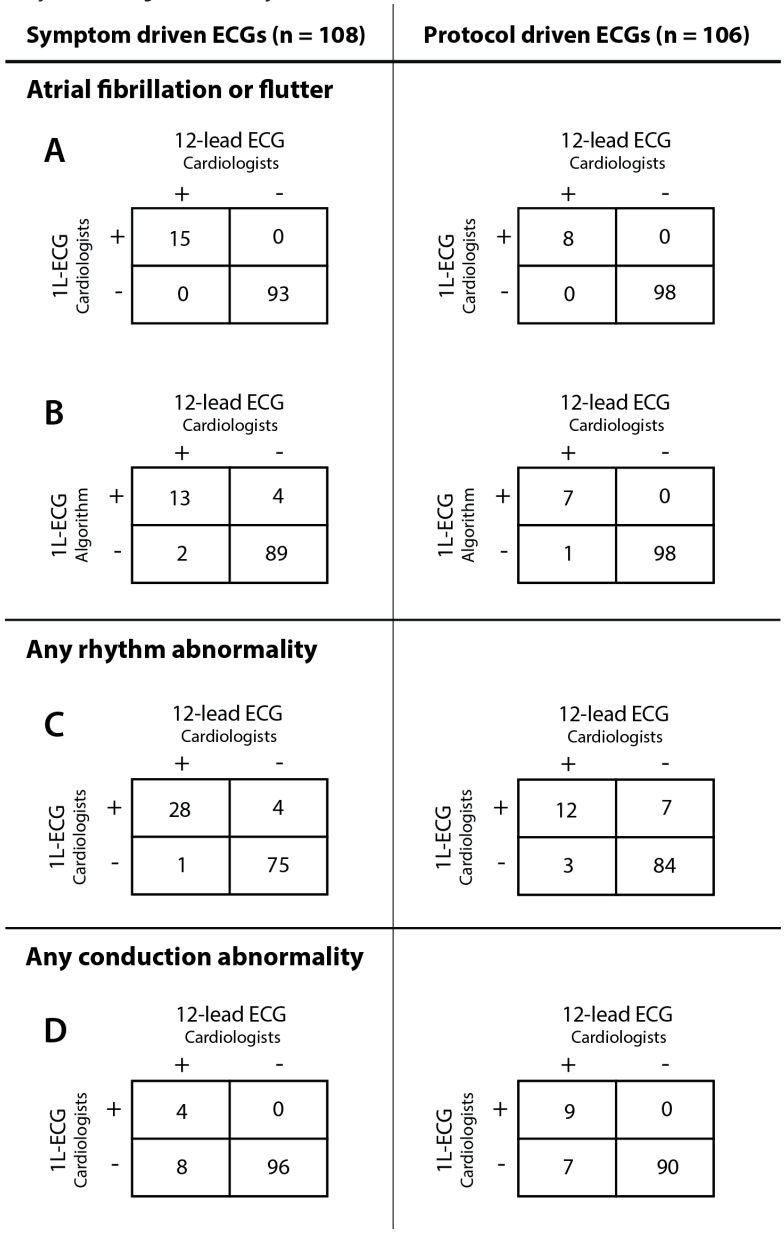
SUPPLEMENTARY FIGURES

Supplementary Figure S1. *2x2 contingency tables of the primary and secondary analyses (n = 214)*



AF, atrial fibrillation ; Afl, atrial flutter; AVB, atrioventricular block; BBB, bundle branch block; EAR, ectopic atrial rhythm; ECG, electrocardiogram; LAD/LAFB, left axis deviation and/or left anterior fascicular block; LBBB, left bundle branch block; NCT, narrow complex tachycardia; PAC, premature atrial complex; PVC, premature ventricular complex; RBBB, right bundle branch block; SR, sinus rhythm; 1L, single-lead.
A: primary analysis on AF/Afl, the 1L-ECG as assessed by cardiologists versus 12-lead ECG;
B: primary analysis on AF/Afl, the 1L-ECG as assessed by the smartphone algorithm versus 12-lead ECG;
C: secondary analysis on any rhythm abnormality, the 1L-ECG as assessed by cardiologists versus 12-lead ECG;
D: secondary analysis on any conduction abnormality, the 1L-ECG as assessed by cardiologists versus 12-lead ECG.

Supplementary Figure S2. *Diagnostic accuracy of the KardiaMobile single-lead ECG: 2x2 contingency tables of the stratified analysis according to indication for ECG*



ACA, any conduction abnormality; AF, atrial fibrillation ; Afl, atrial flutter; ARA, any rhythm abnormality; ECG, electrocardiogram; KM, KardiaMobile; 1L, single-lead.

A: primary analysis on AF/Afl, the 1L-ECG as assessed by cardiologists versus 12-lead ECG;

B: primary analysis on AF/Afl, the 1L-ECG as assessed by the smartphone algorithm versus 12-lead ECG;

C: secondary analysis on any rhythm abnormality, the 1L-ECG as assessed by cardiologists versus 12-lead ECG;

D: secondary analysis on any conduction abnormality, the 1L-ECG as assessed by cardiologists versus 12-lead ECG.

REFERENCES

1. Zwietering P, Knottnerus A, Gorgels T, Rinkens P. Occurrence of arrhythmias in general practice. Scand J Prim Health Care. 1996; 14(4): 244-250.
2. Raviele A, Giada F, Bergfeldt L, Blanc JJ, Blomstrom-Lundqvist C, Mont L, et al., European Heart Rhythm Association. Management of patients with palpitations: a position paper from the European Heart Rhythm Association. Europace. 2011; 13(7): 920-934.
3. Zwietering PJ, Knottnerus JA, Rinkens PE, Kleijne MA, Gorgels AP. Arrhythmias in general practice: diagnostic value of patient characteristics, medical history and symptoms. Fam Pract. 1998; 15(4): 343-353.
4. Kannel WB, Wolf PA, Benjamin EJ, Levy D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. Am J Cardiol. 1998; 82(8A): 2N-9N.
5. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Jr., et al., American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol. 2014; 64(21): e1-e76.
6. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991; 22(8): 983-988.
7. Jones C, Pollit V, Fitzmaurice D, Cowan C, Guideline Development Group. The management of atrial fibrillation: summary of updated NICE guidance. BMJ. 2014; 348: g3655.
8. Chan PH, Wong CK, Pun L, Wong YF, Wong MM, Chu DW, et al. Head-to-head comparison of the Alivecor heart monitor and Microlife WatchBP office AFIB for atrial fibrillation screening in a primary care setting. Circulation. 2017; 135(1): 110-112.
9. Ramkumar S, Nerlekar N, D'Souza D, Pol DJ, Kalman JM, Marwick TH. Atrial fibrillation detection using single lead portable electro-cardiographic monitoring: a systematic review and meta-analysis. BMJ Open. 2018; 8(9): e024178.
10. 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(1): 193-194.
11. 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(6): 1167-1176.
12. Economic Impact Evaluation Case Study: AliveCor Kardia Mobile. Heslington, York, United Kingdom: York Health Economics Consortium; 2018.
13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al., STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015; 351: h5527.
14. Gardner IA, Greiner M. Receiver-operating characteristic curves and likelihood ratios: improvements over traditional methods for the evaluation and application of veterinary clinical pathology tests. Vet Clin Pathol. 2006; 35(1): 8-17.
15. Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. Stat Med. 2007; 26(10): 2170-2183.

16. 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(7616): 383.
17. Liu BA, Juurlink DN. Drugs and the QT interval - caveat doctor. N Engl J Med. 2004; 351(11): 1053-1056.
18. Chung EH, Guise KD. QTC intervals can be assessed with the Alive-Cor heart monitor in patients on dofetilide for atrial fibrillation. J Electrocardiol. 2015; 48(1): 8-9.
19. 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(5): 520-526.
20. Koshy AN, Sajeev JK, Negishi K, Wong MC, Pham CB, Cooray SP, et al. Accuracy of blinded clinician interpretation of single-lead smartphone electrocardiograms and a proposed clinical workflow. Am Heart J. 2018; 205: 149-153.
21. Hoefman E, van Weert HCPM, Reitsma JB, Koster RW, Bindels PJE. Diagnostic yield of patient-activated loop recorders for detecting heart rhythm abnormalities in general practice: a randomised clinical trial. Fam Pract. 2005; 22(5): 478-484.
22. Brasier N, Raichle CJ, Dorr M, Becke A, Nohturfft V, Weber S, et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). Europace. 2019; 21(1): 41-47.
23. Lowres N, Mulcahy G, Gallagher R, Ben Freedman S, Marshman D, Kirkness A, et al. Self-monitoring for atrial fibrillation recurrence in the discharge period post-cardiac surgery using an iPhone electrocardiogram. Eur J Cardiothorac Surg. 2016; 50(1): 44-51.
24. Orchard J, Lowres N, Freedman SB, Ladak L, Lee W, Zwar N, et al. Screening for atrial fibrillation during influenza vaccinations by primary care nurses using a smartphone electrocardiograph (iECG): a feasibility study. Eur J Prev Cardiol. 2016; 23(2 suppl): 13-20.
25. Tarakji KG, Wazni OM, Callahan T, Kanj M, Hakim AH, Wolski K, et al. Using a novel wireless system for monitoring patients after the atrial fibrillation ablation procedure: the iTransmit study. Heart Rhythm. 2015; 12(3): 554-559.
26. William AD, Kanbour M, Callahan T, Bhargava M, Varma N, Rickard J, et al. Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: the iREAD Study. Heart Rhythm. 2018; 15(10): 1561-1565.
27. Williams J, Pearce K, Benett I. The effectiveness of a mobile ECG device in identifying AF: sensitivity, specificity and predictive value. Br J Cardiol. 2015; 22: 70-72.
28. Chan PH, Wong CK, Poh YC, Pun L, Leung WW, Wong YF, et al. Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting. J Am Heart Assoc. 2016; 5(7): e003428.
29. Desteghe L, Raymaekers Z, Lutin M, Vijgen J, Dilling-Boer D, Koopman P, et al. Performance of handheld electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. Europace. 2017; 19(1): 29-39.

7

14-day Holter monitoring for atrial fibrillation after ischemic stroke: The yield of guideline-recommended monitoring duration

Jelle C.L. Himmelreich, Wim A.M. Lucassen, Jonathan M. Coutinho, Ralf E. Harskamp, Joris R. de Groot, Henk C.P.M. van Weert

European Stroke Journal, Volume 8, Issue 1,

March 2023, Pages 157-167,

https://doi.org/10.1177/23969873221146027

ABSTRACT

Introduction: Current European Stroke Organisation (ESO) guidelines recommend >48 h of continuous electrocardiographic monitoring for atrial fibrillation (AF) in all patients with ischemic stroke or transient ischemic attack (TIA) with undetermined origin. We assessed the yield of the guideline-recommended monitoring for AF, as well as of extending monitoring up to 14 days.

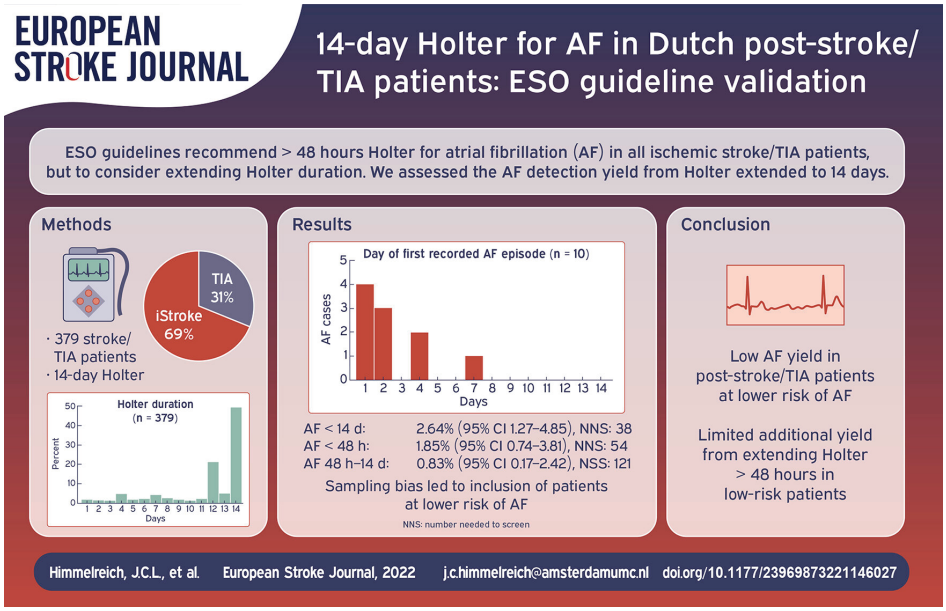
Patients and methods: We included consecutive patients with stroke/TIA without AF in an academic hospital in The Netherlands. We reported AF incidence and number needed to screen (NNS) in the overall sample after 48 h and 14 days of Holter monitoring.

Results: Among 379 patients with median age 63 years (IQR 55–73), 58% male, Holter monitoring detected 10 cases of incident AF during a median of 13 (IQR 12–14) days of monitoring. Seven AF cases were detected within the first 48 hours (incidence 1.85%, 95% CI 0.74–3.81; NNS 54), and three additional AF cases were recorded among the 362 patients with >48 h of monitoring and without AF ≤ 48 h (incidence 0.83%, 95% CI: 0.17–2.42; NNS 121). All AF cases were detected within the first 7 days of monitoring. Our sample was subject to sampling bias favoring inclusion of participants with low AF risk.

Discussion: Strengths of this work were the broad inclusion criteria as recommended by ESO guidelines, and high Holter adherence among participants. The analysis was limited by inclusion of lower-risk cases and a relatively small sample size.

Conclusion:In low-risk patients with recent stroke or TIA, ESO guideline-recommended screening for AF resulted in a low AF yield, with limited additional value of monitoring up to 14 days. Our results underline the need for a personalized approach in determining a patient's optimum duration for post-stroke non-invasive ambulatory monitoring.

GRAPHICAL ABSTRACT



INTRODUCTION

Atrial fibrillation (AF) is an arrhythmia associated with an elevated risk of ischemic stroke that can be effectively lowered with oral anticoagulation.[1] Confirming AF on electrocardiogram (ECG) is however complicated due to its often paroxysmal and/or asymptomatic nature. Screening for paroxysmal AF is therefore warranted in high-risk populations in order to achieve early AF detection to lower the burden of (recurrent) stroke.[1,2]

It is estimated that up to a quarter of ischemic strokes are AF-related.[3,4] Ambulatory post-stroke monitoring yields between 10.7% and 14.7% AF cases, with higher yield among selected patients.[5,6] Therefore, in patients with recent brain ischemia there is consensus to screen for AF if no other cause for stroke is detected.[1,2,7,8] However, there is no uniform recommendation on optimal rhythm monitoring duration.[2] The European Stroke Organisation (ESO), in its Guideline on screening for subclinical atrial fibrillation after stroke or transient ischemic attack of undetermined origin (2022) recommends more than 48 h of cardiac monitoring for AF after stroke or transient attack (TIA). While the American Heart Association/American Stroke Association (AHA/ASS) provide no specific recommendations on monitoring duration,[8] the European Society of Cardiology (ESC) recommends at least 72 h of post-stroke monitoring for AF in all patients with recent ischemic stroke or TIA.[1] Guidelines generally advise to consider extending monitoring duration in order to increase the chance of detecting silent AF, without specific recommendations whom to select for such prolonged monitoring.[1,2,8-10]

We investigated the merits of guideline recommendations for post-stroke ECG monitoring duration by presenting the yield of post-stroke monitoring with 14-day ambulatory Holter in consecutive patients who presented with ischemic stroke or TIA and who were free of AF at inclusion. The primary aim of the current analysis was to assess the overall yield of newly diagnosed AF in an unselected post-stroke or TIA population using 14-day Holter. Secondary aims were to assess AF yields when monitoring according to, as well as beyond, the ESO and ESC guideline-recommended 48 and 72 h, respectively, and to assess AF yield in clinically relevant subgroups of ischemic stroke and TIA patients.

PATIENTS AND METHODS

The current work is presented within the framework of the Risk Assessment for the Identification of Paroxysmal AF (RAPID-AF) study which was initiated to validate novel techniques of assessing the risk of new AF in a combined dataset two high-risk populations: post-stroke and elderly primary care patients [Netherlands Trial Register, NTR6489]. The current analysis aims to provide detailed results on AF yield in RAPID-AF's post-stroke arm. Data on the primary care arm of the RAPID-AF study has been published previously.[11]

Patients

For the post-stroke arm of the RAPID-AF study we included consecutive patients with ischemic stroke or TIA (stroke/TIA) without a history of AF and free of AF on admission ECG who presented to the Neurology department of the Amsterdam University Medical Centers, location Academic Medical Center (AUMC-AMC). We defined ischemic stroke and TIA as an acute loss of focal cerebral or ocular function with symptoms lasting more than or under 24 hours, respectively, and which after adequate investigation was presumed to be due to embolic or thrombotic vascular disease.[7] Inclusion was active from 18 July 2017 through 12 March 2020, and from 11 June 2020 through 17 December 2020, with the intermission and premature end date (distributing n = 400 out of the protocol's stated aim of n = 500 Holter monitors) determined by clinical research restrictions in AUMC-AMC relating to the SARS-CoV-2 pandemic.

Patients were eligible for inclusion in the post-stroke RAPID-AF ambulatory Holter monitoring cohort if they: neither had a history of AF nor de novo AF on ambulance, admission or inpatient ECG or bedside monitor before Holter initiation; were 18 years or older; did not use oral anticoagulation; were free of a pacemaker and/or implantable cardioverter defibrillator; had a life expectancy ≥1 year as estimated by the neurologist in charge; would be able to wear a Holter device for 14 days, and; provided informed consent. We excluded patients who had an alternative explanation for stroke or in whom AF-related stroke was highly unlikely (e.g. periprocedural stroke, symptomatic internal carotid artery (ICA) dissection, or symptomatic ICA occlusion). In patients who presented more than once to our clinic during active inclusion we assessed eligibility only at the first encounter for stroke/TIA.

Study procedures

Standard of care in our center for those presenting with stroke/TIA at time of enrollment consisted of clinical assessment, brain imaging (CT in all patients; MRI at the discretion of the physician e.g. in case of doubt regarding diagnosis or stroke loca-

tion with consequences for treatment), 12-lead resting ECG, laboratory tests, carotid imaging by ultrasound and/or CT-angiogram in case of non-vertebrobasilar stroke/TIA, and ambulatory Holter monitoring (up to 14 days when consenting to RAPID-AF study participation, or up to 72 h in absence of study consent). Initiation and duration of bedside cardiac monitoring in those admitted for stroke was at the discretion of the physician. Patients with TIA or mild stroke generally were not admitted. Patients aged 50 or younger at presentation were given more elaborate investigation including additional laboratory tests, brain MRI and echocardiography as part of the young stroke protocol. For other patients, such additional investigations were at the discretion of the treating physician. Study Holter was the first form of ambulatory cardiac monitoring in all patients included in RAPID-AF.

Additional study procedures were as follows. A study nurse included eligible patients either in the Neurology department, the emergency unit or in the RAPID-AF study outpatient clinic. We aimed to include patients within 90 days after onset of symptoms of the qualifying stroke/TIA. After informed consent, the nurse collected baseline data and instructed patients on the use of the Holter device (Fysiologic ECG Services, Amsterdam, The Netherlands). The study used 2-lead Holters corresponding to leads V1 and V5 of the standard 12-lead ECG, with 8 bit resolution and sampling rate 100 Hz. The leads were applied on the patient's body by three patches attached to one wire leading to a wallet-sized device which was worn in a pouch around the patient's neck. We instructed patients to wear the device continuously except when bathing. We encouraged patients to wear the Holter for the maximum of 14 days, but indicated that they were free to return the device earlier. We instructed patients to return their device either at a return clinical visit or through a prepaid return envelope provided by the study team.

Study procedures were in accordance with the Declaration of Helsinki on medical research involving human subjects.

Baseline data

We collected baseline data at the baseline visit and from the hospital's electronic health records (EHR). At the baseline visit, study personnel asked the patient for data on ethnicity, family history for AF, height, weight, smoking and alcohol consumption. Baseline data retrieved from the EHR consisted of age, sex, index ischemia type and location, stroke severity as per the National Institutes of Health Stroke Scale (NIHSS; score ranging from 0 to 42 with higher scores indicating clinically more severe stroke),[12] blood pressure, baseline 12-lead ECG parameters, medication use, medical history, and relevant routine care laboratory findings. We defined stroke/TIA location

as either retinal, vertebrobasilar, lacunar or non-lacunar anterior, middle or posterior cerebral artery (ACA, MCA, and PCA, respectively) territory, as assessed by clinical symptoms (primarily) and/or available brain imaging. We distinguished the subgroup of patients with non-lacunar hemispheric stroke, defined as retinal, ACA, MCA, or PCA ischemic stroke, due to its relevance in post-stroke AF detection.[13,14] We were unable to distinguish other subgroups relevant to post-stroke AF detection such as cryptogenic stroke or embolic stroke of undetermined source due to lack of systematic pre-enrollment cardiac monitoring and/or echocardiography in our center's standard post-stroke/TIA work-up (see under "Study procedures").[14,15]

We defined vascular disease as history of coronary artery disease, myocardial infarction, peripheral arterial disease, aorta dilatation or known arterial plaques. We defined prior cardiac intervention as a history of coronary artery bypass grafting, percutaneous coronary intervention or cardiothoracic surgery. We determined stroke location based on brain imaging and/or clinical symptoms. We calculated the CHA_2DS_2-VASc score from baseline medical history EHR data.[16]

In order to assess potential sampling bias we collected a selection of baseline variables from stroke/TIA presentations from a random sample of 25% of potentially eligible patients who were not included in our study, as permitted by the Dutch Medical Research Involving Human Subjects Act (WMO) on the use of de-identified routinely collected medical data. In those excluded for a de novo AF diagnosis, we recorded the time at which the AF diagnosis was made (at presentation for, during admission for, or after discharge for their stroke/TIA).

Outcome definitions

The primary outcome was the overall incidence of newly diagnosed (incident) AF as per ambulatory Holter monitoring, with AF defined as AF or atrial flutter lasting ≥30 s.[17] Fysiologic ECG Services (Amsterdam, The Netherlands) performed analysis of all study Holters through digital pre-selection of relevant recordings, followed by manual assessment by trained cardiologists.

Secondary outcomes were the number of days until first AF detection in those with AF diagnosed on study Holter, and AF incidence during monitoring after the guideline-recommended 48 h of Holter monitoring in those who wore their Holter >48 h.[2]

Statistical analysis

We reported medians and interquartile range (IQR) for continuous variables, and numbers and percentages for categorical variables. In case of missing data we reported the

percentage of missing data for each baseline variable with missingness. We displayed baseline characteristics at first presentation for the overall sample, as well as stratified by AF presence on Holter. We plotted the distribution of first day of AF diagnosis in those with AF on Holter, as well as the distribution of the number of days of Holter recording per patient. We calculated incidence and 95% confidence interval (95% CI) and number needed to screen (NNS) using the exact method for AF diagnosed during overall (up to 14 days) Holter monitoring as well as at 24, 48, and 72 h in order to assess the merits of different guidelines.[1,2,10]

We provided a sensitivity analysis of AF incidence and factors associated with AF detection in those with Holter duration over 48 h and without AF detected in the first 48 h of monitoring, in order to assess the added value of monitoring beyond ESO's currently recommended 48 h post-stroke.[2]

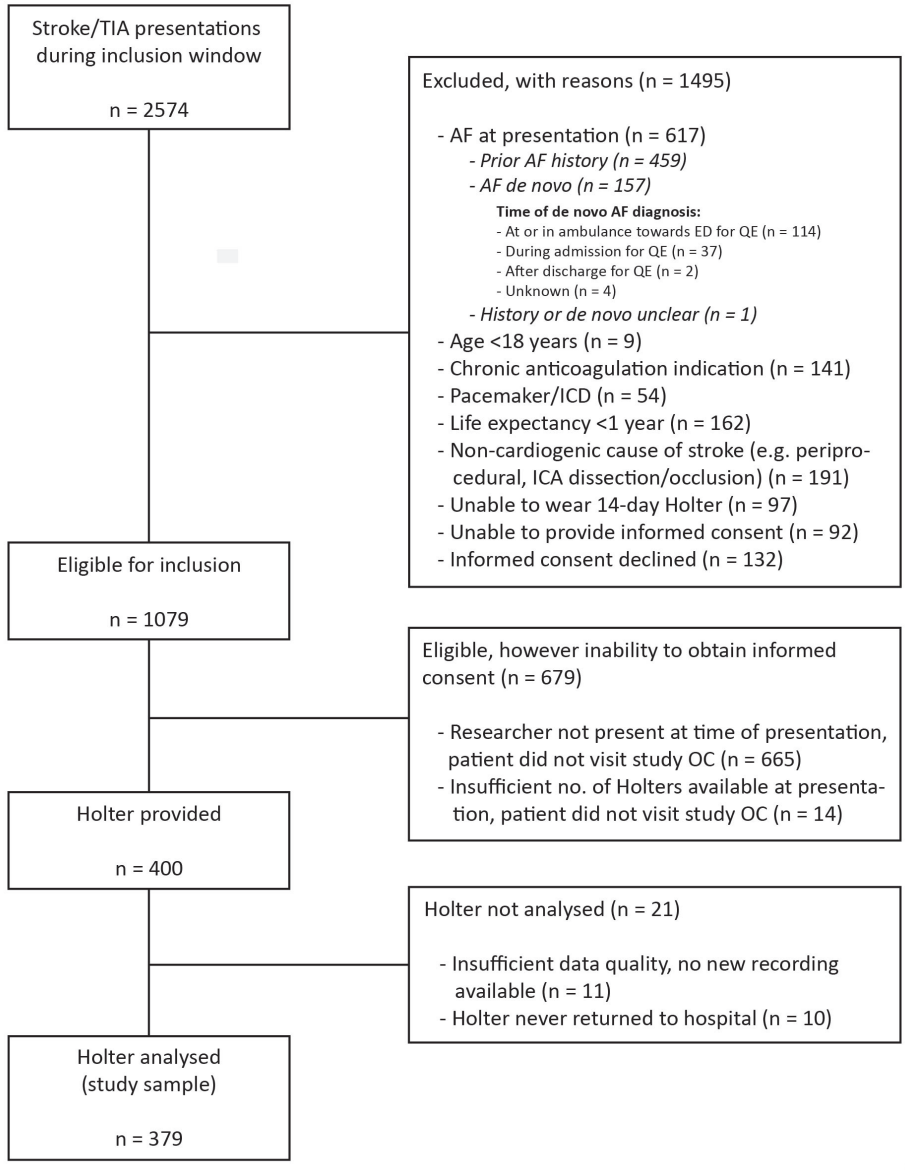
In order to assess whether the application of selection criteria for post-stroke monitoring could have increased AF yield, we presented AF incidence and NNS in subgroups of patients with stroke (not TIA), non-lacunar hemispheric stroke/TIA, and patients with moderate-severe stroke at presentation (NIHSS ≥ 7),[6,18,19] and we presented discrimination of multivariable prediction models developed for post-stroke AF detection. We assessed model discrimination by the C-statistic and 95% CI, using the AS5F (Age, Stroke Severity NIHSS > 5 to Find AF),[20] Re-CHARGE-AF (Recalibrated Cohorts for Heart and Aging Research in Genomic Epidemiology for Atrial Fibrillation), and STAF (Score for the Targeting of AF)[21] risk models, with risk calculated for each patient at baseline using the models' originally published coefficients, and with 95% CIs calculated using 2000 bootstrap samples.

We used R version 3.6.1[22] using the epiR, expss, ggplot2, lubridate, pROC, scales, and table1 packages for our analyses.

RESULTS

Out of 2574 patients who presented with stroke/TIA during the study inclusion windows, 1079 patients (41.9%) were eligible for inclusion (see flowchart, Figure 1). Of these, 400 patients provided written consent and were given a study Holter. We collected analyzable Holter recordings from 379 of these 400 patients, constituting the study sample (35.1% of all eligible patients).

Figure 1. *Study flowchart.*



AF, atrial fibrillation; ED, emergency department; ICA, internal carotid artery; ICD, implantable cardioverter-defi-brillator; OC, outpatient clinic; QE, qualifying event; TIA, transient ischemic attack.

Of the 1495 patients with one or more exclusion criteria, the main reason for exclu-sion was known AF at time of study eligibility assessment (n = 617, 41.3%). Of these, 157 (25.4%) were de novo AF diagnoses, a majority of whom were diagnosed at first presentation (Figure 1).

Patient characteristics

Table 1 shows the main baseline characteristics of included patients. Median age of the included patients was 63 years (IQR: 55–73), 57.8% was male and 69% Caucasian/white. Most patients were included with stroke as qualifying event (68.9% vs 31.1% with TIA), with the middle cerebral artery (MCA) being the most common location for stroke/TIA. Median NIHSS of the total sample was 1 (IQR: 0–3). Median CHA_2DS_2-VASc was 4 (IQR: 3–5). Platelet inhibitors and statins were used at presentation by 34.3% and 36.9%, respectively. Time to Holter was median 35 days (IQR: 14–60), with 88.9% of Holters initiated within 90 days (Figure 2).

Table 1. *Main baseline characteristics of the study sample*

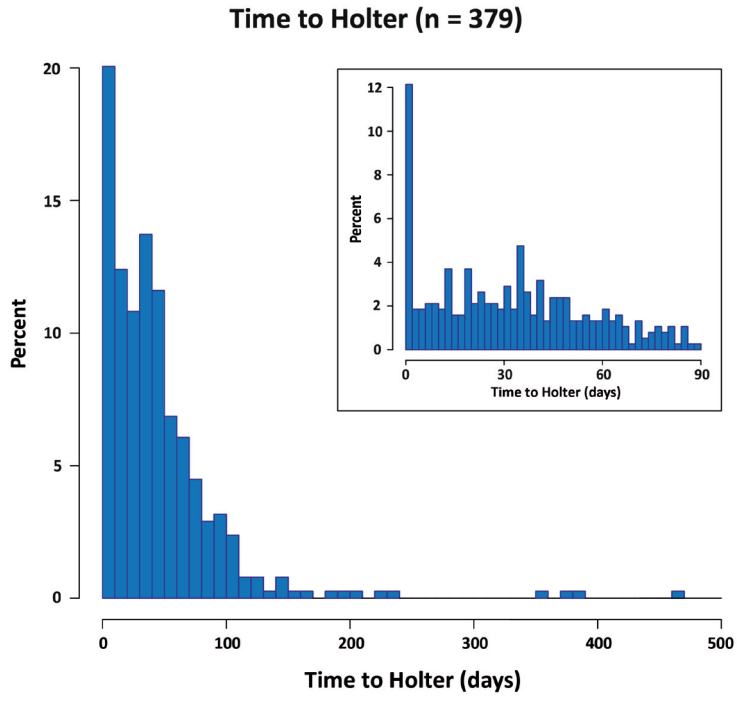
	Overall (n=379)
Female sex (n,%)	160 (42.2%)
Age, years (Median [Q1-Q3])	63.0 (55.0-73.0)
Ethnicity	
Caucasian/white (n,%)	261 (68.9%)
African/black (n,%)	71 (18.7%)
South Asian (n,%)	19 (5.0%)
Asian (n,%)	18 (4.7%)
Other (n,%)	10 (2.6%)
Qualifying event	
Ischemic stroke (n,%)	261 (68.9%)
TIA (n,%)	118 (31.1%)
Stroke/TIA location	
Middle cerebral artery (n,%)	169 (44.6%)
Anterior cerebral artery (n,%)	7 (1.8%)
Posterior cerebral artery (n,%)	20 (5.3%)
Lacunar (n,%)	47 (12.4%)
Retinal (n,%)	16 (4.2%)
Vertebrobasilar (n,%)	120 (31.7%)
Non-lacunar hemispheric stroke (n,%)	131 (34.6)
NIHSS at first presentation (Median [Q1-Q3])	1 (0-3)
Ipsilateral carotid artery stenosis >50% (n,%)	10 (2.6%)
Intravenous thrombolysis (n,%)	74 (19.5%)
Intra-arterial thrombectomy (n,%)	27 (7.1%)
Time to Holter, days (Median [Q1-Q3])	35 (14-60)
Time to Holter ≤90 days (n,%)	337 (88.9%)
Holter duration, days (Median [Q1-Q3])	13 (12-14)
Heart failure (n,%)	10 (2.6%)
Hypertension (n,%)	176 (46.4%)

Table 1. *Main baseline characteristics of the study sample (continued)*

	Overall (n=379)
Diabetes (n,%)	71 (18.7%)
Prior myocardial infarction (n,%)	29 (7.7%)
Prior stroke/TIA/SE (n,%)	83 (21.9%)
Vascular disease (n,%)	49 (12.9%)
CHA_2DS_2-VASc (Median [Q1-Q3])	4 (3-5)
Antiplatelet use (n,%)	130 (34.3%)
ACE/ARB use (n,%)	109 (28.8%)
Calcium antagonist use (n,%)	70 (18.5%)
Diuretics use (n,%)	50 (13.2%)
Statin use (n,%)	140 (36.9%)
Insulin use (n,%)	20 (5.3%)
Metformin use (n,%)	49 (12.9%)

ACE/ARB, angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker; CHA_2DS_2-VASc, congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; NIHSS, National Institutes of Health Stroke Scale; Q1, 1st quartile; Q3, 3rd quartile; SE, systemic embolism; TIA, transient ischemic attack.

Figure 2. *Time between qualifying event and Holter initiation (n = 379)*



Insert: detailed view of time between qualifying event and Holter initiation as percentage of total patients within the first 90 days after qualifying event.

Supplementary Table S1 shows a comparison of baseline characteristics between the study sample and a 25% random sample of non-included eligible patients. Included patients were younger and more often female, had lower NIHSS at presentation, more often had a TIA as qualifying event, less often had an MCA stroke, less frequently underwent intravenous thrombolysis (IVT) or intra-arterial thrombectomy (IAT), and had lower cardiovascular comorbidity and medication use. Of those eligible but not included, 30.1% were presented to our center for tertiary IAT care after which they were discharged to their referring secondary care centers. A further 26.6% were discharged during admission due to shortage of beds in our center and/or residence outside the Amsterdam region.

Holter results

Patients wore their Holter for a median of 13 (IQR: 12–14) days, with 96.0% wearing the device more than 3 days, and 83.1% using the Holter more than 7 days (Figure 3). Overall, 14-day study Holter recorded 10 AF diagnoses (2.64%, 95% CI 1.27–4.85; NNS 38). Four cases were diagnosed in the first 24 h of Holter monitoring (incidence 1.06%, 95% CI 0.29–2.70; NNS 94), and seven were recorded within the first 48 hours with no additional cases in the third day of monitoring (48- and 72-h incidence 1.85%, 95% CI 0.74–3.81; NNS 54). All cases in our sample were detected within the first week of Holter monitoring (Figure 4). Time to Holter initiation was not associated with AF detection in our sample (Supplementary Figure S1).

Figure 3. *Total Holter recording duration in the overall sample (n = 379).*

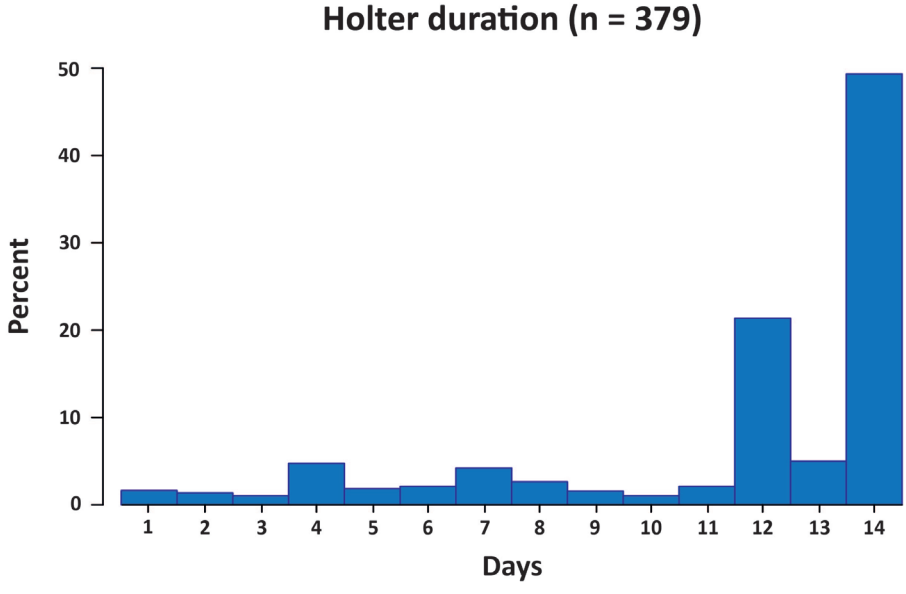
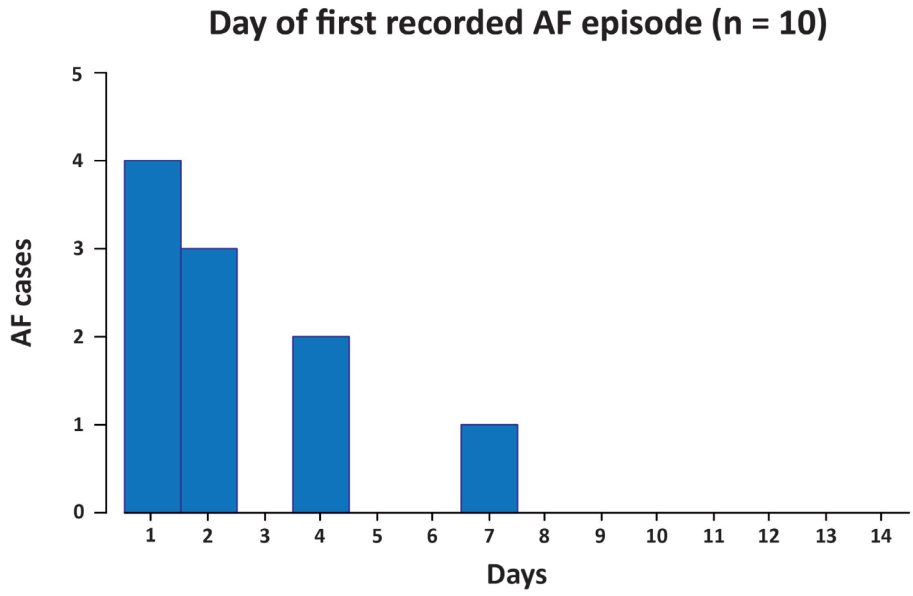


Figure 4. *Day of first AF recorded in those with AF on Holter (n = 10). AF: atrial fibrillation.*



AF, atrial fibrillation.

Baseline characteristics and AF detection

There were a number of notable differences in baseline characteristics between patients who were and who were not diagnosed with AF on 14-day Holter (Supplementary Table S2). These included higher age (median 69.5 vs 63.0 years), higher rate of IAT (40.0% vs 6.2%) or IVT (50.0% vs 18.7%) on admission, higher rate of diuretics use (40.0% vs 12.5%), and lower serum triglycerides (median 0.90 vs 1.36 mmol/l) among those with and those without AF on 14-day monitor. Time from index stroke/TIA onset to Holter initiation was similar in both groups (median 45 days, IQR: 8–60, among those with and 35 days, IQR: 14–59, in those without AF).

Patients with Holter duration >48 h

There were 362 patients who had worn their Holter >48 h and without documented AF during the first 2 days of monitoring (95.5% of the study sample). During the following median 12 days (IQR: 10–12), the study Holter detected three additional AF cases (0.83%, 95% CI 0.17–2.42; NNS 121). Baseline characteristics of patients with >48 h of Holter recording available are shown in Supplementary Table S3. Notable differences were seen in age (median 70.0 and 63.0 years), NIHSS (median 18 vs 1), proportion of patients who underwent IAT at presentation (66.7% vs 5.6%), and family history of AF (66.7% vs 12.4%) in those with and those without AF on Holter beyond the first 48 h, respectively.

Subgroup analysis and prediction models

Among patients with ischemic stroke as qualifying event 9/261 had an AF during 14-day monitoring (incidence 3.45%, 95% CI 1.58–6.55; NNS 29). In patients with non-lacunar hemispheric stroke/TIA 9/212 had AF (incidence 4.25%, 95% CI 1.94–8.06; NNS 24), and among those with non-lacunar hemispheric stroke (not TIA) 8/131 showed AF on 14-day Holter (incidence 6.11%, 95% CI 2.64–12.03; NNS 16). Among patients with moderate-severe stroke at presentation 5/32 were diagnosed with AF during 14-day monitoring (incidence 15.63%, 95% CI 5.07–36.46; NNS 6).

C-statistics (95% CI) of the AS5F, Re-CHARGE-AF, and STAF risk models for 14-day AF detection were 0.75 (0.61–0.89), 0.59 (0.43–0.75), and 0.70 (0.56–0.84), respectively.

DISCUSSION

In low-risk patients with recent ischemic stroke or TIA attending an academic hospital in The Netherlands, screening for AF using 14-day Holter resulted in an overall AF yield of 2.6%, with an NNS of 38. Extending Holter monitoring beyond the ESO guideline-recommended 48 h for post-stroke patients resulted in 0.83% of new AF cases (NNS 121), with no new AF cases detected after the first 7 days of ambulatory monitoring within our study sample. Selection of higher-risk patients to screen for AF using clinically relevant criteria would have increased AF yield. Our results underline the need for a personalized approach in determining a patient's optimum duration for post-stroke non-invasive ambulatory monitoring.

Comparison with previous work

The incidence of device-detected AF during 14-day Holter monitoring in our sample was considerably lower than generally found in previous studies of prolonged ambulatory Holter monitoring among unselected patients with recent stroke/TIA.[5,6,23,24] Even among patients further selected for cryptogenic stroke, thus having undergone at least 24 h of continuous ECG monitoring before further monitoring, AF yield was generally higher in the first 14 days of recording than in our sample.[25-29] Studies that reported Holter-detected post-stroke AF yields similar to ours were generally performed with 24-h rather than 14-day Holter[30-32] or involved retrospective rather than prospective data.[33]

Potentially the main reason for the low AF yield was sampling bias toward inclusion of patients with a lower overall likelihood of post-stroke AF detection. The logistics of our inclusion process, with the requirement of active consent before Holter installa-

tion and with inclusion in part through a study outpatient clinic, has likely contributed to the undersampling of patients who attended our hospital only for IAT (short admission time) or with severe stroke (barriers to attending the study outpatient clinic) who are known to have a higher likelihood of post-stroke AF detection.[34] As shown in our comparison between included patients and the random sample of eligible non-included patients, this resulted in a sample of patients who were younger, with less severe strokes and lower overall cardiovascular risk factor burden than the overall stroke/TIA population. Since AF is associated with increased stroke severity,[35,36] and studies have shown clinical stroke scores to be associated with post-stroke AF detection,[34] inclusion of lower-NIHSS patients could have led to a lower proportion of patients with AF in our study than in the overall stroke/TIA population. Likewise, IAT, an intervention to remove large thrombi from the intracranial anterior circulation, is typically performed in patients with higher stroke severity.[37,38] It can be theorized that such thrombi are more frequently associated with AF, whether directly through AF-associated cardiac emboli or as a result of a cardiovascular risk profile that increases the risk of both large thrombi and atrial cardiomyopathy, including AF.[39] It is therefore likely that the AF yield in our sample cannot be extrapolated to reflect that of the general Dutch stroke/TIA population.

In previous AF screening studies of unselected post-stroke/TIA patients participants were generally older.[23,40-42] With age being one of the most important risk factors for AF incidence,1 this is likely to explain to a large degree the higher AF incidence seen in these studies. The proportion of TIA versus stroke as qualifying event as well as overall NIHSS in our sample were similar compared to previous unselected TIA/stroke studies on post-stroke AF.[40-42] As these studies did not provide a comparison with potentially eligible non-included patients it is not known to what extent their baseline characteristics were affected by sampling bias.

Another possible explanation for the low AF incidence is the time from the qualifying event to commencement of Holter monitoring. Although data suggesting this association is limited, time to Holter initiation is likely related to the probability of AF detection after stroke.[2,43-45] Research on the Stroke-Heart Syndrome (SHS) has indicated that post-stroke major adverse cardiovascular events and AF peak in the first 3–30 days after stroke onset.[45] Our data showed no significant difference in time to Holter between patients with and without detected AF as reported in previous work.[28,46] However, with median time to Holter of 35 days in our sample, we largely included patients outside the peak window for potential SHS cases. Our time to monitoring initiation far exceeded that of other studies which often had timely monitoring commencement in their inclusion criteria.[23,40] These and other post-stroke studies

with shorter mean duration to Holter have shown higher rates of AF detection.[27,42,46] Our study excluded a number of patients with AF detected during admission for their stroke/TIA as de novo AF cases. As such cases may have been included in previous studies on post-stroke Holter yield, which would hinder a comparison with our work, we explicitly reported time of de novo AF detection in our Flowchart for comparison purposes. It is not known to what extent a further reduction in the time to Holter would have resulted in a higher AF yield in our population, or whether the sampling bias alone sufficiently accounts for the lower AF yield in our sample. In line with work on the SHS and its peak in the first 30 days, one could even speculate that the higher AF rate in studies with monitoring initiation directly after stroke were higher due to transient stroke-induced cardiomyopathy which we "missed" with mean 35 days to Holter initiation.[45] A post-hoc sensitivity analysis applying inclusion criteria of previous unselected stroke/TIA AF screening studies to our data showed no significant increase in AF yield compared to our overall sample (data not shown),[23,40] however sample size and event rate were severely limited in these analyses as in the overall analysis. Thus, the question whether every patient with stroke or TIA would benefit from rhythm monitoring directly following their cerebral event – which is currently not routinely performed in our center for logistical reasons – remains unanswered from our data.

A final explanation for our low AF yield is the proportion of patients who already had a known AF diagnosis at time of assessment for study Holter eligibility – a quarter of whom had de novo AF in our study. With pre-Holter AF prevalence of 41% in our stroke population this was considerably higher than in previous post-stroke AF screening studies that reported prior AF as reason for exclusion (generally below 20%).[30,32,47-50] This potentially indicates that, in the presence of an already high proportion of post-stroke patients with (a history of) documented AF, prolonged rhythm monitoring has a relatively low additional yield. This view is supported by a recently published randomized controlled AF opportunistic screening trial among Dutch primary care patients of 65 years and over. The intervention did not achieve higher AF yield than usual care over a 1-year period, mainly underscoring the efficacy of detecting AF in routine primary care in the Netherlands.[11,51]

Our work subscribed to previous work which indicated that stroke location among patients with AF on post-stroke monitoring was most often non-lacunar hemispheric, with none or very few AF cases among patients with lacunar stroke/TIA.[6,41] We note here that stroke location in our data was primarily based on clinical symptoms as MRI was not routinely performed in all stroke/TIA patients. As in previous work, we saw most AF cases detected during the first days of monitoring.[52] Our data also concur with

a recent systematic review and meta-analysis which found an association between IVT treatment, higher age, and lower triglycerides with AF detection.[34] The yield of post stroke rhythm monitoring may thus increase when clinical risk factors for AF are taken into account.

Clinical implications

Our findings are relevant for neurologists in similar settings who aim to optimize efficacy of their post-stroke rhythm monitoring strategy in low-risk stroke/TIA patients. The current data show that the yield of 48-h ambulatory Holter monitoring in a low-risk post-stroke patients of a Dutch academic hospital was lower than expected based on recent international literature. The additional value of monitoring beyond 48 h or 72 h as recommended by the ESC and similarly by the Dutch Neurological Society, was even more limited.[1,53] Given that we detected a minority of cases within the first 24 h, 24-h monitoring as currently recommended by The National Institute for Health and Care Excellence (NICE) may be too short for post-stroke AF diagnosis.[10] It is not known to what extent the more recent NICE diagnostic guidance to consider implanting implantable cardiac monitors in cryptogenic stroke patients will contribute to the detection of occult AF after brain ischemia.[54]

Recent publications have emphasized that the optimal screening strategy for AF after stroke/TIA is yet to be determined.[55] The current work underlines the potential for a more personalized approach than the current recommendation to screen all stroke/TIA patients for more than 48 h.[2] While the question of risk stratification is often viewed from the perspective of identifying those at highest risk (safety driven), we now add a low-risk perspective: are there low-risk patients whom we can spare potentially costly and burdensome prolonged monitoring beyond the guideline-recommended minimum? Given the limitations of our work, our data are especially relevant for those at lowest risk of AF, and in those who are not able to commence monitoring immediately after symptom onset. Our data indicated that 14-day monitoring in low-risk stroke/TIA patients results in surprisingly low AF yields, while selecting for clinically relevant risk factors increases AF detection rates considerably. Due to the low overall AF yield and relatively low number of patients in our sample, we were unable to provide definitive answers to this question. Still, our data on AF yield in clinically relevant subgroups as well as our validation of risk models for post-stroke AF could be combined with that of other observational studies in order to increase our understanding of optimal screening strategies for AF detection after stroke or TIA. We emphasize, however, that it is ultimately up to physicians and other stakeholders in each particular care setting to decide whether the numbers needed to screen as reflected by our and previous studies are deemed sufficiently (cost-)efficient in their situation.

Our data underscore the need for a reliable triage test to identify patients in whom prolonged rhythm monitoring after TIA or stroke is associated with a fair chance of capturing AF. Given the low apparent yield in low-risk stroke/TIA patients, but with uncertainty around the optimal use of biomarkers as triage test for prolonged monitoring,[2] further research could focus on strategies to use clinical parameters to select for prolonged monitoring. Depending on the available resources and expected burden to the patient of wearing extended ambulatory ECG monitoring, clinicians can use such work to decide on whether to extend monitoring duration in their particular patients.

Strengths and limitations

A strength of our work is the prolonged Holter monitoring duration up to 14 days, which far exceeds the currently recommended continuous ECG recording duration for post-stroke patients. Moreover, there was a high rate of patient compliance with a large majority of patients wearing their Holter for 12 days or more. A further strength of our study was the relatively broad inclusion criteria compared to other work that focused ECG monitoring strategies on subgroups of patients, most notably patients with cryptogenic stroke or embolic stroke of undetermined source.[14,25,26,52] This allowed for a closer validation of current guidelines whose recommendations for rhythm monitoring regard all ischemic stroke and TIA cases with undetermined origin.[1,2,8] Another strength is our detailed documentation of reasons for exclusion to the study. This enabled us to demonstrate that almost a quarter of all stroke presentations arrived in our hospital with a known AF diagnosis at time of presentation, which potentially provided further context to the relatively low AF incidence during Holter monitoring. However, by carefully excluding patients with a history of AF or with de novo AF detected during comprehensive clinical and early outpatient clinical observation, our study allows for the assessment of truly new-onset AF. A final strength was the presentation of selected baseline characteristics among a random sample of non-included eligible patients which allowed a better assessment of the extent of sampling bias within our study.

The primary limitation of our study was sampling bias, resulting in a study population with lower risk of AF than the overall ischemic stroke/TIA population. The study's logistic limitations as described above lowered the likelihood of severe stroke or tertiary care IAT patients to be included in our sample. To address this issue we presented limited baseline data from a random sample of non-included study-eligible patients in order to better understand the extent of sampling bias in our sample, which was considerable. Due to limitations imposed by the European Union's GDPR we were unable to compare complete baseline characteristics of non-included eligible patients with those included in our study.[56] The low AF incidence in our sample, as well as

limitations to the scope of our dataset (e.g. incomplete data on MRI for stroke location particularly insular cortex, echocardiography for presence of patent foramen ovale, or biomarkers such as cardiac troponin) impaired our ability to assess the significance of risk factors and biomarkers associated with post-stroke AF detection.[34,45,57] By still presenting baseline data stratified by AF detection during Holter monitoring, we aimed to contribute to potential future work on personalized monitoring approaches.[2] While adherence to our 14-day study design was high among included patients, recent evidence points to the superiority implantable devices in detecting silent AF, which is reflected in the current ESO guideline's recommendations.[2] The use of 14-day Holter was thus a limitation in comparison to AF screening studies that employ implantable loop recorders, and potentially contributed to the low AF yield in our low-risk sample. The use of 2-lead Holter monitors in our study has been shown not to be associated with lower AF yield during post-stroke monitoring than 3- or 6-lead ambulatory monitors.[6] Finally, our study does not contain follow-up for outcomes after AF detection, and is therefore unable to contribute to the work on stroke recurrence in post-stroke AF patients.[55]

CONCLUSION

In conclusion, in low-risk patients with recent stroke or TIA, ESO guideline-recommended screening for AF resulted in a low AF yield, with limited additional value of monitoring up to 14 days. Our results underline the need for a personalized approach in determining a patient's optimum duration for post-stroke non-invasive ambulatory monitoring.

Acknowledgments: The authors wish to thank all participating patients, as well as Maxim Annink, Erik de Bruin, Guido Clerx, Nina Groeneveld, Josje Mangnus, Janice Mitsinga, and Ibtisam Yahya for their support in patient enrollment.

Declaration of conflicting interests: The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: JCLH has received a grant from Stichting Stoffels-Hornstra and consultancy fees from Thrombosis Research Institute (outside the submitted work); WAML reports no disclosures relevant to the manuscript; JMC reported receiving grants from Dutch Heart Foundation and from Boehringer Ingelheim (outside the submitted work); REH reported receiving grants from the Dutch Research Council (outside the submitted work); JRdG has been supported by research grants from Abbott, AtriCure, Boston Scientific, Bayer, Daiichi Sankyo, Johnson & Johnson, and Medtronic Servier and has

received consultancy fees from AtriCure, Bayer, Daiichi Sankyo, Johnson & Johnson, and Medtronic (outside the submitted work); HCPMvW served on the editorial board of the European Journal of General Practice, 2012-2021.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Netherlands Organization for Health Research and Development (ZonMw) [80-83910-98-13046]. The authors had full autonomy in design, conduct, and reporting of the manuscript.

Ethical approval: The study was granted a waiver for the requirement of informed consent under the Medical Research Involving Human Subjects Act (WMO) by the Medical Review Board of the AUMC-AMC on the grounds that post-stroke Holter monitoring is regarded as standard of post-stroke care (AMC METC No. W16_168). All included patients, however, provided written permission for use of their deidentified routine care EHR and Holter data as well as study questionnaires for research purposes under the General Data Protection Regulation (GDPR).

Informed consent: See under "Ethical approval".

Guarantor: JCLH

Author contributions: JCLH was principally responsible for conceptualization of the study as well as acquisition, analysis and interpretation of the data, and was responsible for the drafting and revision of the manuscript. WAML and HCPMvW designed and conceptualized the study and provided valuable input to the manuscript. JMC and JRdG contributed to design and conceptualization of the study as well as to acquisition, analysis and interpretation of the data, and provided valuable input to the manuscript. REH provided valuable input to analysis and interpretation of the data, and provided valuable input to the manuscript.

Trial registration: Netherlands Trial Register, NTR6489.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLES

Supplementary Table S1. *Comparison of selected baseline characteristics between the study sample (n = 379) and a 25% random sample of potentially eligible, non-included patients (n = 169)*

	Study sample (All, n=379)	Eligible non-included (25% RS, n=169)
Female sex (%)	42.2	49.7
Age, years (Median [Q1-Q3])	63.0 (55.0-73.0)	71.0 (60.0-80.0)
Qualifying event		
Ischemic stroke (%)	68.9	81.1
TIA (%)	31.1	18.9
Stroke/TIA location		
Middle cerebral artery (%)	44.6	60.4
Anterior cerebral artery (%)	1.8	1.2
Posterior cerebral artery (%)	5.3	2.4
Lacunar (%)	12.4	13.0
Retinal (%)	4.2	3.0
Vertebrobasilar (%)	31.7	20.1
Non-lacunar hemispheric stroke (%)	34.6	56.8
NIHSS at first presentation (Median [Q1-Q3])	1 (0-3)	3 (1-9)
Intravenous thrombolysis (%)	19.5	34.3
Intra-arterial thrombectomy (%)	7.1	17.2
Heart failure (%)	2.6	0
Hypertension (%)	46.4	45.6
Diabetes (%)	18.7	19.5
Prior myocardial infarction (%)	7.7	9.5
Prior stroke/TIA/SE (%)	21.9	31.4
Vascular disease (%)	12.9	21.9
CHA_2DS_2-VASc (Median [Q1-Q3])	4 (3-5)	5 (3-6)
Antiplatelet use (%)	34.3	39.1
ACE/ARB use (%)	28.8	31.4
Calcium antagonist use (%)	18.5	17.8
Diuretics use (%)	13.2	21.3
Statin use (%)	36.9	38.5
Insulin use (%)	5.3	5.9
Metformin use (%)	12.9	14.8

ACE/ARB, angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker; AF, atrial fibrillation; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; NIHSS, National Institutes of Health Stroke Scale; Q1, 1st quartile; Q3, 3rd quartile; RS, random sample; SE, systemic embolism; TIA, transient ischemic attack.

Supplementary Table S2. *Extended baseline characteristics of the total study sample (n = 379) stratified by AF presence on study Holter*

	AF (n = 10)	No AF (n = 369)
Female sex (n,%)	4 (40.0%)	156 (42.3%)
Age, years (Median [Q1-Q3])	69.5 (63.8-71.8)	63.0 (55.0-73.0)
Ethnicity		
Caucasian/white (n,%)	10 (100%)	251 (68.0%)
African/black (n,%)	0 (0%)	71 (19.2%)
South Asian (n,%)	0 (0%)	19 (5.1%)
Asian (n,%)	0 (0%)	18 (4.9%)
Other (n,%)	0 (0%)	10 (3.3%)
Qualifying event		
Ischemic stroke (n,%)	9 (90.0%)	252 (68.3%)
TIA (n,%)	1 (10.0%)	117 (31.7%)
Stroke/TIA location		
Middle cerebral artery (n,%)	8 (80.0%)	161 (43.6%)
Anterior cerebral artery (n,%)	0 (0%)	7 (1.9%)
Posterior cerebral artery (n,%)	1 (10.0%)	19 (5.1%)
Lacunar (n,%)	0 (0%)	47 (12.7%)
Retinal (n,%)	0 (0%)	16 (4.3%)
Vertebrobasilar (n,%)	1 (10.0%)	119 (32.2%)
Non-lacunar hemispheric stroke (n,%)	8 (80.0%)	123 (33.3)
NIHSS at first presentation (Median [Q1-Q3])	7 (2-13)	1 (0-3)
Ipsilateral carotid artery stenosis >50% (n,%)	1 (10.0%)	9 (2.4%)
Intravenous thrombolysis (n,%)	5 (50.0%)	69 (18.7%)
Intra-arterial thrombectomy (n,%)	4 (40.0%)	23 (6.2%)
Time to Holter, days (Median [Q1-Q3])	45 (8-60)	35 (14-59)
Time to Holter ≤90 days (n,%)	9 (90.0%)	328 (88.9%)
Holter duration, days (Median [Q1-Q3])	14 (12-14)	13 (12-14)
BMI, kg/m^2 (Median [Q1-Q3])	25.3 (22.5-30.0)	26.0 (24.0-29.1)
Missing (n,%)	0 (0%)	2 (0.5%)
SBP, mm Hg (Median [Q1-Q3])	156 (138-166)	154 (135-172)
Missing (n,%)	1 (10.0%)	21 (5.7%)
DBP, mm Hg (Median [Q1-Q3])	78 (70-94)	87 (77-99)
Missing (n,%)	1 (10.0%)	21 (5.7%)
Smoking		
Current smoker (n,%)	2 (20.0%)	76 (20.6%)
Never smoked (n,%)	2 (20.0%)	134 (36.3%)
Former smoker (n,%)	6 (60.0%)	159 (43.1%)
Pack years, years (Median [Q1-Q3])	22 (4-45)	7 (0-24)
Missing (n,%)	0 (0%)	8 (2.2%)
Alcohol units/day (Median [Q1-Q3])	2 (1-2)	0 (0-1)
Family history of AF		

Supplementary Table S2. *Extended baseline characteristics of the total study sample (n = 379) stratified by AF presence on study Holter (continued)*

	AF (n = 10)	No AF (n = 369)
Yes (n,%)	3 (30.0%)	45 (12.2%)
No (n,%)	6 (60.0%)	254 (68.8%)
Unknown (n,%)	1 (10.0%)	70 (19.0%)
Heart failure (n,%)	0 (0%)	10 (2.7%)
Hypertension (n,%)	7 (70.0%)	169 (45.8%)
Diabetes (n,%)	0 (0%)	71 (19.2%)
Prior myocardial infarction (n,%)	1 (10.0%)	28 (7.6%)
Prior stroke/TIA/SE (n,%)	3 (30.0%)	80 (21.7%)
Vascular disease (n,%)	1 (10.0%)	48 (13.0%)
Prior cardiac intervention (n,%)	0 (0%)	31 (8.4%)
Asthma or COPD (n,%)	1 (10.0%)	27 (7.3%)
Chronic kidney disease	1 (10.0%)	16 (4.3%)
CHA_2DS_2-VASc (Median [Q1-Q3])	4 (3-5)	4 (3-5)
AS5F (Median [Q1-Q3])	65.9 (61.8-73.4)	58.4 (51.6-65.2)
Re-CHARGE-AF (Median [Q1-Q3])	0.129 (0.095-0.168)	0.112 (0.061-0.187)
Missing (n,%)	1 (10.0%)	23 (6.2%)
STAF (Median [Q1-Q3])	6 (5-6)	5 (3-5)
Missing (n,%)	0 (0%)	22 (6.0%)
Antiplatelet use (n,%)	4 (40.0%)	126 (34.1%)
ACE/ARB use (n,%)	4 (40.0%)	105 (28.5%)
Calcium antagonist use (n,%)	1 (10.0%)	69 (18.7%)
Diuretics use (n,%)	4 (40.0%)	46 (12.5%)
Statin use (n,%)	4 (40.0%)	136 (36.9%)
Insulin use (n,%)	0 (0%)	20 (5.4%)
Metformin use (n,%)	0 (0%)	49 (13.3%)
eGFR, mL/min/1.73m^2 (Median [Q1-Q3])	68 (55-81)	77 (63-88)
Missing (n,%)	0 (0%)	7 (1.9%)
LDL cholesterol, mmol/L (Median [Q1-Q3])	2.32 (2.27-2.95)	2.65 (1.93-3.42)
Missing (n,%)	2 (20.0%)	48 (13.0%)
Triglycerides, mmol/L (Median [Q1-Q3])	0.90 (0.78-1.32)	1.36 (0.90-1.94)
Missing (n,%)	2 (20.0%)	48 (13.0%)

ACE/ARB, angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker; AF, atrial fibrillation; AS5F, Age, Stroke Severity NIHSS >5 to Find AF; BMI, body mass index; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; COPD, chronic obstructive pulmonary disease; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; LDL, low-density lipoprotein; NIHSS, National Institutes of Health Stroke Scale; Q1, 1st quartile; Q3, 3rd quartile; Re-CHARGE-AF, Recalibrated Cohorts for Heart and Aging Research in Genomic Epidemiology for Atrial Fibrillation; SBP, systolic blood pressure; SE, systemic embolism; STAF, Score for the Targeting of Atrial Fibrillation; TIA, transient ischemic attack.
In case of missing data in each variable, number and percentage of missing are indicated.

Supplementary Table S3. *Extended baseline characteristics of patients with over 48 hours of Holter recording and free of AF at 48 hours of monitoring (n = 362) stratified by AF presence on subsequent Holter*

	AF (n = 3)	No AF (n = 359)
Female sex (n,%)	1 (33.3%)	152 (42.3%)
Age, years (Median [Q1-Q3])	70.0 (69.5-71.0)	63.0 (55.0-73.0)
Ethnicity		
Caucasian/white (n,%)	3 (100%)	245 (68.2%)
African/black (n,%)	0 (0%)	68 (19.1%)
South Asian (n,%)	0 (0%)	19 (5.3%)
Asian (n,%)	0 (0%)	18 (5.1%)
Other (n,%)	0 (0%)	9 (2.5%)
Qualifying event		
Ischemic stroke (n,%)	3 (100%)	244 (68.0%)
TIA (n,%)	0 (0%)	115 (32.0%)
Stroke/TIA location		
Middle cerebral artery (n,%)	3 (100%)	158 (44.0%)
Anterior cerebral artery (n,%)	0 (0%)	7 (1.9%)
Posterior cerebral artery (n,%)	0 (0%)	18 (5.1%)
Lacunar (n,%)	0 (0%)	44 (12.4%)
Retinal (n,%)	0 (0%)	16 (4.5%)
Vertebrobasilar (n,%)	0 (0%)	116 (32.6%)
Non-lacunar hemispheric stroke (n,%)	3 (100%)	120 (33.4%)
NIHSS at first presentation (Median [Q1-Q3])	18 (10-20)	1 (0-3)
Ipsilateral carotid artery stenosis >50% (n,%)	0 (0%)	9 (2.5%)
Intravenous thrombolysis (n,%)	2 (66.7%)	67 (18.7%)
Intra-arterial thrombectomy (n,%)	2 (66.7%)	20 (5.6%)
Time to Holter, days (Median [Q1-Q3])	63 (32-70)	36 (15-60)
Time to Holter ≤90 days (n,%)	3 (100%)	319 (88.9%)
Holter duration, days (Median [Q1-Q3])	14 (14-14)	14 (12-14)
BMI, kg/m^2 (Median [Q1-Q3])	21.9 (21.5-23.4)	26.0 (24.0-29.1)
Missing (n,%)	0 (0%)	2 (0.6%)
SBP, mm Hg (Median [Q1-Q3])	148 (143-153)	155 (136-171)
Missing (n,%)	1 (33.3%)	21 (5.8%)
DBP, mm Hg (Median [Q1-Q3])	70 (67-72)	88 (77-99)
Missing (n,%)	1 (33.3%)	21 (5.8%)
Smoking		
Current smoker (n,%)	1 (33.3%)	72 (20.1%)
Never smoked (n,%)	1 (33.3%)	131 (36.5%)
Former smoker (n,%)	1 (33.3%)	156 (43.5%)
Alcohol units/day (Median [Q1-Q3])	2 (1-2)	0 (0-1)
Family history of AF		

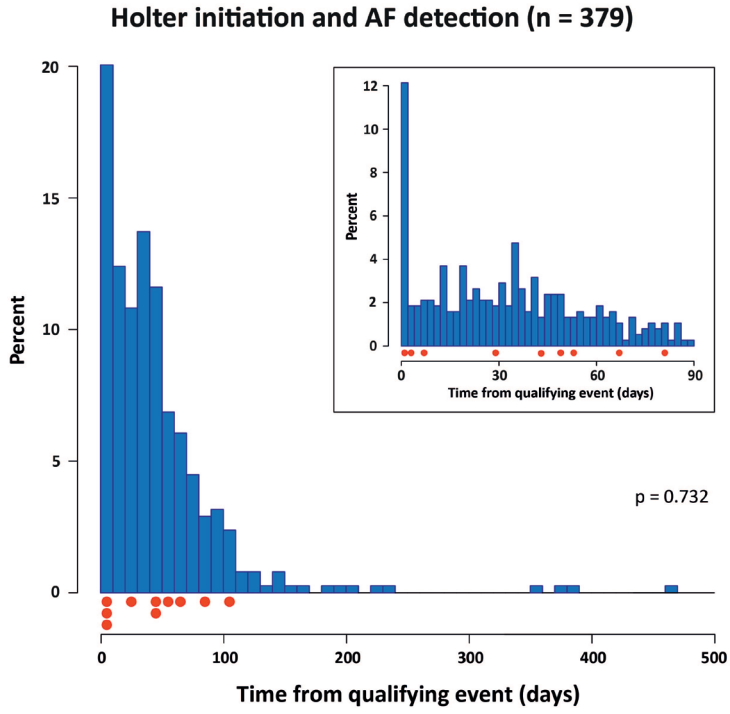
Supplementary Table S3. *Extended baseline characteristics of patients with over 48 hours of Holter recording and free of AF at 48 hours of monitoring (n = 362) stratified by AF presence on subsequent Holter (continued)*

	AF (n = 3)	No AF (n = 359)
Yes (n,%)	2 (66.7%)	44 (12.3%)
No (n,%)	1 (33.3%)	248 (69.1%)
Unknown (n,%)	0 (0%)	67 (18.7%)
Heart failure (n,%)	0 (0%)	9 (2.5%)
Hypertension (n,%)	2 (66.7%)	168 (46.8%)
Diabetes (n,%)	0 (0%)	69 (19.2%)
Prior myocardial infarction (n,%)	0 (0%)	26 (7.2%)
Prior stroke/TIA/SE (n,%)	0 (0%)	79 (22.0%)
Vascular disease (n,%)	0 (0%)	46 (12.8%)
Prior cardiac intervention (n,%)	0 (0%)	29 (8.1%)
Asthma or COPD (n,%)	1 (10.0%)	25 (7.0%)
Chronic kidney disease	1 (10.0%)	16 (4.3%)
CHA_2DS_2-VASc (Median [Q1-Q3])	4 (4-5)	4 (3-5)
AS5F (Median [Q1-Q3])	74.2 (67.8-75.0)	58.4 (51.6-65.2)
Re-CHARGE-AF (Median [Q1-Q3])	0.109 (0.102-0.116)	0.112 (0.062-0.186)
Missing (n,%)	1 (33.3%)	23 (6.4%)
STAF (Median [Q1-Q3])	6 (6-7)	5 (3-5)
Missing (n,%)	0 (0%)	20 (5.6%)
Antiplatelet use (n,%)	0 (0%)	123 (34.3%)
ACE/ARB use (n,%)	1 (33.3%)	104 (29.0%)
Calcium antagonist use (n,%)	0 (0%)	68 (18.9%)
Diuretics use (n,%)	2 (66.7%)	46 (12.8%)
Statin use (n,%)	2 (66.7%)	133 (37.0%)
Insulin use (n,%)	0 (0%)	19 (5.3%)
Metformin use (n,%)	0 (0%)	48 (13.4%)
eGFR, mL/min/1.73m^2 (Median [Q1-Q3])	52 (50-61)	77 (63-88)
Missing (n,%)	0 (0%)	7 (1.9%)
LDL cholesterol, mmol/L (Median [Q1-Q3])	2.32 (2.31-2.32)	2.65 (1.90-3.42)
Missing (n,%)	1 (33.3%)	46 (12.8%)
Triglycerides, mmol/L (Median [Q1-Q3])	0.77 (0.72-0.82)	1.36 (0.90-1.94)
Missing (n,%)	1 (33.3%)	46 (12.8%)

ACE/ARB, angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker; AF, atrial fibrillation; AS5F, Age, Stroke Severity NIHSS >5 to Find AF; BMI, body mass index; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; COPD, chronic obstructive pulmonary disease; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; LDL, low-density lipoprotein; NIHSS, National Institutes of Health Stroke Scale; Q1, 1st quartile; Q3, 3rd quartile; Re-CHARGE-AF, Recalibrated Cohorts for Heart and Aging Research in Genomic Epidemiology for Atrial Fibrillation; SBP, systolic blood pressure; SE, systemic embolism; STAF, Score for the Targeting of Atrial Fibrillation; TIA, transient ischemic attack.
In case of missing data in each variable, number and percentage of missing are indicated.

SUPPLEMENTARY FIGURES

Supplementary Figure S1. *Association between Holter initiation and day of first AF detection after qualifying event (n = 379)*



AF, atrial fibrillation.
Blue bars indicate percentage of included patients with Holter initiated after n days of the qualifying event. Red dots indicate the days from qualifying event at which each AF case was first detected (n = 10). P-value indicates the association between time from qualifying event to Holter and presence of AF on study Holter.

REFERENCES

1. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). Eur Heart J. 2020;0:1-125.
2. Rubiera M, Aires A, Antonenko K, Lemeret S, Nolte CH, Putaala J, et al. European Stroke Organisation (ESO) guideline on screening for subclinical atrial fibrillation after stroke or transient ischaemic attack of undetermined origin. European Stroke Journal. 2022;7(3):CVII–CXXXIX.
3. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke. 1991;22(8):983-8.
4. Marini C, De Santis F, Sacco S, Russo T, Olivieri L, Totaro R, et al. Contribution of atrial fibrillation to incidence and outcome of ischemic stroke - Results from a population-based study. Stroke. 2005;36(6):1115-9.
5. Sposato LA, Cipriano LE, Saposnik G, Ruiz Vargas E, Riccio PM, Hachinski V. Diagnosis of atrial fibrillation after stroke and transient ischaemic attack: a systematic review and meta-analysis. Lancet Neurol. 2015;14(4):377-87.
6. Kishore A, Vail A, Majid A, Dawson J, Lees KR, Tyrrell PJ, et al. Detection of atrial fibrillation after ischemic stroke or transient ischemic attack: a systematic review and meta-analysis. Stroke. 2014;45(2):520-6.
7. Fonseca AC, Merwick A, Dennis M, Ferrari J, Ferro JM, Kelly P, et al. European Stroke Organisation (ESO) guidelines on management of transient ischaemic attack. Eur Stroke J. 2021;6(2):CLXIII-CLXXXVI.
8. Kleindorfer DO, Towfighi A, Chaturvedi S, Cockroft KM, Gutierrez J, Lombardi-Hill D, et al. 2021 Guideline for the Prevention of Stroke in Patients With Stroke and Transient Ischemic Attack: A Guideline From the American Heart Association/American Stroke Association. Stroke. 2021;52(7):e364-e467.
9. Ahmed N, Audebert H, Turc G, Cordonnier C, Christensen H, Sacco S, et al. Consensus statements and recommendations from the ESO-Karolinska Stroke Update Conference, Stockholm 11-13 November 2018. Eur Stroke J. 2019;4(4):307-17.
10. National Institute for Health and Care Excellence. Atrial fibrillation: diagnosis and management (NICE guideline NG196). 2021. https://www.nice.org.uk/guidance/ng196. p. 1-43.
11. Uittenbogaart SB, Verbiest-van Gurp N, Lucassen WAM, Winkens B, Nielen M, Erkens PMG, et al. Opportunistic screening versus usual care for detection of atrial fibrillation in primary care: cluster randomised controlled trial. BMJ. 2020;370:m3208.
12. Brott T, Adams HP, Jr., Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke. 1989;20(7):864-70.
13. Yaghi S, Raz E, Yang D, Cutting S, Mac Grory B, Elkind MS, et al. Lacunar stroke: mechanisms and therapeutic implications. J Neurol Neurosurg Psychiatry. 2021;92:823–30.
14. Hart RG, Diener H-C, Coutts SB, Easton JD, Granger CB, O'Donnell MJ, et al. Embolic strokes of undetermined source: the case for a new clinical construct. The Lancet Neurology. 2014;13(4):429-38.

15. Adams HP, Jr., Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke. 1993;24(1):35-41.
16. 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(2):263-72.
17. Calkins H, Hindricks G, Cappato R, Kim YH, Saad EB, Aguinaga L, et al. 2017 HRS/EHRA/ECAS/APHRS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation: Executive summary. Heart Rhythm. 2017;14(10):e445-e94.
18. Rost NS, Bottle A, Lee JM, Randall M, Middleton S, Shaw L, et al. Stroke Severity Is a Crucial Predictor of Outcome: An International Prospective Validation Study. Journal of the American Heart Association. 2016;5(1):1-7.
19. Jackson C, Sudlow C. Are lacunar strokes really different? A systematic review of differences in risk factor profiles between lacunar and nonlacunar infarcts. Stroke. 2005;36(4):891-901.
20. Uphaus T, Weber-Kruger M, Grond M, Toenges G, Jahn-Eimermacher A, Jauss M, et al. Development and validation of a score to detect paroxysmal atrial fibrillation after stroke. Neurology. 2019;92(2):e115-e24.
21. Suissa L, Bertora D, Lachaud S, Mahagne MH. Score for the targeting of atrial fibrillation (STAF): a new approach to the detection of atrial fibrillation in the secondary prevention of ischemic stroke. Stroke. 2009;40(8):2866-8.
22. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
23. Wachter R, Groschel K, Gelbrich G, Hamann GF, Kermer P, Liman J, et al. Holter-electrocardiogram-monitoring in patients with acute ischaemic stroke (Find-AFRANDOMISED): an open-label randomised controlled trial. Lancet Neurol. 2017;16(4):282-90.
24. Korompoki E, Del Giudice A, Hillmann S, Malzahn U, Gladstone DJ, Heuschmann P, et al. Cardiac monitoring for detection of atrial fibrillation after TIA: A systematic review and meta-analysis. Int J Stroke. 2017;12(1):33-45.
25. Gladstone DJ, Spring M, Dorian P, Panzov V, Thorpe KE, Hall J, et al. Atrial fibrillation in patients with cryptogenic stroke. N Engl J Med. 2014;370(26):2467-77.
26. Sanna T, Diener HC, Passman RS, Di Lazzaro V, Bernstein RA, Morillo CA, et al. Cryptogenic stroke and underlying atrial fibrillation. N Engl J Med. 2014;370(26):2478-86.
27. Haeusler KG, Kirchhof P, Kunze C, Tutuncu S, Fiessler C, Malsch C, et al. Systematic monitoring for detection of atrial fibrillation in patients with acute ischaemic stroke (MonDAFIS): a randomised, open-label, multicentre study. Lancet Neurol. 2021;20(6):426-36.
28. Miller DJ, Khan MA, Schultz LR, Simpson JR, Katramados AM, Russman AN, et al. Outpatient cardiac telemetry detects a high rate of atrial fibrillation in cryptogenic stroke. J Neurol Sci. 2013;324(1-2):57-61.
29. Haeusler KG, Tutuncu S, Schnabel RB. Detection of Atrial Fibrillation in Cryptogenic Stroke. Curr Neurol Neurosci Rep. 2018;18(10):66.
30. 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(2):253-7.

31. 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(8):1112-6.
32. Rizos T, Guntner 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(10):2689-94.
33. Liran O, Banon T, Grossman A. Detection of occult atrial fibrillation with 24-hour ECG after cryptogenic acute stroke or transient ischaemic attack: A retrospective cross-sectional study in a primary care database in Israel. Eur J Gen Pract. 2021;27(1):152-7.
34. Cameron A, Cheng HK, Lee RP, Doherty D, Hall M, Khashayar P, et al. Biomarkers for Atrial Fibrillation Detection After Stroke: Systematic Review and Meta-analysis. Neurology. 2021;97(18):e1775-e89.
35. Lin HJ, Wolf PA, Kelly-Hayes M, Beiser AS, Kase CS, Benjamin EJ, et al. Stroke severity in atrial fibrillation. The Framingham Study. Stroke. 1996;27(10):1760-4.
36. Dulli DA, Stanko H, Levine RL. Atrial fibrillation is associated with severe acute ischemic stroke. Neuroepidemiology. 2003;22(2):118-23.
37. Yoo AJ, Khatri P, Mocco J, Zaidat OO, Gupta R, Frei D, et al. Impact of Thrombus Length on Outcomes After Intra-Arterial Aspiration Thrombectomy in the THERAPY Trial. Stroke. 2017;48(7):1895-900.
38. Berkhemer OA, Fransen PS, Beumer D, van den Berg LA, Lingsma HF, Yoo AJ, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. N Engl J Med. 2015;372(1):11-20.
39. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model. Stroke. 2016;47(3):895-900.
40. Sejr MH, May O, Damgaard D, Sandal BF, Nielsen JC. External continuous ECG versus loop recording for atrial fibrillation detection in patients who had a stroke. Heart. 2019;105(11):848-54.
41. Jabaudon D, Sztajzel J, Sievert K, Landis T, Sztajzel R. Usefulness of ambulatory 7-day ECG monitoring for the detection of atrial fibrillation and flutter after acute stroke and transient ischemic attack. Stroke. 2004;35(7):1647-51.
42. Higgins P, MacFarlane PW, Dawson J, McInnes GT, Langhorne P, Lees KR. Noninvasive cardiac event monitoring to detect atrial fibrillation after ischemic stroke: a randomized, controlled trial. Stroke. 2013;44(9):2525-31.
43. Schnabel RB, Haeusler KG, Healey JS, Freedman B, Boriani G, Brachmann J, et al. Searching for Atrial Fibrillation Poststroke A White Paper of the AF-SCREEN International Collaboration. Circulation. 2019;140(22):1834-50.
44. Wang Y, Qian Y, Smerin D, Zhang S, Zhao Q, Xiong X. Newly Detected Atrial Fibrillation after Acute Stroke: A Narrative Review of Causes and Implications. Cardiology. 2019;144(3-4):112-21.
45. Scheitz JF, Sposato LA, Schulz-Menger J, Nolte CH, Backs J, Endres M. Stroke-Heart Syndrome: Recent Advances and Challenges. J Am Heart Assoc. 2022;11(17):e026528.
46. Miyazaki Y, Toyoda K, Iguchi Y, Hirano T, Metoki N, Tomoda M, et al. Atrial Fibrillation After Ischemic Stroke Detected by Chest Strap-Style 7-Day Holter Monitoring and the Risk Predictors: EDUCATE-ESUS. J Atheroscler Thromb. 2021;28(5):544-54.

47. Wachter R, Weber-Kruger M, Seegers J, Edelmann F, Wohlfahrt J, Wasser K, et al. Age-dependent yield of screening for undetected atrial fibrillation in stroke patients: the Find-AF study. J Neurol. 2013;260(8):2042-5.
48. Sanak D, Hutyra M, Kral M, Bartkova A, Zapletalova J, Fedorco M, et al. Paroxysmal atrial fibrillation in young cryptogenic ischemic stroke: A 3-week ECG Holter monitoring study. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2015;159(2):283-7.
49. Wohlfahrt J, Stahrenberg R, Weber-Kruger M, Groschel S, Wasser K, Edelmann F, et al. Clinical predictors to identify paroxysmal atrial fibrillation after ischaemic stroke. Eur J Neurol. 2014;21(1):21-7.
50. 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(7):991-5.
51. Karregat EPM, Gurp NV, Bouwman AC, Uittenbogaart SB, Himmelreich JCL, Lucassen WAM, et al. Screening for paroxysmal atrial fibrillation in primary care using Holter monitoring and intermittent, ambulatory single-lead electrocardiography. Int J Cardiol. 2021;345:41-6.
52. Lumikari TJ, Putaala J, Kerola A, Sibolt G, Pirinen J, Pakarinen S, et al. Continuous 4-week ECG monitoring with adhesive electrodes reveals AF in patients with recent embolic stroke of undetermined source. Ann Noninvasive Electrocardiol. 2019;24(5):e12649.
53. Nederlandse Vereniging voor Neurologie. Richtlijn Hersenifarct en Hersenbloeding. 2017. https://richtlijnendatabase.nl/richtlijn/herseninfarct_en_hersenbloeding/startpagina_herseninfarct_-bloeding.html. p. 1-1151.
54. National Institute for Health and Care Excellence. Implantable cardiac monitors to detect atrial fibrillation after cryptogenic stroke (NICE diagnostics guidance DG41). 2020. https://www.nice.org.uk/guidance/dg41. p. 1-50.
55. Sposato LA, Chaturvedi S, Hsieh CY, Morillo CA, Kamel H. Atrial Fibrillation Detected After Stroke and Transient Ischemic Attack: A Novel Clinical Concept Challenging Current Views. Stroke. 2022;53(3):E94-E103.
56. van der Ree MH, Scholte RA, Postema PG, de Groot JR. Playing by the rules: Impact of the new General Data Protection Regulation on retrospective studies: A researcher's experience. Eur Heart J. 2019;40(24):1900-2.
57. Scheitz JF, Erdur H, Haeusler KG, Audebert HJ, Roser M, Laufs U, et al. Insular Cortex Lesions, Cardiac Troponin, and Detection of Previously Unknown Atrial Fibrillation in Acute Ischemic Stroke Insights From the Troponin Elevation in Acute Ischemic Stroke Study. Stroke. 2015;46(5):1196-201.

8

Validation of an algorithm for assessing risk of paroxysmal atrial fibrillation on continuous ECG versus 14-day Holter in primary care and post-stroke patients

Jelle C.L. Himmelreich, Lukas de Clercq, Jonathan M. Coutinho, Ralf E. Harskamp, Wim B. Busschers, Henk C.P.M. van Weert, Wim A.M. Lucassen

Submitted

ABSTRACT

Background: Early detection of paroxysmal atrial fibrillation (pAF) could contribute to preventing (recurrent) ischemic stroke through effective prophylaxis.

Objective: To validate a computer algorithm that assesses risk of pAF during non-AF rhythm on Holter electrocardiogram (ECG) recordings with 14-day Holter monitoring as reference test.

Methods: We included primary care AF screening trial participants (n=264) and consecutive patients with recent brain ischemia in The Netherlands (n=359), free of AF at baseline, who underwent 14-day Holter. We applied a computer algorithm which weighs ECG parameters to assess the risk of pAF during non-AF Holter recordings on snippets of the first 1, 2, 6, 12 and 24 hours of each Holter recording. We validated the algorithm's assessment against the outcome (AF or no AF) at the end of the study Holter.

Results: Median age in the overall cohort was 69.3 years, 45.4% were female. Holter detected 13 pAF cases during median 12 days of recording. Specificity (95% CI) of the computer algorithm for subsequent pAF rose with snippet length: 74.4% (70.5-78.1) in 1-hour snippets; 84.5% (81.3-87.3) in 24-hour snippets. Sensitivity (95% CI) was low, with estimates declining from 50.0% (15.7-84.3) in 1-hour snippets to 28.6% (3.7-71.0) in 24-hour snippets.

Conclusions: A computer algorithm showed low to moderate diagnostic accuracy for pAF in elderly primary care and post-stroke patients with low AF incidence. Due to low sensitivity for pAF, the algorithm was not effective as a stand-alone triage test for 14-day Holter monitoring in our low-risk sample. The validation was limited by the low number of positive cases in our sample, preventing definitive conclusions on the algorithm's value for pAF risk assessment.

INTRODUCTION

Atrial fibrillation (AF) is a common arrhythmia associated with increased risk of ischemic stroke for which effective anticoagulation is available.[1] Screening for AF in patients at high risk for ischemic stroke is therefore warranted.[2] Since AF detection can be costly and burdensome for patients, there is increasing interest in methods to identify patients at highest risk of AF in order for them to be monitored more extensively.[3] One patient group with established indication for AF screening are patients after ischemic stroke or transient ischemic attack (TIA) who are recommended to undergo at least 72 hours of continuous electrocardiogram (ECG) monitoring.[1] Another group are community-dwelling elderly, in whom opportunistic case finding is currently recommended to increase the likelihood of early AF detection.[1]

In recent years, researchers have looked at risk factors[4,5] as well as clinical prediction models[6-8] as tools for patient selection for AF screening. With the increased opportunities arising from automated algorithms and artificial intelligence on ECG data, new potential methods for AF risk stratification have arisen.[1,3] One such method is the Stroke Risk Analysis (SRA) algorithm that uses continuous ECG data to assess whether a person who is currently in non-AF rhythm has a high risk for paroxysmal AF (pAF) when monitored for an extended period of time.[9] The SRA algorithm has been validated in case-control settings using 1- and 24-hour ECG data as input, as well as in post-stroke patients using 1-hour ECG data for predicting 72-hour pAF, with remarkable results.[9-11] Among the questions remaining from current research are SRA's accuracy when using ECG input over one but under 24 hours, in order to assess whether an optimum duration for pAF risk prediction can be derived within this time window. Also, SRA validation for predicting pAF beyond the guideline-recommended monitoring duration of ≥72 hours for post-stroke patients is warranted.[1] Finally, the algorithm has never been tested for assessing risk of pAF in community-dwelling elderly patients.

If the SRA algorithm were found to have clinically relevant predictive abilities in community-dwelling elderly or post-stroke patients, it could serve as a triage test for prolonged monitoring. We therefore validated the SRA algorithm's high and low pAF risk assessment categories against the outcome of 14-day Holter in participants from a primary care AF screening study as well as in consecutive patients presenting for ischemic stroke or TIA.

METHODS

Primary care elderly and post-stroke datasets

For the current analysis we used data from primary care elderly as well as from post-stroke patients who underwent 14-day Holter as part of two separate studies conducted by our research group.[12,13]

Primary care elderly patients were recruited as part of the Detecting and Diagnosing AF (D_2AF) study, a cluster randomized controlled trial comparing opportunistic case finding for AF with care as usual in Dutch primary care patients 65 years or over and free of AF (Netherlands Trial Register [NTR] No NL4776 (old NTR4914)). Patients from the D_2AF intervention arm who had one or more positive tests out of three index tests and a 10% random sample of index-negative patients underwent 12-lead ECG. Those without AF on study ECG were subsequently invited to undergo 14-day Holter in search of paroxysmal AF. The D_2AF-study performed 266 Holters between September 2015 and August 2018, resulting in four new AF cases.[14]

For the post-stroke cohort we included consecutive adult patients free of AF at baseline who were treated in an academic hospital in the Netherlands (Amsterdam University Medical Centers, location AMC (AUMC-AMC), Amsterdam, The Netherlands) for TIA or ischemic stroke (NTR6489). Ischemic stroke and TIA were defined as an acute loss of focal cerebral or ocular function with symptoms lasting more than or under 24 hours, respectively, and which after adequate investigation was presumed to be due to embolic or thrombotic vascular disease.[15] All included post-stroke patients underwent 14-day Holter monitoring in search of paroxysmal AF. Inclusion ran from July 2017 to June 2020, resulting in 379 Holter recordings which detected 10 new AF cases.[13]

Patient selection

Participants were eligible if raw Holter data was available and convertible to SRA-compatible format, and if at least one hour within the first 24 hours was analysable by the SRA algorithm. For validation of the SRA algorithm in each of the first 1-, 2-, 6-, 12- and 24-hour snippets of Holter recording we subsequently excluded patients who had manifest AF as per the reference standard during the snippet of interest. We did this in accordance with clinical practice, where such patients would have been confirmed AF positive upon snippet analysis, and prediction of AF risk to triage for subsequent Holter recording would have been futile. For instance, a patient with AF first detected during the 5th hour of recording would only be eligible for the 1- and 2-hour snippet validation as that patient would still have been at risk of AF detection after 1 and 2 hours. However, that same patient would not be eligible for the 6-, 12- and 24-hour

snippet analyses as AF would have already been detected within these snippets with no subsequent need for AF risk prediction.

Data collection

The D_2AF study remotely extracted baseline data including age, sex, ethnicity, and history of heart failure, hypertension, diabetes, stroke, TIA, systemic embolism (SE), and vascular disease, from electronic health records (EHRs) from participating primary care practices. Additional data on ethnicity as well as systolic and diastolic blood pressure were taken at the index visit. We calculated the CHA_2DS_2-VASc score from its individual components recorded at baseline.[16]

We extracted baseline data of RAPID-AF participants from the hospital's routine care EHR data, with additional study data taken during the index visit. RAPID-AF baseline data included the same parameters as D_2AF, with additional data on the qualifying event, NIHSS at first presentation (National Institutes of Health Stroke Scale; ranging 0-42 where a higher score indicates clinically more severe stroke),[17] stroke/TIA location, presence of >50% ipsilateral carotid artery stenosis as per vascular imaging, treatment with intravenous thrombolysis and/or intra-arterial thrombectomy at first presentation, time from qualifying event to Holter, height, weight, smoking status, prior myocardial infarction, chronic kidney disease, use of antiplatelet drugs, angiotensin-converting enzyme inhibitors or angiotensin receptor blockers, calcium antagonists, diuretics, statins or antidiabetic drugs (oral or parenteral), and laboratory measurements including estimated glomerular filtration rate and triglycerides.

Reference standard and outcome definition

All participants in both D_2AF and RAPID-AF underwent Holter recording up to 14 days as the reference standard for presence or absence of AF. We used 2-lead Holters (Fysiologic, Amsterdam, The Netherlands) in all participants. Trained cardiologists, blinded to index test results, assessed Holters using automated signal processing followed by visual assessment of selected sections. We defined AF in accordance with recent guidelines as presence of ≥30 seconds of irregularly irregular RR intervals without discernible P waves as detected on study Holter.[1]

The index-test: SRA algorithm

The SRA algorithm (Apoplex Medical Technologies, Pirmasens, Germany) is an automated screening software program to assess presence of AF, or risk of having paroxysmal AF in those without AF, based on continuous ECG data.[18] Details of the algorithm have been published previously.[9] In short, its AI-enabled assessment is based on automated QRS recognition followed by time series analysis of linear and

non-linear parameters such as RR difference, frequency of premature contractions, and the ratio between shortest and longest interval of maximum 6 consecutive R-R intervals, within 1-hour segments of ECG data.[9,18] The software subsequently derives one of four possible outcomes to the overall amount of ECG data fed to the algorithm at each time: 1) no AF and low risk of pAF ('low risk'); 2) no AF but increased risk of pAF ('high risk'); 3) manifest AF; 4) not analysable. When manifest AF is detected, the SRA software provides the relevant ECG section to allow for visual verification.

In case of disagreement in manifest AF assessment between SRA and the reference standard, two investigators assessed the type of rhythm and presence of any artefacts that could explain the discrepancy in the relevant ECG section, with a third acting as arbiter in case of disagreement. The researchers operated independently in study conduct and reporting from, and had no financial ties with, the SRA algorithm manufacturer.

Statistical analysis

We reported the descriptives of continuous variables as medians and interquartile range (IQR), and of categorical variables as numbers and percentages. We reported the percentage of missing data for each baseline variable (Supplementary Table S1). We reported baseline characteristics for the overall cohort, as well as for the individual primary care elderly and post-stroke cohort stratified by AF diagnosis on study Holter. In order to assess generalisability of our validation results to overall primary care elderly and post-stroke populations, we compared baseline characteristics of included primary care elderly patients with patients from the D_2AF intervention arm who did not undergo Holter, as well as of included post-stroke patients with a 25% random sample of non-included stroke/TIA presentations at AUMC-AMC who would have been eligible for study Holter.

In our primary analysis we validated SRA's low risk and high risk assessment categories in snippets of participants' first 1-, 2-, 6-, 12- and 24-hours of recording against the presence or absence of AF on the remaining hours of Holter recording following each snippet. In a secondary analysis we validated the algorithm's manifest AF assessment category against presence of AF within the first 24 hours of patients' Holter recording. We validated the SRA algorithm in the overall sample (D_2AF + RAPID-AF) as the primary validation cohort and in the individual D_2AF and RAPID-AF cohorts as secondary validation cohorts.

In a sensitivity analysis we validated SRA's risk assessment categories in high-risk subgroups of each cohort in order to assess whether SRA accuracy could be increased

when combined with established clinical variables. In D_2AF participants we defined high risk as CHA_2DS_2-VASc ≥2 in men or ≥3 in women, i.e. those with an indication for oral anticoagulation upon AF detection.[1] In RAPID-AF patients we defined high risk as AS5F (Age, Stroke Severity NIHSS >5 to Find AF) risk score ≥67.5,[7] STAF (Score for the Targeting of AF) risk score ≥5[6] or presence of non-lacunar hemispheric stroke.[19] We also validated the SRA algorithm in all RAPID-AF participants with over 72 hours of Holter data and without AF detected in the first 72 hours of monitoring in order to validate SRA as a potential triage test for whom to select for monitoring beyond the current 72-hour minimum in the post-stroke setting as per recent European Society of Cardiology guidelines.[1]

We presented sensitivity, specificity, positive predicting value (PPV) and negative predicting value (NPV) as primary measures of validation. We presented the number needed to evaluate (NNE; the inverse of the PPV) and the positive and negative likelihood ratio (LR+ and LR-, respectively) as secondary measures of validation.[20] Here, a LR+ and 95% confidence interval (CI) >1, and LR- significantly <1, indicated that high and low risk as per the SRA algorithm were significantly associated with presence and absence, respectively, of AF on subsequent Holter. An LR+ >10 or LR- <0.1 indicated a strong triage test.[21]

Finally, we presented an exploratory analysis on the proportion of patients transitioning between SRA results (low risk, high risk, and manifest AF) in subsequent snippet duration (from 1 to 2 hours, from 2 to 6 hours, etc.) in the overall cohort. We did so in order to assess whether there was a snippet duration after which saturation had arguably been reached in terms of SRA prediction variability, or whether longer snippets would continuously increase variability in SRA results.

We used R version 3.6.1[22] using the dplyr, expss, haven, and table1 packages for the analyses.

Ethics and approval

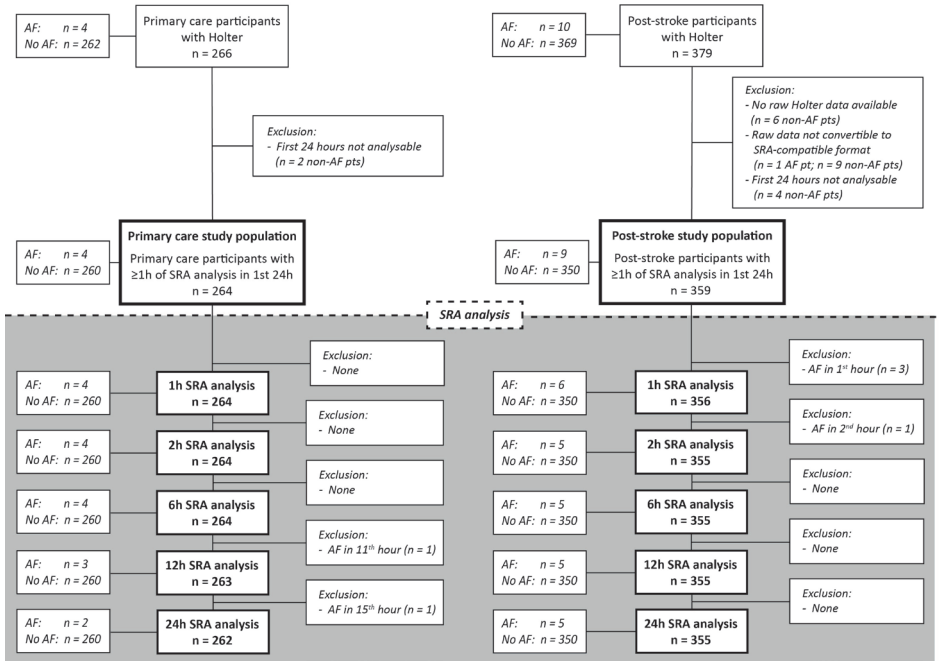
All study procedures were in accordance with the Declaration of Helsinki on medical research involving human subjects. The D_2AF study was approved by the medical research ethics committee (MREC) of the AUMC, Amsterdam (No. NL48215.018.14, 2014). The inclusion of post-stroke patients was granted a waiver for formal informed consent requirement under the Medical Research Involving Human Subjects Act (WMO) by the MREC of the AUMC, Amsterdam, as post-stroke Holter monitoring was regarded as standard of post-stroke care (No. W16_168, 2016). All post-stroke participants, however, provided written permission for use of their de-identified Holter and routine

care data as well as data acquired for study purposes (index visit questionnaire) under the General Data Protection Regulation (GDPR).

RESULTS

The primary care elderly and post-stroke cohorts included 266 and 379 participants, respectively, with available Holter data. Of these, 264 and 359 participants, respectively, had ≥1 hour out of the first 24 hours available for SRA algorithm validation. Together, these 623 participants constituted the study population (see Figure 1 for study flowchart and reasons for exclusion).

Figure 1. *Patient flowchart*

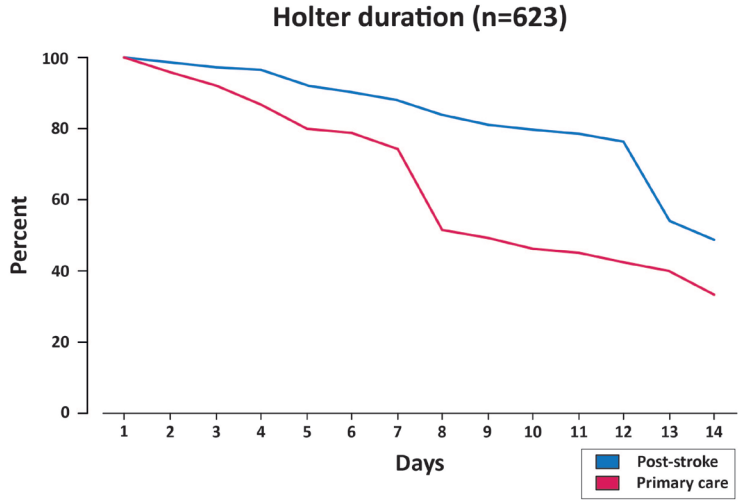


AF, atrial fibrillation; SRA, Stroke Risk Analysis algorithm.

Holter results

Median Holter recording duration was 12 (IQR 7-14) days in the overall sample, and 8 (IQR 6-14) and 13 (IQR 12-14) days in the individual primary care elderly and post-stroke cohorts, respectively (Figure 2)

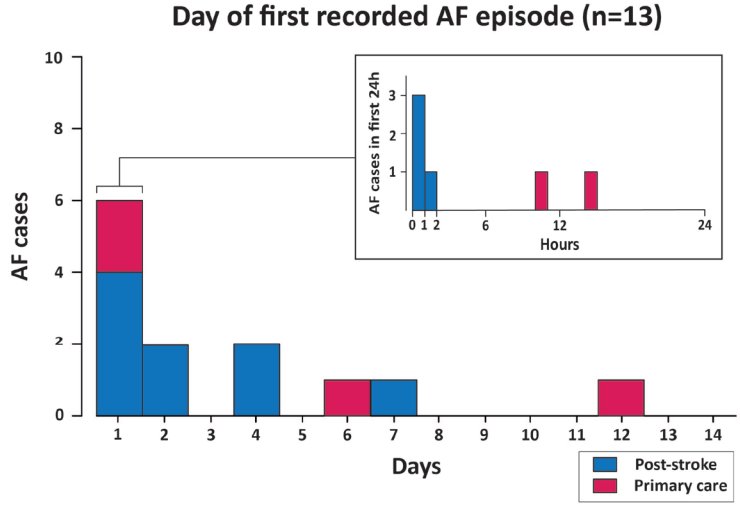
Figure 2. *Holter duration within the primary care elderly and post-stroke cohorts.*



Plot depicts the percentage of patients (y axis) with Holter duration up to n days (x axis).

The overall sample contained 13 cases of Holter-detected AF. All but one AF cases were detected within the first week of monitoring (Figure 3). Among the six patients with AF detected during the first 24 hours, three post-stroke patients had AF within the 1st hour, one AF case was detected in the 2nd hour in the post-stroke cohort, one AF case from the primary care cohort was first seen in the 11th hour, and one AF case from the primary care cohort was detected in the 15th hour (Figure 3, inset).

Figure 3. *Day of first recorded AF episode within the primary care elderly and post-stroke cohorts.*



AF, atrial fibrillation.
Inset: histogram detailing the number of AF cases detected within each of the first 24 hours of recording.

Baseline characteristics

Table 1 shows the main baseline characteristics among the overall sample. Median age was 69.3 (IQR 61.0-75.1) years with 45.4% female and a majority Caucasian/white (80.6%). Median CHA_2DS_2-VASc score was 3 (IQR 2-4), with hypertension as the most common comorbid risk factor (48.6%).

Table 1. *Baseline characteristics of the overall sample (n = 623).*

	All (n = 623)
Female sex	283 (45.4%)
Age (years)	69.3 (61.0-75.1)
Ethnicity	
Caucasian/white	502 (80.6%)
African/black	69 (11.1%)
Other	51 (8.2%)
Holter duration (days)	12 (7-14)
SBP (mm Hg)	145 (130-162)
DBP (mm Hg)	82 (74-92)
Heart failure	16 (2.6%)
Hypertension	303 (48.6%)
Diabetes	118 (18.9%)
Stroke/TIA/SE	104 (16.7%)
Vascular disease	98 (15.7%)
CHA_2DS_2-VASc	3 (2-4)
CHA_2DS_2-VASc ≥2 in men, ≥3 in women	555 (89.1%)
Baseline ECG	
QTc (ms)	423 (401-448)
PAC	52 (8.3%)
PVC	29 (4.7%)
LVH	49 (7.9%)

CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; DBP, diastolic blood pressure; ECG, electrocardiogram; LVH, left ventricular hypertrophy; ms, milliseconds; PAC, premature atrial contraction; PVC, premature ventricular contraction; QTc, corrected QT interval; SBP, systolic blood pressure; SE, systemic embolism; TIA, transient ischemic attack.
Data are number (percentage) or median (interquartile range).

Primary care elderly participants were older and more often Caucasian/white, and more often had vascular disease than post-stroke participants. Prior stroke/TIA/SE was present in under 10% among primary care elderly participants, reflected in lower median CHA_2DS_2-VASc compared to the post-stroke cohort. All AF cases in each cohort were of Caucasian/white ethnicity (Supplementary Table S2). Post-stroke participants with Holter-detected AF had higher median age and higher rates of hypertension,

higher rate of ischemic stroke vs TIA as qualifying event, higher rate of non-lacunar hemispheric stroke and higher percentage of IVT and IAT at baseline, but lower prevalence of diabetes, compared to those without AF on Holter (Supplementary Tables S2 & S3).

In assessing generalisability of patient characteristics of study participants compared to non-included patients who would have been eligible for inclusion, included post-stroke patients were younger and had lower cardiovascular risk burden compared to non-included stroke/TIA presentations. Among primary care patients, those who had undergone Holter were younger and more often male, but cardiovascular risk factors were more evenly distributed compared to eligible non-included patients (Supplementary Table S4).

Primary analysis: high/low risk category validation

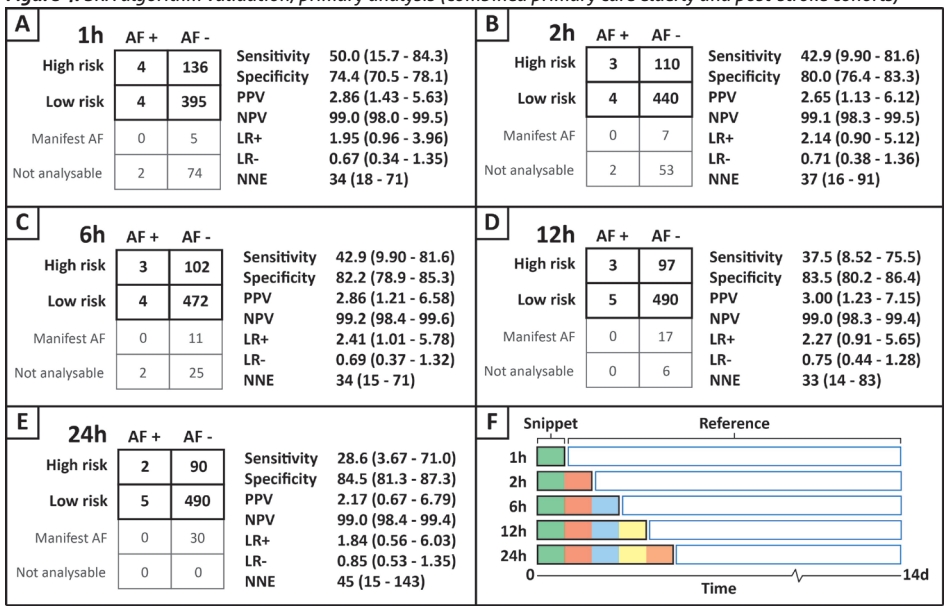
Figure 4 shows the results of the validation of the SRA algorithm for risk of pAF with Holter-detected AF beyond the snippet's duration in the overall sample as reference standard. The sensitivity decreased from of 50.0% (95% CI: 15.7-84.3) in the 1-hour snippets to 28.6% (3.7-71.0) in the 24-hour snippets while the specificity (95% CI) increased from 74.4% (70.5-78.1) in the 1h snippets to 84.5% (81.3-87.3) in the 24-hour snippets. In all snippets, the NPV was around 99.0%, while the PPV was between 2-3%. The NNE was 34 in the 1-hour snippets and 45 in the 24-hour snippets. Only the 6-hour snippets within the overall sample showed a statistically significant LR+ with 2.4 (95% CI: 1.01-5.8).

In validating the SRA algorithm within the individual D_2AF and RAPID-AF cohorts, diagnostic accuracy was generally higher in post-stroke patients (Table 2). SRA analysis within the shorter snippets in RAPID-AF participants showed the highest validity for AF on subsequent Holter, with sensitivity 66.7% (95% CI: 22.3-95.7), specificity 80.0% (95% CI: 75.1-84.3), PPV 6.2% (95% CI: 3.4-10.8; NNE 16) and NPV 99.2% (95% CI: 97.5-99.7) in post-stroke patients' 1-hour snippets. In the RAPID-AF cohort all snippets except the 24-hour snippets showed a statistically significant LR+, while none were significant in the D_2AF cohort.

Secondary analysis: manifest AF category validation

The SRA algorithm's manifest AF assessment correctly diagnosed all Holter-detected AF episodes within the first 24 hours (100% sensitivity and NPV for manifest AF detection). However, 30 patients in the overall sample's 24-hour snippets were incorrectly assessed as having manifest AF (specificity 95.1%; 95% CI: 93.1-96.7). False-positive manifest AF assessments by SRA were most often misdiagnosed ectopy

Figure 4. *SRA algorithm validation, primary analysis (combined primary care elderly and post-stroke cohorts)*



AF, atrial fibrillation; LR+, positive likelihood ratio; NNE, number needed to evaluate; NPV, negative prediction value; PPV, positive predicting value.

Panels A-E depict the 2x4 tables of SRA analysis results (rows) versus the Holter reference standard (columns) when assessing the 1-hour, 2-hour, 6-hour, 12-hour and 24-hour snippets, respectively. Diagnostic accuracy parameters (95%CI) are provided for each snippet based on the top rows ('high risk' and 'low risk' SRA results) against the reference standard. Patients with manifest AF during each snippet as per the reference standard were excluded from each snippet's 2x4 table, hence true positives for 'Manifest AF' are shown as '0'; Panel F shows a schematic diagram of the division of a patient's recording into snippets of the recording's first 1, 2, 6, 12 and 24 hours, with each consecutive snippet containing all information of the entire previous snippet plus additional hours, and with the remaining duration of that patient's Holter monitor serving as the reference standard against which the snippet's SRA algorithm results are validated.

during sinus rhythm, with a majority of false-positive recordings showing significant noise and/or baseline drift (Supplementary Table S3). Within the individual D_2AF and RAPID-AF cohorts specificity (95% CI) was 91.2% (87.0-94.3) and 98.0% (95.9-99.2), respectively.

Sensitivity analysis: high/low risk category validation in subgroups

Validation of SRA in high-risk participants saw a trend towards more favourable parameters of diagnostic accuracy than in the overall sample (Supplementary Tables S5 & S6). The analyses were however hindered by a lack of power due to low number of positive cases. A more detailed description of SRA validation in high-risk subgroups is provided in the Supplementary Results.

Table 2. Validation of SRA algorithm's high and low pAF risk categories in the individual primary care elderly and post-stroke cohorts.

Diagnostic accuracy parameter	Cohort	n (n AF)	Snippet				
			1h	2h	6h	12h	24h
Sensitivity	Primary care elderly	264 (4)	0 (0-84.2)	0 (0-84.2)	50.0 (1.3-98.7)	33.3 (0.8-90.6)	50.0 (1.3-98.7)
	Post-stroke	356 (9)	66.7 (22.3-95.7)	60.0 (14.7-94.7)	40.0 (5.3-85.3)	40.0 (5.3-85.3)	20.0 (0.5-71.6)
Specificity	Primary care elderly	264 (4)	66.8 (60.3-72.9)	73.1 (66.9-78.7)	74.7 (68.7-80.1)	76.6 (70.8-81.8)	77.6 (71.8-82.8)
	Post-stroke	356 (9)	80.0 (75.1-84.3)	85.1 (80.7-88.9)	87.7 (83.7-91.0)	88.3 (84.5-91.5)	89.2 (85.4-92.3)
PPV	Primary care elderly	264 (4)	0	0	1.6 (0.4-6.3)	1.7 (0.4-8.1)	1.9 (0.5- 7.2)
	Post-stroke	356 (9)	6.2 (3.4-10.8)	6.0 (2.9-12.0)	4.7 (1.6-12.9)	4.8 (1.6-13.2)	2.6 (0.5-13.8)
NPV	Primary care elderly	264 (4)	98.7 (98.6-99.8)	98.8 (98.8-99.9)	99.5 (97.8-99.9)	98.9 (97.7-99.5)	99.5 (97.9-99.9)
	Post-stroke	356 (9)	99.2 (97.5-99.7)	99.3 (97.9-99.8)	98.9 (97.9-99.5)	99.0 (98.0-99.5)	98.7 (98.0-99.2)
LR+	Primary care elderly	264 (4)	0	0	2.0 (0.5-8.0)	1.4 (0.3-7.2)	2.2 (0.6-9.1)
	Post-stroke	356 (9)	3.3 (1.8-6.1)	4.0 (1.9-8.7)	3.3 (1.1-9.9)	3.4 (1.1-10.4)	1.9 (0.3-11.0)
LR-	Primary care elderly	264 (4)	1.5 (1.4 - 1.6)	1.4 (1.3 - 1.5)	0.7 (0.2 - 2.7)	0.9 (0.4 - 1.9)	0.6 (0.2 - 2.6)
	Post-stroke	356 (9)	0.4 (0.1 - 1.3)	0.5 (0.2 - 1.4)	0.7 (0.3 - 1.4)	0.7 (0.3 - 1.4)	0.9 (0.6 - 1.4)
NNE	Primary care elderly	264 (4)	-	-	63 (16-250)	59 (12-250)	53 (14-200)
	Post-stroke	356 (9)	16 (9-29)	17 (8-34)	21 (8-63)	21 (8-63)	38 (7-200)

AF, atrial fibrillation; LR-, negative likelihood ratio; LR+, positive likelihood ratio; NNE, number needed to evaluate NPV, negative prediction value; PPV, positive predicting value. Values indicate percentage (95%CI) or ratio (95%CI).

SRA transitions from previous snippets

Supplementary Figure S1 shows the transitioning from SRA result 0 (low AF risk; Panel A) to 1 (risk of AF) or 2 (suspected manifest AF), and from SRA result 1 (risk of AF; Panel B) to 0 (low AF risk) or 2 (suspected manifest AF) at each subsequent snippet duration in the overall cohort. There were no transitions down from SRA result 2 (suspected manifest AF) in subsequent snippets indicating that a manifest AF assessment was consistently adjudicated by the algorithm. Most transitions in SRA results were seen within the first 6 hours of recording.

DISCUSSION

Validation of the SRA algorithm in post-stroke and elderly primary care cohorts at lower risk of AF resulted in low to moderate diagnostic accuracy for pAF during subsequent 14-day Holter recording. Our analyses were severely limited by the low AF incidence among both primary care and post-stroke participants despite good overall adherence to 14-day monitoring.

Clinical relevance

Previous studies in the field of AF risk prediction have often focused on validating tools for stratifying patients according to risk of AF detection over an extended time window, e.g. 3 months or 5 years.[6,8,23] While providing useful insights into clinical markers associated with a risk of AF, such studies are often lacking in practical guidance for which clinical actions to take in the presence of high risk. The current study was aimed at bridging the gap between prediction and clinical consequence by validating SRA as a potential triage test for immediate further monitoring. We found that in Dutch elderly primary care patients at lower risk of AF, the SRA algorithm showed no merits in informing a decision whether or not to extend monitoring up to 14 days. In Dutch post-stroke/TIA patients at lower risk of AF the SRA algorithm showed somewhat better performance than in primary care patients. However, with LR+ point estimates ranging 3.2-4.3 in those post-stroke analyses that reached statistical significance – far below the commonly accepted LR+ ≥10 required to assess a test's performance as strong[21] – SRA's predictive performance was still moderate at best.

Both the primary care elderly and post-stroke datasets were subject to sampling bias. This resulted in a study cohort with a relatively low AF yield in both samples.[13,14] Our study results may therefore not be generalizable to all elderly primary care or post-stroke patients. Given the low a priori risk in our sample, if the SRA algorithm had shown a higher sensitivity and highly significant negative likelihood

ratios in our analyses, it could still have been a useful triage tool to exclude patients of further prolonged monitoring. Since this does not seem to be the case, we conclude that the SRA algorithm is likely not the best candidate for this purpose in populations resembling our study cohorts. Validation of SRA in predicting pAF in higher-risk patients in post-stroke or community settings is, however, still warranted.

Our results of low AF screening yield in Dutch primary care and post-stroke patients leads to further questions on the extent to which researchers should aim for intensive AF screening in elderly primary care or post-stroke patients at lower risk of AF. If more intensive screening were at all investigated, more efficient efforts could likely be attempted in those with the highest estimated benefit in terms of AF detection and subsequent stroke prophylaxis. Whether ECG signal-based algorithms such as SRA could contribute to the latter aim is a question for potential future investigations.

Comparison to previous work

Early case-control studies on the SRA algorithm's high and low pAF risk categories showed around 50% sensitivity and up to 99% specificity for pAF.[9,10] A later study in post-stroke patients showed a PPV of 38.5% (95% CI: 25–52) for AF detected during or after hospitalisation for ischemic stroke or TIA.[24] While we were able to replicate the sensitivity of 50% in some of our (sensitivity) analyses, the other measures of diagnostic accuracy were consistently lower in our study. A possible explanation is that we externally validated SRA in samples with lower AF incidence than in previous work. A difference in cardiovascular risk profile of included patients could affect SRA's relative ability to discern pAF risk, given that the algorithm is based on ECG features associated with higher cardiovascular risk. The high sensitivity for manifest AF in our study concurred with previous findings.[10,24-27]

Previous work on the SRA algorithm indicated a potential increase in diagnostic accuracy with increased duration of the snippet fed to the algorithm.[10] This led us to validate SRA on snippets of increased duration in order to assess whether an optimum could be deduced at which diagnostic accuracy and burden to the patient could be balanced. Conversely, in the overall cohort we saw trends of declining sensitivity with increasing snippet duration for which we have no plausible explanation other than chance from our low incidence and subsequently broad and largely overlapping confidence intervals. Significant LR+ and overall association between SRA high risk and subsequent AF were only seen in the shorter snippets. Moreover, most transitions in SRA results occurred within the first 6 hours of recording. These data suggest that application of the SRA algorithm on 12- or 24-hour snippets provided relatively little

additional information over 1-, 2- or 6-hour snippets in terms of pAF prediction in our sample.

With its use of a multitude of ECG markers for cardiomyopathy in its risk assessment for pAF, the SRA algorithm's philosophy fits well within the current understanding of AF as one of many ECG markers for the continuum that constitutes a patient's cardiovascular disease burden – albeit one with a rich body of evidence on how to act in its presence.[28] Why, then, did the SRA algorithm not show the diagnostic accuracy that it had shown in previous work?[9,10] One possible explanation is the low incidence of Holter-detected AF in our sample; lower than would be expected in samples of elderly primary care or post-stroke patients. Previous reports have pointed to a relatively high quality of routine care in the overall Dutch medical system, as assessed by a high rate of 1-year AF incidence through routine primary care in the D_2AF control arm, as well as a high rate of prevalent or de novo AF at stroke/TIA presentation.[13,14] It is possible that those who remain at risk of AF in such a healthcare setting have a different cardiovascular make-up in terms of clinical and/or electrocardiographic profile than those patients on which the SRA algorithm has been trained – an effect that may have been amplified by the aforementioned sampling bias.

Another potential explanation for both the low AF incidence and the low validity of the SRA algorithm is that we simply have not monitored long enough. Despite good overall adherence to our 14-day protocol – especially in the post-stroke cohort – it could be that the use of loop recorders or consecutive monitoring episodes as used in other AF screening studies could have led to less uncertainty and possibly even different results.[29,30] Whether the difference in monitoring duration between AF and non-AF patients seen in the D_2AF cohort played a role in its low Holter-detected AF rate is not sure.

Strengths and limitations

This work had a number of strengths. Adherence to the 14-day protocol was high, especially in the post-stroke cohort. Few participants were excluded for non-analysable data or inability to convert data to SRA-compatible format. We were the first to validate the SRA in primary care, and the first to validate it against subsequent continuous monitoring longer than 72 hours in post-stroke patients. This enabled us to assess SRA's merits as a triage test for immediate prolonged ambulatory monitoring in primary care, and as a potential triage test for monitoring longer than the guideline-recommended minimum in post-stroke patients.[1,31] We included a comparison of baseline characteristics between study participants and patients eligible for inclusion but who were not included in each of our study cohorts. This increased transparency

into the existence of sampling bias, strengthening our understanding of whom to generalize our results to.

The principal limitations of our study were the relatively small sample size and low incidence of AF in both the D_2AF and RAPID-AF cohorts. This increased uncertainty in our overall validation analyses, and especially in our subgroup analyses aimed at identifying populations where SRA could have a higher accuracy. The limitation was further exacerbated by a considerable proportion of patients with AF detected during the first 24 hours, precluding these from further snippet validation. In all, this inhibited definitive conclusions on the value of the SRA algorithm for assessing risk of pAF on prolonged Holter monitoring based on our analysis. The use of 14-day Holter rather than e.g. implantable loop recorder devices could have underestimated AF yield in our samples.[29] Further limitations were the drop in Holter data after day 7 of Holter monitoring in the primary care sample of which the reason is improperly understood. It is not known to what extent a change and/or loss of data resolution played a role in SRA results in our conversion from the Holter provider's format to an SRA-compatible format. We mitigated this problem as much as possible by working closely together with engineers of both parties in order to achieve data compatibility, resulting in relatively few cases that were excluded for data incompatibility. Due to the low number of AF cases as well as the relatively low number of transitions between snippet lengths in our dataset we were unable to formally test the presence of an optimum snippet duration for SRA prediction variability.

Future work

The question whether our monitoring was too short for a thorough validation of the SRA algorithm due to false-negatives for pAF in our reference standard could be answered by validation of the SRA algorithm on the first hours of AF screening studies with longer follow-up employing e.g. loop recorder or repeated ambulatory monitoring.[29,30] Given the indications for higher validity in post-stroke patients at higher risk of AF, researchers could combine SRA with clinical variables to devise a more accurate triage test for prolonged monitoring, or could expand on previous work that investigates whether addition of SRA results to established clinical risk models could increase the performance of such models for AF risk prediction.[11] Finally, given SRA's reliance on ECG parameters associated with elevated stroke risk,[28] it would be interesting to test whether SRA could assist in informing treatment strategies in screening-detected pAF patients when applying SRA to non-AF snippets in these pAF patients. The latter is especially relevant in light of recent indications that not all AF may be worth screening for – and not all screening-detected AF may warrant

anticoagulation – which should urge researchers to look for alternative strategies of assessing stroke risk in screening-detected pAF patients.[29]

CONCLUSION

Validation of the SRA algorithm in elderly primary care and post-stroke patients with low AF incidence resulted in low to moderate point estimates on diagnostic accuracy for pAF during 14-day Holter recording. The SRA algorithm was not effective as a stand-alone triage test for 14-day Holter monitoring in our sample. The validation was considerably limited by the low number of positive cases, inhibiting definitive conclusions on the value of the SRA algorithm for assessing risk of pAF on prolonged Holter monitoring.

Acknowledgments: We thank all patients and staff from D_2AF and RAPID-AF sites for their participation in our research efforts. We are grateful for the support received from Apoplex Medical Technologies, in particular Albert Hirtz and Christian Beer, in conducting the SRA analyses. We thank Fysiologic ECG Services, particularly John den Engelsman, Robert den Engelsman and Dave Hopman, for their efforts in harmonizing raw Holter data with the SRA format. We thank prof. dr. Joris R. de Groot for adjudicating Holter and ECG interpretation discrepancies.

Author contributions: JCLH drafted protocol, performed data collection, performed analyses, drafted and finalised manuscript. LdC contributed to statistical analyses, provided valuable input to manuscript. WBB contributed to statistical analyses, provided valuable input to manuscript. JMC supervised inclusion, provided valuable input to methods and manuscript. REH provided valuable input to methods and manuscript. HCPMvW conceived study, provided valuable input to methods and manuscript. WAML conceived study, acquired funding, drafted protocol, provided valuable input to methods and manuscript.

Funding: This analysis was funded by grants from The Netherlands Organisation for Health Research and Development (ZonMw) (839110006 [D_2AF] & 80-83910-98-13046 [RAPID-AF]) as well as the European Research Council under the European Union's Horizon 2020 research and innovation programme (648 131).

Declarations: JCLH has received a grant from Stichting Stoffels-Hornstra and consultancy fees from Thrombosis Research Institute (outside the submitted work); JMC reported receiving grants from Dutch Heart Foundation and from Boehringer Ingelheim

(outside the submitted work); REH reported receiving grants from the Dutch Research Council (outside the submitted work); HCPMvW served on the editorial board of the European Journal of General Practice, 2012-2021; LdC, WBB and WAML report no disclosures relevant to the manuscript;

Cooperation with the software developers: The investigated software was made available free of charge by the developers. The researchers operated independently from, and had no financial ties with, the manufacturer of the investigated software.

Trial registration: Netherlands Trial Register, NTR6489.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY RESULTS

Sensitivity analysis: high/low risk category validation in subgroups

In elderly primary care participants with CHA_2DS_2-VASc ≥2 in men or ≥3 in women SRA's high and low risk categories showed poor accuracy for AF on subsequent Holter in the shorter snippets, but 100% sensitivity and significant LR+ (4.0, 95%CI: 3.1-5.1) in the 24-hour snippets (Supplementary Table S5).

High-risk subgroups in the post-stroke cohort generally saw higher diagnostic accuracy from SRA in the shorter snippets (Supplementary Table S6). Application of SRA in patients with STAF≥5 and with non-lacunar hemispheric stroke showed significant LR+ in snippets of 1- through 12-hour duration (LR+ range 2.4-4.6), with NNE being particularly low in those with non-lacunar hemispheric stroke (range 6-7 in 1- through 12-hour snippets). In patients with high AS5F risk SRA showed no significant association with subsequent AF on Holter as per the LR+.

In the subgroup of post-stroke patients with over 72 hours of monitoring data and free of AF within the first 72 hours, diagnostic accuracy parameter point estimates were similar to those of the overall RAPID-AF cohort. The LR+ remained significantly positive in the 1- and 2-hours snippets recorded in the over-72-hour post-stroke subgroup at 3.2 (95%I: 1.4-7.4) and 4.3 (95%CI: 1.9-10.1), respectively.

SUPPLEMENTARY TABLES

Supplementary Table S1. *Missing data per baseline variable in included participants*

	Primary care elderly		Post-stroke patients	
	AF (N=4)	No AF (N=260)	AF (N=9)	No AF (N=350)
Female sex	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Age (years)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Ethnicity	0 (0%)	1 (0.4%)	0 (0%)	0 (0%)
Holter duration (days)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
SBP (mm Hg)	0 (0%)	1 (0.4%)	1 (11.1%)	21 (6.0%)
DBP (mm Hg)	0 (0%)	1 (0.4%)	1 (11.1%)	21 (6.0%)
Heart failure	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Hypertension	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Diabetes	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Stroke/TIA/SE	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Vascular disease	0 (0%)	0 (0%)	0 (0%)	0 (0%)
CHA_2DS_2-VASc	0 (0%)	0 (0%)	0 (0%)	0 (0%)
QTc	0 (0%)	16 (6.2%)	0 (0%)	17 (4.9%)
PAC	0 (0%)	16 (6.2%)	0 (0%)	17 (4.9%)
PVC	0 (0%)	16 (6.2%)	0 (0%)	17 (4.9%)
LVH	0 (0%)	16 (6.2%)	0 (0%)	17 (4.9%)
Qualifying event	-	-	0 (0%)	0 (0%)
NIHSS at first presentation	-	-	0 (0%)	0 (0%)
Stroke/TIA location	-	-	0 (0%)	0 (0%)
Non-lacunar hemispheric stroke	-	-	0 (0%)	0 (0%)
Ipsilateral stenosis >50%	-	-	0 (0%)	0 (0%)
Intravenous thrombolysis	-	-	0 (0%)	0 (0%)
Intra-arterial thrombectomy	-	-	0 (0%)	0 (0%)
Time to Holter	-	-	0 (0%)	0 (0%)
Ethnicity	-	-	0 (0%)	0 (0%)
Height	-	-	0 (0%)	1 (0.3%)
Weight	-	-	0 (0%)	1 (0.3%)
Current smoker	-	-	0 (0%)	0 (0%)
Prior myocardial infarction	-	-	0 (0%)	0 (0%)
Chronic kidney disease	-	-	0 (0%)	0 (0%)
AS5F			0 (0%)	0 (0%)
STAF			0 (0%)	22 (6.3%)
Antiplatelet use	-	-	0 (0%)	0 (0%)
ACE/ARB use	-	-	0 (0%)	0 (0%)
Calcium antagonist use	-	-	0 (0%)	0 (0%)

Supplementary Table S1. *Missing data per baseline variable in included participants* (*continued*)

	Primary care elderly		Post-stroke patients	
	AF (N=4)	No AF (N=260)	AF (N=9)	No AF (N=350)
Diuretics use	-	-	0 (0%)	0 (0%)
Statin use	-	-	0 (0%)	0 (0%)
Antidiabetic drug use	-	-	0 (0%)	0 (0%)
eGFR	-	-	0 (0%)	7 (2.0%)
Triglycerides	-	-	2 (22.2%)	46 (13.1%)

ACE/ARB, angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker; AF, atrial fibrillation; AS5F, Age, Stroke Severity NIHSS >5 to Find AF; CHA_2DS_2-VA, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; COPD, chronic obstructive pulmonary disease; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; LVH, left ventricular hypertrophy; NIHSS, National Institutes of Health Stroke Scale; PAC, premature atrial contraction; PVC, premature ventricular contraction; QTc, corrected QT interval; SBP, systolic blood pressure; SE, systemic embolism; STAF, Score for the Targeting of AF; TIA, transient ischemic attack.
Data are number (percentage).

Supplementary Table S2. *Baseline characteristics of included participants by cohort and AF status*

	Primary care elderly		Post-stroke patients	
	AF (N=4)	No AF (N=260)	AF (N=9)	No AF (N=350)
Female sex	1 (25.0%)	130 (50.0%)	3 (33.3%)	149 (42.6%)
Age (years)	71.6 (70.9-72.2)	72.5 (69.2-77.1)	69.0 (63.0-71.0)	63.0 (55.0-72.8)
Ethnicity				
Caucasian/white	4 (100%)	253 (97.3%)	9 (100%)	236 (67.4%)
African/black	0 (0%)	1 (0.4%)	0 (0%)	68 (19.4%)
Other	0 (0%)	5 (1.9%)	0 (0%)	46 (13.1%)
Holter duration (days)	14 (14-14)	8 (6-14)	14 (12-14)	13 (12-14)
SBP (mm Hg)	133 (128-137)	139 (128-152)	149 (136-160)	154 (135-171)
DBP (mm Hg)	80 (71-84)	78.0 (71-84)	82 (73-94)	88 (78-99)
Heart failure	0 (0%)	7 (2.7%)	0 (0%)	9 (2.6%)
Hypertension	3 (75.0%)	137 (52.7%)	7 (77.8%)	156 (44.6%)
Diabetes	0 (0%)	52 (20.0%)	0 (0%)	66 (18.9%)
Stroke/TIA/SE	0 (0%)	25 (9.6%)	9 (100.0%)	350 (100.0%)
Vascular disease	0 (0%)	52 (20.0%)	1 (11.1%)	45 (12.9%)
CHA_2DS_2-VASc	2 (2-2)	3 (2-4)	4 (3-5)	4 (3-5)
CHA_2DS_2-VASc ≥2 in men, ≥3 in women	3 (75.0%)	193 (74.2%)	9 (100%)	350 (100%)
Baseline ECG				
QTc (ms)	392 (384-399)	403 (386-421)	449 (416-464)	437 (418-460)
PAC	1 (25.0%)	37 (14.2%)	2 (22.2%)	12 (3.4%)
PVC	0 (0%)	17 (6.5%)	0 (0%)	12 (3.4%)
LVH	0 (0%)	6 (2.3%)	0 (0%)	43 (12.3%)

AF, atrial fibrillation; CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; DBP, diastolic blood pressure; ECG, electrocardiogram; LVH, left ventricular hypertrophy; ms, milliseconds; PAC, premature atrial contraction; PVC, premature ventricular contraction; QTc, corrected QT interval; SBP, systolic blood pressure; SE, systemic embolism; TIA, transient ischemic attack.
Data are number (percentage) or median (interquartile range).

Supplementary Table S3. *Additional baseline characteristics of included post-stroke participants*

	AF (N=9)	No AF (N=350)
Qualifying event		
Ischemic stroke	8 (88.9%)	245 (70.0%)
TIA	1 (11.1%)	105 (30.0%)
NIHSS at first presentation	2 (2-13)	1 (0-3)
Stroke/TIA location		
ACA	0 (0%)	7 (2.0%)
MCA	7 (77.8%)	153 (43.7%)
PCA	1 (11.1%)	18 (5.1%)
Lacunar	0 (0%)	46 (13.1%)
Retinal	0 (0%)	12 (3.4%)
Vertebrobasilar	1 (11.1%)	114 (32.6%)
Non-lacunar hemispheric stroke	7 (77.8%)	119 (34.0%)
Ipsilateral stenosis >50%	0 (0%)	9 (2.6%)
Intravenous thrombolysis	4 (44.4%)	67 (19.1%)
Intra-arterial thrombectomy	3 (33.3%)	21 (6.0%)
Time to Holter (days)	47 (1-63)	35 (14-61)
Height (cm)	173 (168-181)	173 (165-180)
Weight (kg)	80.7 (63.0-85.0)	78.8 (70.0-89.0)
Current smoker	2 (22.2%)	70 (20.0%)
Prior myocardial infarction	1 (11.1%)	25 (7.1%)
Chronic kidney disease	1 (11.1%)	15 (4.3%)
AS5F	64.3 (61.4-71.2)	58.4 (51.6-65.2)
STAF	6 (5-6)	5 (3-5)
Antiplatelet use	3 (33.3%)	119 (34.0%)
ACE/ARB use	4 (44.4%)	96 (27.4%)
Calcium antagonist use	1 (11.1%)	62 (17.7%)
Diuretics use	4 (44.4%)	43 (12.3%)
Statin use	4 (44.4%)	127 (36.3%)
Antidiabetic drug use	0 (0%)	55 (15.7%)
eGFR (mL/min/1.73m^2)	69 (52-82)	77 (65-88)
Triglycerides (mmol/L)	0.87 (0.74-1.32)	1.38 (0.90-1.94)

ACA, anterior cerebral artery; ACE/ARB, angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker; AF, atrial fibrillation; AS5F, Age, Stroke Severity NIHSS >5 to Find AF; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate; MCA, middle cerebral artery; NIHSS, National Institutes of Health Stroke Scale; PCA, posterior cerebral artery; STAF, Score for the Targeting of AF; TIA, transient ischemic attack. Data are number (percentage) or median (interquartile range).

Supplementary Table S3. *ECG signal assessment in false-positive 24-hour snippet SRA 'manifest AF' results (n=30)*

	Rhythm				Artefacts	
ECG result	SR	Sinus arrhythmia	SR + ectopy	Not interpretable	BL drift	Noise
number	8	1	20	1	6	18

BL, baseline; ECG, electrocardiogram; SR, sinus rhythm.
Rhythm and artefact results were not mutually exclusive. Noise and/or baseline drift was present in all cases with 'sinus rhythm' or 'not interpretable' rhythm assessment.

Supplementary Table S4. Comparison of included and non-included eligible patients in the primary care elderly and post-stroke cohorts.

	Primary care elderly		Post-stroke patients	
	Included (n = 264)	D_2AF intervention arm without Holter (n = 8952)	Included (n = 359)	Eligible stroke/ TIA presentations without Holter (25% RS, n = 169)
Female sex	49.6	55.2	42.3	49.7
Age (years)	72 (69-77)	74 (70-80)	63 (55-72)	71 (60-80)
Heart failure	2.7	3.8	2.5	0
Hypertension	53.0	49.5	45.4	45.6
Diabetes	19.7	19.3	18.4	19.5
Stroke/TIA/SE	9.5	14.3	100.0	100.0
Vascular disease	19.7	20.2	12.8	21.9
CHA_2DS_2-VASc	3 (2-4)	3 (2-4)	4 (3-5)	5 (3-6)
Qualifying event				
Ischemic stroke	-	-	70.5	81.1
TIA	-	-	29.5	18.9
NIHSS at first presentation	-	-	1 (0-3)	3 (1-9)
Intravenous thrombolysis	-	-	19.8	34.3
Intra-arterial thrombectomy	-	-	6.7	17.2

CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; D_2AF, Detecting and Diagnosing Atrial Fibrillation; NIHSS, National Institutes of Health Stroke Scale; RS, random sample; SE, systemic embolism; TIA, transient ischemic attack.
Data are percentage or median (interquartile range).

Supplementary Table S5. *Sensitivity analysis validating SRA algorithm's high and low pAF risk categories in high-risk primary care elderly participants*

Diagnostic accuracy parameter	Cohort	n total (n AF)	Snippet				
			1h	2h	6h	12h	24h
Sensitivity	CHA_2DS_2-VASc ≥2 (men), ≥3 (women)	196 (3)	0 (0-84.2)	0 (0-84.2)	50.0 (1.3-98.7)	50.0 (1.3-98.7)	100.0 (2.5-100.0)
Specificity	CHA_2DS_2-VASc ≥2 (men), ≥3 (women)	196 (3)	61.5 (53.6-68.9)	68.8 (61.3-75.6)	71.4 (64.1-77.9)	74.0 (67.0-80.3)	74.7 (67.6-81.0)
PPV	CHA_2DS_2-VASc ≥2 (men), ≥3 (women)	196 (3)	0	0	1.9 (0.5-7.4)	2.1 (0.5-8.0)	2.2 (1.7-2.9)
NPV	CHA_2DS_2-VASc ≥2 (men), ≥3 (women)	196 (3)	98.1 (97.8-98.3)	98.4 (98.2-98.5)	99.2 (96.9-99.8)	99.3 (97.1-99.8)	100.0
LR+	CHA_2DS_2-VASc ≥2 (men), ≥3 (women)	196 (3)	0	0	1.8 (0.4-7.1)	1.9 (0.5-7.9)	4.1 (3.1-5.1)
LR-	CHA_2DS_2-VASc ≥2 (men), ≥3 (women)	196 (3)	1.6 (1.4-1.8)	1.5 (1.3-1.6)	0.7 (0.2-2.8)	0.7 (0.2-2.7)	0
NNE	CHA_2DS_2-VASc ≥2 (men), ≥3 (women)	196 (3)	-	-	52 (14-200)	48 (13 – 200)	45 (34-59)

CHA_2DS_2-VASc, Congestive heart failure, Hypertension, Age ≥75 years (doubled), Diabetes mellitus, prior Stroke or TIA or thromboembolism (doubled), Vascular disease, Age 65 to 74 years, Sex category; LR-, negative likelihood ratio; LR+, positive likelihood ratio; NNE, number needed to evaluate NPV, negative prediction value; PPV, positive predicting value.

Values indicate percentage (95%CI), ratio (95%CI) or number (95%CI).

Supplementary Table S6. *Sensitivity analysis validating SRA algorithm's high and low pAF risk categories in high-risk post-stroke participants*

Diagnostic accuracy parameter	Cohort	n total (n AF)	Snippet				
			1h	2h	6h	12h	24h
Sensitivity	AS5F ≥67.5	63 (4)	50.0 (1.3-98.7)	50.0 (1.3-98.7)	50.0 (1.3-98.7)	50.0 (1.3-98.7)	0.0 (0.0-84.2)
	STAF ≥5	187 (7)	60.0 (14.7-94.7)	75.0 (19.4-99.4)	50.0 (6.8-93.2)	50.0 (6.8-93.2)	25.0 (0.6-80.6)
	NLH stroke	126 (7)	66.7 (22.3- 95.7)	60.0 (14.7-94.7)	40.0 (5.3-85.3)	40.0 (5.3-85.3)	20.0 (0.5-71.6)
	>72h Holter, no AF <72h	341 (3)	66.7 (9.4-99.2)	66.7 (9.4-99.2)	33.3 (0.8-90.6)	33.3 (0.8-90.6)	0.0 (0.0-70.8)
Specificity	AS5F ≥67.5	63 (4)	69.2 (54.9-81.3)	76.5 (62.5-87.2)	77.8 (64.4-88.0)	80.0 (67.0-89.6)	80.4 (67.6-89.8)
	STAF ≥5	187 (7)	75.2 (67.5-81.8)	82.3 (75.4-87.9)	82.6 (76.0-88.1)	83.9 (77.6-89.0)	83.8 (77.5-89.0)
	NLH stroke	126 (7)	80.2 (71.1-87.5)	84.8 (76.4-91.0)	89.4 (82.2-94.4)	91.3 (84.6-95.8)	90.4 (83.5-95.1)
	>72h Holter, no AF <72h	341 (3)	79.4 (74.3-83.9)	84.6 (80.1-88.5)	85.7 (81.4-89.3)	88.0 (84.0-91.3)	88.8 (84.9-92.0)
PPV	AS5F ≥67.5	63 (4)	5.9 (1.5-21.0)	7.7 (1.9-26.6)	7.7 (1.9-26.7)	8.3 (2.0-28.6)	0
	STAF ≥5	187 (7)	7.3 (3.5-14.5)	9.7 (5.3-17.1)	6.5 (2.4-16.3)	6.7 (2.5-16.8)	3.5 (0.6-16.8)
	NLH stroke	126 (7)	16.7 (9.1-28.5)	15.8 (7.5-30.4)	14.3 (4.8-35.6)	16.7 (5.5-40.5)	8.3 (1.4-36.4)
	>72h Holter, no AF <72h	341 (3)	3.2 (1.4-7.0)	4.1 (1.8-9.0)	2.1 (0.4-9.7)	2.4 (0.5-11.3)	0
NPV	AS5F ≥67.5	63 (4)	97.3 (89.9-99.3)	97.5 (90.6-99.4)	97.7 (91.2-99.4)	97.8 (91.6-99.4)	95.7 (95.2-96.2)
	STAF ≥5	187 (7)	98.3 (95.1-99.4)	99.2 (96.0-99.8)	98.6 (96.3-99.5)	98.7 (96.5-99.5)	98.0 (96.5-98.8)
	NLH stroke	126 (7)	97.6 (92.9-99.2)	97.8 (93.8-99.2)	97.1 (94.3-99.6)	97.2 (94.5-98.6)	96.3 (94.4-97.6)
	>72h Holter, no AF <72h	341 (3)	99.6 (97.9-99.9)	99.6 (98.1-99.9)	99.3 (98.4-99.7)	99.3 (98.5-99.7)	99.0 (99.0-99.0)
LR+	AS5F ≥67.5	63 (4)	1.6 (0.4-6.9)	2.1 (0.5-9.3)	2.3 (0.5-9.8)	2.5 (0.6-11.0)	0
	STAF ≥5	187 (7)	2.4 (1.1-5.2)	4.2 (2.2-8.2)	2.9 (1.0-8.1)	3.1 (1.1-8.8)	1.5 (0.3-8.7)
	NLH stroke	126 (7)	3.4 (1.7-6.7)	3.9 (1.7-9.2)	3.8 (1.1-12.5)	4.6 (1.4-15.7)	2.1 (0.3-13.2)
	>72h Holter, no AF <72h	341 (3)	3.2 (1.4-7.4)	4.3 (1.9-10.1)	2.3 (0.5-11.8)	2.8 (0.5-14.1)	0

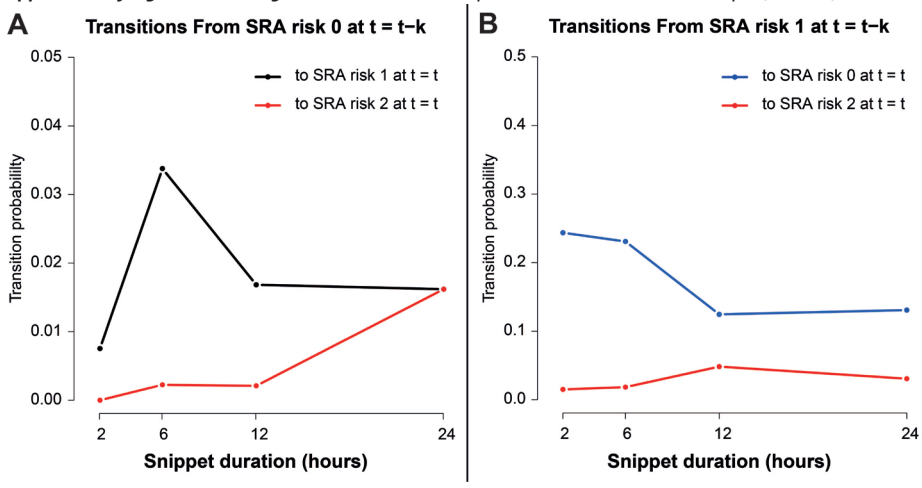
Supplementary Table S6. *Sensitivity analysis validating SRA algorithm's high and low pAF risk categories in high-risk post-stroke participants (continued)*

Diagnostic accuracy parameter	Cohort	n total (n AF)	Snippet				
			1h	2h	6h	12h	24h
LR-	AS5F ≥67.5	63 (4)	0.7 (0.2-2.9)	0.7 (0.2-2.6)	0.6 (0.2-2.6)	0.6 (0.2-2.5)	1.2 (1.1-1.4)
	STAF ≥5	187 (7)	0.5 (0.2-1.6)	0.3 (0.1-1.7)	0.6 (0.2-1.6)	0.6 (0.2-1.6)	0.9 (0.5-1.6)
	NLH stroke	126 (7)	0.4 (0.1-1.3)	0.5 (0.2-1.4)	0.7 (0.3-1.4)	0.7 (0.3-1.4)	0.9 (0.6-1.4)
	>72h Holter, no AF <72h	341 (3)	0.4 (0.1 - 2.1)	0.4 (0.1 - 2.0)	0.8 (0.4 - 1.7)	0.8 (0.3 - 2.6)	1.1 (1.1 - 1.2)
NNE	AS5F ≥67.5	63 (4)	17 (5-67)	13 (4-53)	13 (4-53)	12 (3-50)	-
	STAF ≥5	187 (7)	14 (7-29)	10 (6-19)	15 (6-42)	15 (6-40)	29 (6-167)
	NLH stroke	126 (7)	6 (4-11)	6 (3-13)	7 (3-21)	6 (2-18)	12 (3-71)
	>72h Holter, no AF <72h	341 (3)	31 (14-71)	24 (11-56)	48 (10-250)	42 (9-200)	-

AF, atrial fibrillation; AS5F, Age, Stroke Severity NIHSS >5 to Find AF; LR-, negative likelihood ratio; LR+, positive likelihood ratio; NLH, non-lacunar hemispheric; NNE, number needed to evaluate NPV, negative prediction value; PPV, positive predicting value; STAF, Score for the Targeting of AF.
Values indicate percentage (95%CI), ratio (95%CI) or number (95%CI).

SUPPLEMENTARY FIGURES

Supplementary Figure S1. *SRA algorithm result transition probabilities in the overall sample (n = 623).*



Data points and trend lines depict the probability of transitioning from SRA result 0 (low AF risk, Panel A) or 1 (risk of AF, Panel B) to SRA result 0 (low AF risk, blue, Panel B), 1 (risk of AF, black, Panel A) or 2 (suspected manifest AF, red, Panels A & B) at each snippet duration (t) compared to the SRA result of the previous snippet (k). There were no transitions from SRA result 2 to SRA result 0 or SRA result 1 in our data.

REFERENCES

1. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). Eur Heart J. 2021;42(5):373-498.
2. Freedman B, Camm J, Calkins H, Healey JS, Rosenqvist M, Wang JG, et al. Screening for Atrial Fibrillation A Report of the AF-SCREEN International Collaboration. Circulation. 2017;135(19):1851-+.
3. Schnabel RB, Haeusler KG, Healey JS, Freedman B, Boriani G, Brachmann J, et al. Searching for Atrial Fibrillation Poststroke A White Paper of the AF-SCREEN International Collaboration. Circulation. 2019;140(22):1834-50.
4. Himmelreich JCL, Lucassen WAM, Heugen M, Bossuyt PMM, Tan HL, Harskamp RE, et al. Frequent premature atrial contractions are associated with atrial fibrillation, brain ischaemia, and mortality: a systematic review and meta-analysis. Europace. 2019;21(5):698-707.
5. Cameron A, Cheng HK, Lee RP, Doherty D, Hall M, Khashayar P, et al. Biomarkers for Atrial Fibrillation Detection After Stroke: Systematic Review and Meta-analysis. Neurology. 2021;97(18):e1775-e89.
6. Suissa L, Bertora D, Lachaud S, Mahagne MH. Score for the targeting of atrial fibrillation (STAF): a new approach to the detection of atrial fibrillation in the secondary prevention of ischemic stroke. Stroke. 2009;40(8):2866-8.
7. Uphaus T, Weber-Kruger M, Grond M, Toenges G, Jahn-Eimermacher A, Jauss M, et al. Development and validation of a score to detect paroxysmal atrial fibrillation after stroke. Neurology. 2019;92(2):e115-e24.
8. Himmelreich JCL, Veelers L, Lucassen WAM, Schnabel RB, Rienstra M, van Weert H, et al. Prediction models for atrial fibrillation applicable in the community: a systematic review and meta-analysis. Europace. 2020;22(5):684–94.
9. Duning T, Kirchhof P, Wersching H, Hepp T, Reinhardt R. Extended Electrocardiographic Poincare Analysis (EPA) for Better Identification of Patients with Paroxysmal Atrial Fibrillation. J Clin Exp Cardiolog. 2011;02(02).
10. Schaefer JR, Leussler D, Rosin L, Pittrow D, Hepp T. Improved detection of paroxysmal atrial fibrillation utilizing a software-assisted electrocardiogram approach. PLoS One. 2014;9(2):e89328.
11. Groschel S, Lange B, Wasser K, Hahn M, Wachter R, Groschel K, et al. Software-based analysis of 1-hour Holter ECG to select for prolonged ECG monitoring after stroke. Ann Clin Transl Neurol. 2020;7(10):1779-87.
12. Uittenbogaart SB, Verbiest-van Gurp N, Erkens PM, Lucassen WA, Knottnerus JA, Winkens B, et al. Detecting and Diagnosing Atrial Fibrillation (D_2AF): study protocol for a cluster randomised controlled trial. Trials. 2015;16:478.
13. Himmelreich JCL, Lucassen WAM, Coutinho JM, Harskamp RE, de Groot JR, van Weert HCPM. 14-day Holter monitoring for atrial fibrillation after ischemic stroke: The yield of guideline-recommended monitoring duration. European Stroke Journal. 2022;00(0):1-11.
14. Uittenbogaart SB, Verbiest-van Gurp N, Lucassen WAM, Winkens B, Nielen M, Erkens PMG, et al. Opportunistic screening versus usual care for detection of atrial fibrillation in primary care: cluster randomised controlled trial. BMJ. 2020;370:m3208.

15. Fonseca AC, Merwick A, Dennis M, Ferrari J, Ferro JM, Kelly P, et al. European Stroke Organisation (ESO) guidelines on management of transient ischaemic attack. Eur Stroke J. 2021;6(2):CLXIII-CLXXXVI.
16. 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(2):263-72.
17. Brott T, Adams HP, Jr., Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke. 1989;20(7):864-70.
18. 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(4):410-7.
19. Yaghi S, Raz E, Yang D, Cutting S, Mac Grory B, Elkind MS, et al. Lacunar stroke: mechanisms and therapeutic implications. J Neurol Neurosurg Psychiatry. 2021;92:823–30.
20. Romero-Brufau S, Huddleston JM, Escobar GJ, Liebow M. Why the C-statistic is not informative to evaluate early warning scores and what metrics to use. Critical Care. 2015;19.
21. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ. 2004;329(7458):168-9.
22. R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
23. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB, et al. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc. 2013;2(2):e000102.
24. Adami A, Gentile C, Hepp T, Molon G, Gigli GL, Valente M, et al. Electrocardiographic RR Interval Dynamic Analysis to Identify Acute Stroke Patients at High Risk for Atrial Fibrillation Episodes During Stroke Unit Admission. Transl Stroke Res. 2019;10(3):273-8.
25. Rizos T, Guntner 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(10):2689-94.
26. Groschel S, Lange B, Grond M, Jauss M, Kirchhof P, Rostock T, et al. Automatic Holter electrocardiogram analysis in ischaemic stroke patients to detect paroxysmal atrial fibrillation: ready to replace physicians? Eur J Neurol. 2020;27(7):1272-8.
27. Ross LS, Bettin M, Kochhauser S, Ritter M, Minnerup J, Eckardt L, et al. Sensitive Detection of Atrial Fibrillation in Acute Stroke Patients by Short-Term Bedside Electrocardiography Monitoring Software Analysis. Cerebrovasc Dis. 2018;45(1-2):54-60.
28. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model. Stroke. 2016;47(3):895-900.
29. Svendsen JH, Diederichsen SZ, Hojberg S, Krieger DW, Graff C, Kronborg C, et al. Implantable loop recorder detection of atrial fibrillation to prevent stroke (The LOOP Study): a randomised controlled trial. Lancet. 2021;398(10310):1507-16.
30. Brik T, Lucassen WAM, Harskamp RE, Karregat EPM, Himmelreich JCL, Busschers WB, et al. Personalized approach using wearable technology for early detection of atrial fibrillation in high-risk primary care patients (PATCH-AF): Study protocol for a cluster randomized controlled trial. Am Heart J. 2022;254:172-82.
31. Rubiera M, Aires A, Antonenko K, Lemeret S, Nolte CH, Putaala J, et al. European Stroke Organisation (ESO) guideline on screening for subclinical atrial fibrillation after stroke or transient ischaemic attack of undetermined origin. European Stroke Journal. 2022;7(3):CVII–CXXXIX.

9

General discussion and perspectives for future research

GENERAL DISCUSSION

Main findings

This thesis was initiated to investigate methods to improve patient selection for atrial fibrillation (AF) screening and to facilitate early identification of AF, after our group's experience from the neutral Detecting and Diagnosing AF (D_2AF) primary AF screening trial which selected patients for high age alone (see also **Chapter 1**).[1] It was theorised that a triage test was necessary to increase efficiency of future interventions by restricting the screening effort only to those at highest risk of AF detection. We therefore set out to investigate and validate multiple methods with potential use for stratifying patients into higher and lower risk of AF. These could then be used for patient selection in future targeted primary AF screening efforts.

First, we systematically reviewed current literature and found that CHARGE-AF (Cohorts for Heart and Aging Research in Genomic Epidemiology model for Atrial Fibrillation)[2] seemed best equipped for use in primary AF screening (**Chapter 2**). In a subsequent validation in a large database of Dutch routine primary care data, however, we found that only a minority of older patients had complete data for all CHARGE-AF variables (**Chapter 3**). This could limit the use of CHARGE-AF for remote risk stratification based on electronic health record (EHR) data without requiring a baseline visit. The CHA_2DS_2-VASc risk score, which had slightly lower predictive performance but was universally applicable to routine primary care data, was therefore validated as a potential alternative.

Subsequently, we looked at premature atrial contractions (PACs) as a potential marker for AF, and found that frequent PACs on continuous electrocardiogram (ECG) recordings were indeed associated with AF, as well as with ischemic stroke and all-cause mortality (**Chapter 4**). There was a trend towards, but not yet a definitive association between finding 1 or more PACs on standard ECG and AF. A validation of our findings in a cohort of primary care patients with type 2 diabetes confirmed that detecting a PAC on 12-lead ECG was associated with later AF detection, although the association with brain ischaemia and/or TIA could not be replicated in these patients (**Chapter 5**). Our findings fitted well within the framework laid out by Kamel and colleagues who proposed the concept of an abnormal atrial substrate – of which frequent PACs as well as AF are electrocardiographic 'fingerprints' – as the mechanistic link between cardiovascular risk profile and subsequent risk of stroke and mortality.[3]

We then turned to artificial intelligence (AI) enabled methods for identifying (risk of) AF based on ECG signal. We validated a single-lead ECG device for AF detection

(**Chapter 6**), and found good diagnostic accuracy from its in-built AF diagnosing algorithm and perfect properties when the visual ECG signal was assessed by expert ECG readers. This indicated that the single-lead ECG device could be a valuable addition for unobtrusive rhythm diagnosis in patients suspected of AF, whether in the primary care clinic or during house visits.

Finally, we validated an AI-enabled algorithm for the assessment of risk of underlying paroxysmal AF (pAF) during 14-days Holter (continuous ECG) recordings. In order to supplement our existing database of Holter participants from the D_2AF trial, we collected an additional sample of patients who had presented at the department of Neurology of Amsterdam UMC, location AMC, for transient ischemic attack (TIA) or ischemic stroke. In data collection for this 'post-stroke' cohort, we found a surprisingly low yield of newly detected AF compared to international literature (**Chapter 7**). Although in a post-hoc analysis we found that sampling bias had taken place which caused us to have included post-stroke patients generally at lower risk of AF, we saw that our patients' baseline characteristics were still similar to those reported in prior international post-stroke AF screening studies with higher AF yield from screening. Combined with a relatively high prevalence of known AF at TIA/stroke presentation, these findings were in line with prior conclusions from D_2AF and other work on AF screening in The Netherlands[4] that the Dutch routine care system already seems comparatively conducive to AF detection.

In the subsequent validation (**Chapter 8**), the AI algorithm showed insufficient sensitivity to serve as a triage test for 14-day Holter monitoring in the combined cohort of elderly primary care and post-stroke at lower risk of AF. The analyses were, however, severely limited by the low AF yield in both cohorts, leading to a high degree of uncertainty of the validation results.

BENEFITS OF RISK-STRATIFIED AF SCREENING

The yield of risk-stratified screening for AF in the community

The past years have seen an increase in work that relates to the overarching aim of this thesis: community AF screening with patient selection through risk stratification beyond age alone. Further validation studies of existing risk models for routine care AF risk prediction have underwritten the work shown in **Chapter 2**.[5] Moreover, new risk models have recently been developed with the aim to optimise the use of data contained in primary care EHRs for predicting AF risk in the community.[6-8] These attempted to advance the field by going beyond classically used variables such as age,

sex, or medical history that are stored as pre-set variables in each EHR ('coded data'), and involving free text and other non-coded data in their models, as well as by using more advanced modelling techniques. In doing so, these models have shown promising results compared to previously used risk models in the field.[6,7]

Recent work in primary AF screening has also increasingly employed risk factors or clinical risk prediction models to select patients for their screening intervention, over selecting for age alone.[9] A selection of published community AF screening studies that used risk factors beyond high age alone in selecting for screening to achieve higher AF yield is provided in Table 1.[10-15] In all, these results indicate that screening in patients selected for age and additional risk factors can be highly effective in achieving high AF yield in community settings. In most of these trials, risk stratification was based on one single additional risk factor beyond age, and two others on the CHA_2DS_2-VASc risk model. This thesis, which focused on identifying markers beyond age alone with potential use in patient selection for primary AF screening (**Chapters 2-5 & 8**), therefore fits well within this field of research.

Ongoing trials on risk-stratified AF screening in primary care

Several other AF screening trials with patient selection based on risk stratification beyond age alone in primary care or community settings are still ongoing.[16,17] Of particular interest in the context of this thesis is the PATCH-AF (Personalized approach using wearable technology for early detection of atrial fibrillation in high-risk primary care patients) trial, conducted by our AmstelHeart research unit in collaboration with the Amsterdam UMC Primary Care Network (ANHA).[16] PATCH-AF is a cluster-randomized, controlled trial that includes high-risk primary care patients, defined as 65 years or older and with a CHA_2DS_2-VASc score ≥3 for men or ≥4 for women, free of AF at baseline, from 20 ANHA-affiliated primary care practices in the Amsterdam region in The Netherlands randomised 1:1 to intervention or control practice. Risk stratification occurs remotely (prior to baseline visit) based on the GP's EHR as extracted by ANHA. Participants are visited at home by our research team and will undergo annual 7-day Holter monitoring for three years in a row. The three-year AF yield for the total intervention cohort (AF detected through screening or usual care) will be compared with that of an age and comorbidity matched cohort from control ANHA practices.[16] Results of this trial are expected by 2025.

An interesting trial to compare PATCH-AF's results to, will be the ongoing AMALFI study conducted in the United Kingdom (UK). This trial has similar inclusion criteria as PATCH-AF (age ≥65 and with CHA_2DS_2-VASc ≥3 in men or ≥4 in women), but randomises patients to receive one 14-day Holter within a 5-year follow-up period, after which AF

Table 1. *Atrial fibrillation yield of recent community screening studies with risk stratification beyond age alone*

Trial (years of recruitment)	n	Country	Inclusion criteria	Intervention	Follow-up duration	AF yield
LOOP (2014-2016)[10]	6004	DK	Age ≥70 years with at least one additional risk factor (hypertension, diabetes, heart failure or prior stroke)	Randomised 1:3 to ILR, or usual care	Median 64.5 months	31.8% in the intervention arm vs 12.2% in the usual care arm (HR 3.17; 95% CI: 2.81–3.59)
mSToPS (2015-2016)[11]	1738	USA	Age ≥75 years, or men >55 or women >65 with one or more cardiovascular risk factors	Randomised 1:1 to 14-day Holter at baseline and at 3 months (total 4-month intervention), or the same intervention with a 4-month delay; matched observational cohort received usual care	4 months & 12 months	- 4 months: 3.9% in the immediate group vs 0.9% in the delayed group (absolute difference 3.0%; 95% CI: 1.8%-4.1%); - 12 months: 6.7 per 100 PY in monitored vs 2.6 per 100 PY in unmonitored (difference: 4.1; 95% CI: 3.9-4.2)
REHEARSE-AF (2015)[12]	1001	UK	Age ≥65 years and CHA_2DS_2-VASc score ≥2	Randomised 1:1 to an 1L-ECG to be used twice per week for 12 months and when symptomatic, or usual care	12 months	19/500 in the intervention arm vs 5/501 in the control arm (HR: 3.9; 95% CI: 1.4-10.4)
STROKESTOP II (2016)[13]	3766	SE	Respondents with NT-ProBNP ≥125 ng/L among a random sample of all 75- and 76-year-olds in the Stockholm region	Invited to record multiple 1L-ECGs per day for 14 days	14 days	164/3766 (4.4%; 95% CI: 3.7–5.10) new AF during 14-day intervention (no information on 14-day AF yield of usual care arm)
SCREEN-AF (2015-2019)[14]	856	CA	Age 75 or older with hypertension	Randomised 1:1 to 14-day Holter patch and simultaneous twice-daily use of eBPM with AF detection functionality at baseline & at 3 months, or usual care	6 months	5.3% in intervention arm vs 0.5 in control arm (relative risk: 11.2; 95% CI: 2.7-47.1; absolute difference: 4.8%; 95% CI, 2.6%-7.0%)
eBRAVE-AF (2020)[15]	5551	DE	Age ≥50 and CHA_2DS_2-VASc ≥1 in men or ≥2 in women	Randomised 1:1 to using a smartphone app with PPG pulse monitoring followed by ELR in case of irregular pulse, or usual care	6 months	1.33% with new AF leading to OAC treatment in intervention arm vs 0.63% in usual care arm (odds ratio: 2.12; 95% CI: 1.19-3.76)

AF, atrial fibrillation; CA, Canada; CHA_2DS_2-VASc, Cardiac failure or dysfunction, Hypertension, Age >=75 [Doubled], Diabetes, Stroke [Doubled]-Vascular disease, Age 65-74, and Sex category [Female]; CI, confidence interval; DE, Germany; DK, Denmark; eBPM, electronic blood pressure monitor; eBRAVE-AF, eHealth-based Bavarian Alternative Detection of Atrial Fibrillation; ELR, external loop recorder; HR, hazard ration; ILR, implantable loop recorder; LOOP, Atrial Fibrillation Detected by Continuous ECG Monitoring Using Implantable Loop Recorder to Prevent Stroke in High-risk Individuals; NT-ProBNP, N-terminal B-type natriuretic peptide; mSTOPS, mHealth Screening to Prevent Strokes; PPG, photoplethysmogram; PY, person-years; REHEARSE-AF, Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation; SCREEN-AF, Screening for Atrial Fibrillation in the Older Population; SE, Sweden; STROKESTOP, Systematic ECG Screening for Atrial Fibrillation; UK, United Kingdom; USA, United States of America; 1L-ECG, single-lead- electrocardiogram.

incidence will be compared to that of the control group.[17] By comparing AF yield in each trial to their respective predecessor which selected patients for age ≥65 years alone (D_2AF in The Netherlands, and the Screening for AF in the Elderly [SAFE] trial in the UK[18]) the benefit of selecting higher-risk patients for screening could be further estimated. And in comparing AF incidence of each trial's control arm, further light will be shed on the question whether the current standard of care in The Netherlands is already relatively conducive to routine care AF detection compared to international settings (see discussed above). This, in turn, could have ramifications for the generalisability of international AF screening trials to the Dutch setting and vice versa.

Recent developments in post-stroke AF screening

Where in primary care the merits of routine AF screening are still debated, the discussion whether to screen has already been settled in the post-stroke setting.[19-21] The occurrence of recent ischemic stroke or TIA is deemed such an important risk factor for presence of AF that prolonged cardiac monitoring for AF is immediately warranted.[21] Moreover, treatment with OAC is known to be beneficial in preventing stroke recurrence and mortality in patients with AF detected after stroke.[22]

In the post-stroke setting, the questions are therefore rather by what means to monitor, for how long at minimum, and whom to select for extended monitoring beyond the guideline-recommended minimum.[19-21] Randomised trials have been conducted to shed light on these questions. Table 2 lists a selection of relevant trials on post-stroke AF screening and their results in terms of AF yield. Post-stroke screening for AF beyond routine care seems to increase yield of incident AF detection, especially when selecting for further risk factors than the occurrence of recent stroke or TIA itself. Selection was based on prior diagnostic work-up (e.g. cryptogenic stroke, in which work-up with several diagnostic investigations had not yet resulted in finding a cause for stroke), time from stroke onset to inclusion, stroke location, or additional stroke risk factors. Notice, however, that in none of the currently published trials the standard of care to which the intervention was compared was 72 hours of continuous monitoring, as the European Society of Cardiology currently advises.[21] Given the increasing technological abilities for continuous cardiac monitoring, one might expect that the currently advised minimum of up to 72 hours will be progressively increased. However, we presented evidence for there being a sizable subset of patients in whom external monitoring beyond 72 hours holds little benefit (see **Chapter 7** of this thesis). It therefore remains important to weigh the increased burden to the patient from wearing prolonged external monitoring to their relative chance of finding AF during said monitoring. Further studies therefore are awaited, and it would be interesting

Table 2. *Atrial fibrillation yield of recent post-stroke atrial fibrillation screening trials*

Trial (years of recruitment)	n	Country/ region	Inclusion criteria	Intervention	Follow-up duration	AF yield
CRYSTAL-AF (2009-2012)[23]	441	CA, EU, USA	Age ≥40 years with unknown origin for their recent cryptogenic ischemic stroke or TIA after (work-up for cryptogenic stroke included 24-hour Holter, brain and neck imaging, and echocardiography)	Randomised 1:1 to ILR insertion or usual care (ECG monitoring at the discretion of the physician; no predefined monitoring minimum)	6 & 12 months	- 6 months: 8.9% in intervention vs 1.4% in control arm (HR: 6.4%; 95% CI 1.9-21.7); - 12 months: 12.4% in intervention vs 2.0% in control arm (HR: 7.3; 95% CI: 2.6-20.8)
EMBRACE (2009-2012)[24]	572	CA	Age ≥55 years with recent (<6 months) cryptogenic ischemic stroke or TIA (work-up for cryptogenic stroke included 24-hour Holter, brain imaging, and echocardiography)	Randomised 1:1 to 30-day event-triggered recorder or a conventional 24-hour monitor	90 days	16.1% in intervention arm vs 3.2% in control arm (absolute difference: 12.9%; 95% CI: 8.0%-17.6%)
FIND-$AF_{RANDOMISED}$ (2013-2014)[25]	398	DE	Age ≥60 years with recent ischemic stroke (<7 days)	Randomised 1:1 to three 10-day Holters (at baseline, 3 months, and 6 months) or usual care (≥24-hour Holter with further monitoring at the physician's discretion)	6 months & 12 months	- 6 months: 13.5% in the intervention arm vs 4.5% in control arm (absolute difference 9.0%; 95% CI: 3.5%–14.6%) - 12 months: 13.5% in intervention arm vs 6.1% in the control arm (absolute difference 7.4%; 95% CI: 1.6%-13.2%)
MonDAFIS (2014-2017)[26]	3465	DE	Stroke or TIA patients admitted <72 hours after onset of symptoms	Randomised 1:1 to 7-day in-hospital continuous monitoring or routine care (≥24-hour Holter with further monitoring at the physician's discretion)	7 days	5.8% in intervention arm vs 4.0% in control arm (HR: 1.4; 95% CI: 1.0–2.0)
PER DIEM (2015-2017)[27]	300	CA	Patients with arterial ischemic stroke or clinical TIA but with evidence of infarction on brain imaging op to 6 months before inclusion	Randomised 1:1 to ILR insertion for 12 months, or a 30-day external loop recorder at baseline	12 months	15.3% in the intervention arm vs 4.7% in the control arm (absolute difference: 10.7%; 95% CI: 4.0%-17.3%).

Table 2. *Atrial fibrillation yield of recent post-stroke atrial fibrillation screening trials* (*continued*)

Trial (years of recruitment)	n	Country/ region	Inclusion criteria	Intervention	Follow-up duration	AF yield
STROKE-AF (2016-2019)[28]	496	USA	Age ≥60 years with stroke attributed to large- or small-vessel disease, or age 50-59 years with at least 1 additional stroke risk factor	Randomised 1:1 to either ILR insertion or usual care (ECG monitoring according to site usual care; no predefined monitoring minimum)	12 months	12.1% in intervention arm vs 1.8% in control arm (HR: 7.4; 95% CI: 2.6-21.3)

AF, atrial fibrillation; CA, Canada; CRYSTAL-AF, Cryptogenic Stroke and Underlying AF; CI, confidence interval; DE, Germany; EMBRACE, Cardiac Event Monitor Belt for Recording Atrial Fibrillation after a Cerebral Ischemic Event; EU, Europe; FIND-$AF_{RANDOMISED}$, Holter-electrocardiogram-monitoring in patients with acute ischaemic stroke; HR, hazard ration; ILR, implantable loop recorder; MonDAFIS, MONitoring for Detection of Atrial Fibrillation in Ischemic Stroke; PER DIEM, Effect of Implantable vs Prolonged External Electrocardiographic Monitoring on Atrial Fibrillation Detection in Patients With Ischemic Stroke; TIA, transient ischemic attack; USA, United States of America;

to see whether the use of clinical prediction models for selecting – or ruling out – patients for extended monitoring will increase in this field.

Does screening for AF reduce adverse outcomes?

Arguably one of the most important recent developments is the increased emphasis on the question whether screening for AF is ultimately also associated with fewer strokes or other adverse outcomes among those randomised to the screening intervention.[9,29,30] Since the start of this PhD project, several important studies have been conducted that included outcomes such as stroke and mortality during follow-up in their community and post-stroke AF screening trials. The results so far have been surprising, and often counterintuitive, as summarised in Table 3.

In the community screening setting, the mSToPS trial (mHealth Screening to Prevent Strokes) showed a significant benefit from screening in terms of long-term (3-year) ischemic stroke or the composite of death, stroke, systemic embolism, or myocardial infarction,[31] and STROKESTOP (Systematic ECG Screening for Atrial Fibrillation) saw a slightly significant decrease in the composite of ischaemic or haemorrhagic stroke, systemic embolism, bleeding leading to hospitalisation, or all-cause death after 6.9 years of follow-up.[32] However, for ischemic stroke, STROKESTOP was underpowered, with HR: 0.92 (95% CI: 0.83–1.01).[32] In the LOOP study (Atrial Fibrillation Detected by Continuous ECG Monitoring Using Implantable Loop Recorder to Prevent Stroke in High-risk Individuals), no difference in stroke or systemic embolism incidence was seen after over 5 years of follow-up, despite highly increased OAC initiation in the intervention arm.[10] The authors therefore wondered whether all AF is worth screening for, and whether all AF detected through screening warrants anticoagulation treatment.[10] Other community AF screening trials are still in the midst of follow-up for clinical outcomes after AF screening intervention.[13,33-35]

As in primary care, there is the question whether post-stroke screening for AF leads to fewer recurrent stroke or deaths after extended follow-up. Several trials have published results, however to date none have been sufficiently powered to conclusively show the benefit of post-stroke AF screening on their primary endpoint (Table 3). However, the MonDAFIS trial (MONitoring for Detection of Atrial Fibrillation in Ischemic Stroke) showed a decrease in the secondary endpoint all-cause mortality (odds ratio: 0.7; 95% CI: 0.5-0.9).[26] Researchers in the field therefore await with interest the ongoing FIND-AF2 trial (Intensive Rhythm Monitoring to Decrease Ischemic Stroke and Systemic Embolism).[36] This will include 5200 patients with recent ischemic stroke (hospitalisation within last 30 days), who will receive a risk-stratified intervention: 1040 with high risk will receive an ILR, while those with low risk will receive

sequential 7-day Holter monitoring (at baseline, after 3 and 12 months, and annually thereafter). Here, high risk will be defined by increased atrial ectopic activity as per 24-hour Holter prior to randomisation, recognising the importance of this risk factor for probability of AF detection as also featured in **Chapters 4 and 5** of this thesis. Patients will be followed for a minimum of 2 and up to 5 years for the primary efficacy outcome recurrent ischemic stroke or systemic embolism, and with first haemorrhagic stroke as primary safety outcome.[36]

Primary screening for AF: game over?

The mixed, and sometimes disappointing results of currently published trials on the effect of AF screening on long-term adverse outcomes gives rise to the intuitive question whether AF screening should keep being pursued, or whether an AF diagnosis is as relevant in each patient.[10] Indeed, there is increasing evidence that not all AF diagnoses should be regarded as similar, from an electrophysiological nor from a clinical vantage point.[37-41] New approaches have therefore been outlined to better classify subgroups of AF diagnoses in order to optimize treatment decisions.[42] Whether a subclassification in types of AF to screen for will at some point arise, and how this selection should be achieved, remains unknown. What is necessary for this aim, however, is continued research in optimal patient selection strategies, for which this thesis has aimed to provide some additional work.

It should be concluded that the current evidence is insufficient to draw definitive conclusions on the merits for AF screening on long-term outcomes. In order to combine forces in answering this question, a consortium of international researchers has launched an initiative to collect individual patient data from published as well as ongoing AF screening trials around the globe.[9] The aim of this combined effort is to resolve the question whether AF screening is not only associated with increased AF detection, but ultimately whether it also able to do what it was designed to do: prevent stroke and reduce early mortality.[9]

OPPORTUNITIES FOR IMPROVED OUTCOMES AFTER SCREENING-DETECTED AF

Screening-detected AF and the cardiovascular continuum

A possible explanation for the neutral results of past AF screening interventions in terms of clinical outcomes, is that current standard of care after an AF diagnosis still mainly focused on stroke prevention through OAC initiation, while other aspects of AF-associated complications are often given less attention in the provided care.

Table 3. *Results of previous AF screening trials that reported long-term outcomes*

Trial (years of recruitment)	n	Country	Inclusion criteria	Intervention	Follow-up duration	Difference in reported outcomes
Community setting						
mSToPS (2015-2016)[31]	1718	USA	Age ≥75 years, or men >55 or women >65 with one or more cardiovascular risk factors	Randomised 1:1 to 14-day Holter at baseline and at 3 months (total 4-month intervention), or the same intervention with a 4-month delay; matched observational cohort received usual care	3 years	- Composite of death, IS, systemic embolism, or MI*: 3.6 per 100 PY in intervention arm vs 4.5 per 100 PY in control arm (adjusted HR: 0.79; 95% CI: 0.66, 0.96) - IS: 1.7 per 100 PY in intervention arm vs 2.1 per 100 PY in control arm (adjusted HR: 0.75; 95% CI: 0.57- 0.99) - AF: 11.4% in intervention arm vs 7.7% in control arm (p<0.01)
STROKESTOP (2012-2014)[32]	7165	SE	Age 75 or 76 years residing in two regions in Sweden	Randomised 1:1 to recording multiple 1L-ECGs per day for 14 days, or usual care	Median 6.9 years	- Composite of IS or haemorrhagic stroke, systemic embolism, bleeding leading to hospitalisation, and all-cause death*: 5.45 per 100 PY in the intervention arm vs 5.68 per 100 PY in the control arm (HR: 0.96; 95% CI: 0.92-1.00; p=0.045) - IS: 0.90 per 100 PY in the intervention arm vs 0.98 per 100 PY in the control arm (HR: 0.92; 95% CI: 0.83–1.01)
LOOP (2014-2016)[10]	6004	DK	Age 70-90 years with at least one additional risk factor (hypertension, diabetes, heart failure or prior stroke)	Randomised 1:3 to ILR insertion, or usual care	Median 5.4 years	- IS or systemic embolism*: 4.5% in intervention arm vs 5.6 in control arm (HR: 0.80; 95% CI: 0.61-1.05) - OAC initiation: 29.7% in the intervention arm vs 13.1 in the control arm (HR: 2.72; 95% CI: 2.41-3.08)
Post-stroke setting						

Table 3. *Results of previous AF screening trials that reported long-term outcomes (continued)*

Trial (years of recruitment)	n	Country	Inclusion criteria	Intervention	Follow-up duration	Difference in reported outcomes
FIND-AF$_{RANDOMISED}$ (2013-2014)[25]	398	DE	Age ≥60 years with recent ischemic stroke (<7 days)	Randomised 1:1 to three 10-day Holters (at baseline, 3 months, and 6 months) or usual care (≥24-hour Holter with further monitoring at the physician's discretion)	1 year	- Recurrent IS, systemic embolism, or death: 3.7% in intervention arm vs 5.4% in control arm (absolute difference 1.7%; 95% CI: -2.5%-5.9%)
MonDAFIS (2014-2017)[26]	3465	DE	Stroke or TIA patients admitted <72 hours after onset of symptoms	Randomised 1:1 to 7-day in-hospital continuous monitoring or routine care (≥24-hour Holter with further monitoring at the physician's discretion)	12 months (OAC initiation) & 24 months (composite endpoint)	- OAC initiation*: 13.7% in intervention arm vs 11.8% in control arm (OR: 1.2; 95% CI 0.9-1.5) - Composite of recurrent IS, major bleeding, MI, or death: 13.5% in intervention arm vs 14.5% in control arm (HR: 0.9; 95% CI: 0.8-1.1) - All-cause death: 4.3% in intervention arm vs 6.0% in control arm (OR: 0.7; 95% CI: 0.5-0.9)
PER DIEM (2015-2017)[27]	300	CA	Patients with arterial ischemic stroke or clinical TIA but with evidence of infarction on brain imaging op to 6 months before inclusion	Randomised 1:1 to ILR insertion for 1 year, or a 30-day external loop recorder at baseline	1 year	- Recurrent IS: 3.3% in intervention arm vs 5.3% in control arm (absolute difference: -2.0%; 95% CI: -6.6-2.6)

AF, atrial fibrillation; CA, Canada; CI, confidence interval; DE, Germany; DK, Denmark; FIND-AF$_{RANDOMISED}$, Holter-electrocardiogram-monitoring in patients with acute ischaemic stroke; HR, hazard ratio; IS, ischemic stroke; LOOP, Atrial Fibrillation Detected by Continuous ECG Monitoring Using Implantable Loop Recorder to Prevent Stroke in High-risk Individuals; MI, myocardial infarction; MonDAFIS, MONitoring for Detection of Atrial Fibrillation in Ischemic Stroke; mSToPS, mHealth Screening to Prevent Strokes; OAC, oral anticoagulation; OR, odds ratio; PER DIEM, Effect of Implantable vs Prolonged External Electrocardiographic Monitoring on Atrial Fibrillation Detection in Patients With Ischemic Stroke; PY, person-years; SE, Sweden; STROKESTOP, Systematic ECG Screening for Atrial Fibrillation; USA, United States of America.

* Reported as the study's primary endpoint.

One could argue that screening-detected AF (SDAF) is 'earlier' on the journey from atrial cardiomyopathy, to self-limiting bouts of pAF, to more persistent types of AF.[43] There is also evidence that those with SDAF are generally younger and with lower concomitant cardiovascular disease than routine care-detected AF cases.[1,44] This suggests that SDAF patients are generally also 'earlier' on the continuum that is one's cardiovascular burden. Although speculative, it can then be theorised that a treatment approach aimed at long-term reduction of cardiovascular risk to reduce AF-related complications could be especially relevant to SDAF patients, while the gains traditionally seen from immediate OAC initiation are relatively less prominent in the SDAF patient group.

The ABC pathway

Recent years have seen the increasing emphasis on a more holistic approach to AF care that aims to further improve outcomes in AF patients. This framework has become known as the ABC (AF Better Care) approach, consisting of three pillars: **A**void stroke (with **A**nticoagulants); **B**etter symptom management; **C**ardiovascular risk and **C**omorbidity optimization (Table 4).[45]

Table 4 . *The ABC approach to optimal AF treatment*

ABC pillar	Treatment options
Avoid stroke	- Oral anticoagulants
Better symptom management	- Rhythm and rate control, depending on patient-reported and/or observed symptoms and heart rate
Cardiovascular risk and **C**omorbidity optimization	- Identify (cardiovascular) risk factors and comorbidities - Work with patient to improve modifiable risk factors - Provide treatment for (modifiable) risk factors and cardiovascular comorbidities

ABC, **A**void stroke (with **A**nticoagulants), **B**etter symptom management, and **C**ardiovascular and **C**omorbidity risk optimization.

The "A" pillar of the ABC approach involves the prevention of ischemic cerebral events through OAC treatment in those with an indication for stroke prophylaxis.[21,46,47] The "B" pillar encourages physicians to optimise rate as well as rhythm control in their AF patients.[21] While rate and rhythm control's main objective are to improve quality of life, recent work has indicated that rhythm control is also associated with better cardiovascular outcomes when initiated early after AF diagnosis.[48] The "C" pillar entails the early detection and management of the overall spectrum of a patient's cardiovascular risk and comorbidities, such as hypertension, heart failure, coronary artery disease, diabetes and sleep apnoea.[45] Providing good "C" care entails frequent monitoring of modifiable risk factors, with optimisation through e.g. lifestyle changes or treatment of concomitant conditions.[21]

So far, retrospective studies have indicated that AF patients with adherence to the full ABC approach had more favourable clinical outcomes over extended follow-up.[49] A dose-response relationship was observed between number of ABC pillars successfully adhered to based on AF patients' records, and the degree of reduction of adverse outcomes.[50-52] Adherence to the full ABC approach (optimal care provided for all three pillars) was found to be associated with up to 56% reduction of composite adverse cardiovascular events in AF patients.[50,53] However, the comprehensive ABC approach was found to only have been consistently offered in a minority of patients, with only 1 in every 25 complex AF patients (at highest risk of adverse outcomes) complying with all three pillars of ABC care.[50,53,54] Some researchers therefore estimate further reduction in adverse outcomes if the ABC pathway could be implemented in larger proportions of AF patients.[50,55]

All-in on integrated care?

A number of early prospective studies have also already demonstrated the potential of what is dubbed this 'integrated' approach to AF care.[56] Here, integrated care should be understood as treatment according to the holistic ABC principles (intra-patient axis) while recognising the importance of cooperation between healthcare disciplines (primary, secondary and paramedical care; the inter-healthcare provider axis).

A pivotal trial on integrated ABC approach to AF care in primary care was the ALL-IN trial, conducted by researchers from University Medical Centers Utrecht, The Netherlands.[57] This was a cluster-randomised, open-label trial that included AF patients aged ≥65 years from 26 primary care practices in The Netherlands. They were included between 2015-2017, with follow-up for the primary endpoint (mortality after 2 years) ending in 2018-2019.[57] Patients in the intervention arm (n = 527) received integrated AF care based on the ABC principles and coordinated in primary care, including quarterly check-ups. Control patients (n = 713) received what was standard of care at that time: AF care coordinated by cardiologists or anticoagulation clinics, which involved once-yearly check-ups in most patients. After 2 years, mortality in the primary care intervention arm was significantly lower (HR: 0.55; 95% CI: 0.37-0.82) after adjustment for age, sex and frailty index.[57]

The ALL-IN trial thus showed that integrated AF care in The Netherlands, according to the ABC principles and in close cooperation between healthcare disciplines, can safely be coordinated from primary care.[57] An important question that remains, is whether GP-led integrated AF care is superior to current routine primary care, given the continuously evolving standard of care for AF in The Netherlands.[58-60] Moreover, it remains to be seen whether GP-led integrated AF care is as effective among newly-

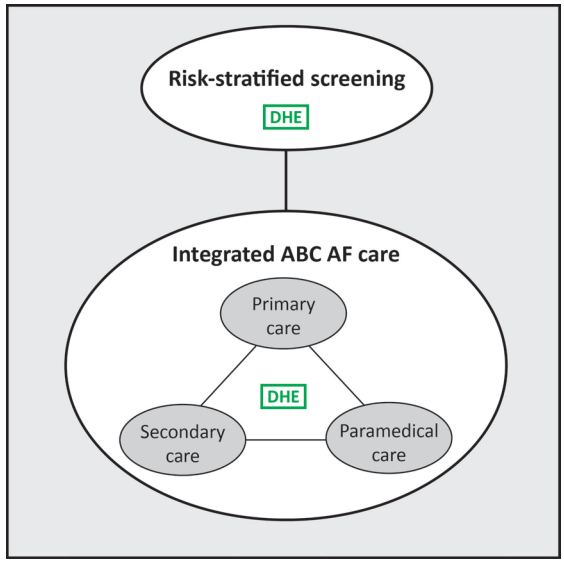
diagnosed SDAF patients as in those included in the ALL-IN trial (with median 4 years of known AF diagnosis).

PERSPECTIVES FOR FUTURE RESEARCH

In this thesis, we aimed to identify potential methods to increase the yield of AF screening through better patient selection. In the previous paragraphs, we discussed how risk-stratified screening indeed achieves higher AF yield than routine care, but that its net clinical benefit in terms of reducing long-term adverse outcomes has been counterintuitively low. We then showed that there is increasing evidence for the clinical benefit of an integrated, holistic approach to AF care, which could be especially relevant to SDAF patients. Finally, we showed evidence that such integrated AF care can safely be coordinated from primary care.

It is therefore worth investigating the potential benefit of combining risk-stratified AF screening with the latest insights on optimised AF care. This could take the form of a coupled intervention, linking risk-stratified screening to integrated ABC AF care. The combined intervention would be coordinated from primary care, in close cooperation with other healthcare disciplines, and would involve extensive follow-up for clinical outcomes (see Figure 1).

Figure 1. *Schematic outline of a potential future combined risk-stratified screening and integrated AF care intervention.*



ABC, Atrial Fibrillation Better Care; AF, atrial fibrillation DHE, digital health environment.

In this intervention, there is a prominent role for a digital health environment (DHE) that optimises both the risk stratification and integrated treatment phases of the trial. This DHE enables better linkage and compatibility of data across healthcare disciplines, employing advanced techniques where possible.[61] It also facilitates the use of mobile health technologies such as single-lead ECG or disease activity monitoring smartphone applications. The DHE expands the use of EHR data from mainly coded data (sex, age, or diagnosis and medication codes) to also include e.g. free text from consultations or specialist correspondence.[62,63] Through imputation as well as employment of AI techniques such as natural language processing and reinforcement learning, the DHE contributes to increased data completeness, higher accuracy of risk assessment, and better treatment recommendations. By actively suggesting addition of new or up-to-date anamnestic, physical or diagnostic information to the system, it further increases accuracy of risk assessment and treatment recommendations.[63,64] Finally, the DHE assists physicians in optimising adherence to the latest guidelines and best practices in light of the continually evolving and ever-increasing body of knowledge in the field. The DHE thereby contributes to an additional type of prevention: the prevention of *preventable suboptimal care provision*.

For the risk stratification scheme, CHARGE-AF seems a logical primary predictor for eligible patients in case of sufficient data, with CHA_2DS_2-VASc to fall back on in case of missing CHARGE-AF data. With the evolving evidence on PACs as a potent predictor for AF as well as clinical outcomes[65] after our publication featured in **Chapter 4**, it could be investigated whether 'frequent PACs' regardless of clinical prediction model score could be added as an entry variable for the screening intervention. Finally, given the ongoing development of AI-based pAF risk prediction based on raw ECG data (whether continuous ECG data analogous to **Chapter 8**, standard 12-lead ECGs[66] or even single-lead ECGs[67]), risk assessment based on routinely collected (raw) ECG data could also be considered for the risk stratification scheme.

As for the screening intervention, the available evidence should be systematically evaluated at time of protocol finalisation for their AF yield versus their burden and applicability. The PATCH-AF trial will show whether sequential (annual) monitoring is feasible and produces higher AF yield in high-risk Dutch primary care patients. As implantable loop recorder (ILR) devices are becoming more prevalent, with increasing understanding of their ease of use and potential complications, ILR insertion would be another candidate for the intervention.[23,27] Hybrid interventions, e.g. with (sequential) external continuous monitoring supplemented by ILR insertion in those in the highest-risk strata could also be considered.

The integrated AF care approach would be provided to all AF patients, whether detected through screening or through routine care. Follow-up would ideally extend to a minimum of 10 years, similar to other studies on cardiovascular risk management in primary care,[68] in order to assess the hypothesis that the integrated care approach is especially relevant for the subset of patients with SDAF. Collaboration should be sought with existing international consortia of AF screening research, aiming to conduct a similar intervention simultaneously across different settings. This would aim to provide further context to the question whether there are distinct features to the Dutch and other settings in terms of conduciveness of established routine care to AF detection and treatment, and the relative benefit of an active screening-and-treatment intervention.

Research into this intervention would also include work on stakeholder experiences with such an intervention, including work on potential harms of screening to the patient.[69] In the context of risk-stratified AF screening, the impact of being labelled 'high risk', or from screening results having inconclusive or incidental findings, should for instance be investigated.

Finally, there is a need for a comprehensive overview of the legal and regulatory framework around risk-stratified primary screening: from data ownership in remote risk prediction (at the discretion of the treating physician, or requirement for consent from each individual patient?) to inviting high-risk patients for screening, to potential legal or insurance implications from assessing a patient as high risk – both from the patient's and from a physician's (liability) perspective.

CONCLUSION

Atrial fibrillation (AF) is a common arrhythmia increasing in incidence with age. Due to its association with an increased risk of complications such as stroke or heart failure, there is extensive research aimed at early diagnosis and subsequent treatment of asymptomatic (silent) AF in order to prevent complication through early treatment initiation. Previous research on screening for AF in the community indicated that better patient selection was necessary in order to increase efficiency of AF screening efforts. In this thesis we therefore aimed to investigate potential triage tests for future AF screening interventions, as well as to validate means to capture AF in low-resource, community settings. We systematically reviewed and subsequently validated clinical risk prediction models, with CHARGE-AF showing high potential for use in primary AF screening, and CHA_2DS_2-VASc a viable alternative. We subsequently investigated PACs

on (continuous) ECG as a single biomarker for AF, and found that frequent PACs were indeed associated with AF, as well as with brain ischemia and stroke. We then validated an AI-enabled device for AF detection in primary care and saw that it had excellent diagnostic accuracy for AF when assessed by expert readers. Finally, we validated an AI algorithm that assesses risk of pAF on continuous Holter snippets for AF in older primary care and post-stroke patients at lower risk of AF, but found that the algorithm was insufficiently sensitive to serve as triage test for 14-day Holter monitoring in these patients. In light of the lessons from this thesis, as well as from previous work on improved outcomes through an integrated approach to AF treatment, we proposed further research which combines risk-stratified screening with an integrated AF care intervention in order to optimise outcomes in all AF patients – whether detected through routine care or through AF screening.

REFERENCES

1. Uittenbogaart SB, Verbiest-van Gurp N, Lucassen WAM, Winkens B, Nielen M, Erkens PMG, et al. Opportunistic screening versus usual care for detection of atrial fibrillation in primary care: cluster randomised controlled trial. BMJ. 2020;370:m3208.
2. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB, et al. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc. 2013;2(2):e000102.
3. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model. Stroke. 2016;47(3):895-900.
4. Kaasenbrood F, Hollander M, de Bruijn SH, Dolmans CP, Tieleman RG, Hoes AW, et al. Opportunistic screening versus usual care for diagnosing atrial fibrillation in general practice: a cluster randomised controlled trial. Br J Gen Pract. 2020.
5. Nadarajah R, Alsaeed E, Hurdus B, Aktaa S, Hogg D, Bates MGD, et al. Prediction of incident atrial fibrillation in community-based electronic health records: a systematic review with meta-analysis. Heart. 2022;108(13):1020-9.
6. Hulme OL, Khurshid S, Weng LC, Anderson CD, Wang EY, Ashburner JM, et al. Development and Validation of a Prediction Model for Atrial Fibrillation Using Electronic Health Records. JACC Clin Electrophysiol. 2019;5(11):1331-41.
7. Nadarajah R, Wu J, Hogg D, Raveendra K, Nakao YM, Nakao K, et al. Prediction of short-term atrial fibrillation risk using primary care electronic health records. Heart. 2023(0):1-8.
8. Segan L, Canovas R, Nanayakkara S, Chieng D, Prabhu S, Ling LH, et al. Development and validation of the HARMS2-AF lifestyle risk score to predict incident AF. Eur Heart J. 2022;43:2278-.
9. Af S, Collaborators A-E. Protocol for a Systematic Review and Individual Participant Data Meta-Analysis of Randomized Trials of Screening for Atrial Fibrillation to Prevent Stroke. Thromb Haemost. 2023;123(3):366-76.
10. Svendsen JH, Diederichsen SZ, Hojberg S, Krieger DW, Graff C, Kronborg C, et al. Implantable loop recorder detection of atrial fibrillation to prevent stroke (The LOOP Study): a randomised controlled trial. Lancet. 2021;398(10310):1507-16.
11. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. JAMA. 2018;320(2):146-55.
12. Halcox JPJ, Wareham K, Cardew A, Gilmore M, Barry JP, Phillips C, et al. Assessment of Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation: The REHEARSE-AF Study. Circulation. 2017;136(19):1784-94.
13. Kemp Gudmundsdottir K, Fredriksson T, Svennberg E, Al-Khalili F, Friberg L, Frykman V, et al. Stepwise mass screening for atrial fibrillation using N-terminal B-type natriuretic peptide: the STROKESTOP II study. Europace. 2020;22(1):24-32.
14. Gladstone DJ, Wachter R, Schmalstieg-Bahr K, Quinn FR, Hummers E, Ivers N, et al. Screening for Atrial Fibrillation in the Older Population: A Randomized Clinical Trial. JAMA Cardiol. 2021.
15. Rizas KD, Freyer L, Sappler N, von Stulpnagel L, Spielbichler P, Krasniqi A, et al. Smartphone-based screening for atrial fibrillation: a pragmatic randomized clinical trial. Nat Med. 2022;28(9):1823-30.

16. Brik T, Lucassen WAM, Harskamp RE, Karregat EPM, Himmelreich JCL, Busschers WB, et al. Personalized approach using wearable technology for early detection of atrial fibrillation in high-risk primary care patients (PATCH-AF): Study protocol for a cluster randomized controlled trial. Am Heart J. 2022;254:172-82.
17. Pessoa-Amorim G, Casadei B, Jones NR, A'Court C, Bulbulia R, Buck G, et al. Active Monitoring for AtriaL FIbrillation (AMALFI): protocol and pilot from a mail-based randomized trial of screening for subclinical atrial fibrillation in high-risk individuals. ESC Heart & Stroke. 2020:9.
18. 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(7616):383.
19. Kleindorfer DO, Towfighi A, Chaturvedi S, Cockroft KM, Gutierrez J, Lombardi-Hill D, et al. 2021 Guideline for the Prevention of Stroke in Patients With Stroke and Transient Ischemic Attack: A Guideline From the American Heart Association/American Stroke Association. Stroke. 2021;52(7):e364-e467.
20. Rubiera M, Aires A, Antonenko K, Lemeret S, Nolte CH, Putaala J, et al. European Stroke Organisation (ESO) guideline on screening for subclinical atrial fibrillation after stroke or transient ischaemic attack of undetermined origin. European Stroke Journal. 2022;7(3):CVII–CXXXIX.
21. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomstrom-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). Eur Heart J. 2021;42(5):373-498.
22. Hsu JY, Liu PP, Sposato LA, Huang HK, Liu AB, Lai EC, et al. Oral anticoagulant decreases stroke recurrence in patients with atrial fibrillation detected after stroke. Front Cardiovasc Med. 2022;9:929304.
23. Sanna T, Diener HC, Passman RS, Di Lazzaro V, Bernstein RA, Morillo CA, et al. Cryptogenic stroke and underlying atrial fibrillation. N Engl J Med. 2014;370(26):2478-86.
24. Gladstone DJ, Spring M, Dorian P, Panzov V, Thorpe KE, Hall J, et al. Atrial fibrillation in patients with cryptogenic stroke. N Engl J Med. 2014;370(26):2467-77.
25. Wachter R, Groschel K, Gelbrich G, Hamann GF, Kermer P, Liman J, et al. Holter-electrocardiogram-monitoring in patients with acute ischaemic stroke (Find-AFRANDOMISED): an open-label randomised controlled trial. Lancet Neurol. 2017;16(4):282-90.
26. Haeusler KG, Kirchhof P, Kunze C, Tutuncu S, Fiessler C, Malsch C, et al. Systematic monitoring for detection of atrial fibrillation in patients with acute ischaemic stroke (MonDAFIS): a randomised, open-label, multicentre study. Lancet Neurol. 2021;20(6):426-36.
27. Buck BH, Hill MD, Quinn FR, Butcher KS, Menon BK, Gulamhusein S, et al. Effect of Implantable vs Prolonged External Electrocardiographic Monitoring on Atrial Fibrillation Detection in Patients With Ischemic Stroke: The PER DIEM Randomized Clinical Trial. JAMA. 2021;325(21):2160-8.
28. Bernstein RA, Kamel H, Granger CB, Piccini JP, Sethi PP, Katz JM, et al. Effect of Long-term Continuous Cardiac Monitoring vs Usual Care on Detection of Atrial Fibrillation in Patients With Stroke Attributed to Large- or Small-Vessel Disease: The STROKE-AF Randomized Clinical Trial. JAMA. 2021;325(21):2169-77.

29. Benjamin EJ, Go AS, Desvigne-Nickens P, Anderson CD, Casadei B, Chen LY, et al. Research Priorities in Atrial Fibrillation Screening: A Report From a National Heart, Lung, and Blood Institute Virtual Workshop. Circulation. 2021;143(4):372-88.
30. Kahwati LC, Asher GN, Kadro ZO, Keen S, Ali R, Coker-Schwimmer E, et al. Screening for Atrial Fibrillation: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2022;327(4):368-83.
31. Steinhubl SR, Waalen J, Sanyal A, Edwards AM, Ariniello LM, Ebner GS, et al. Three year clinical outcomes in a nationwide, observational, siteless clinical trial of atrial fibrillation screening-mHealth Screening to Prevent Strokes (mSToPS). PLoS One. 2021;16(10):e0258276.
32. Svennberg E, Friberg L, Frykman V, Al-Khalili F, Engdahl J, Rosenqvist M. Clinical outcomes in systematic screening for atrial fibrillation (STROKESTOP): a multicentre, parallel group, unmasked, randomised controlled trial. Lancet. 2021;398(10310):1498-506.
33. Singer DE, Atlas SJ, Go AS, Lopes RD, Lubitz SA, McManus DD, et al. ReducinG stroke by screening for UndiAgnosed atRial fibrillation in elderly inDividuals (GUARD-AF): Rationale and design of the GUARD-AF randomized trial of screening for atrial fibrillation with a 14-day patch-based continuous ECG monitor. Am Heart J. 2022;249:76-85.
34. Williams K, Modi RN, Dymond A, Hoare S, Powell A, Burt J, et al. Cluster randomised controlled trial of screening for atrial fibrillation in people aged 70 years and over to reduce stroke: protocol for the pilot study for the SAFER trial. BMJ Open. 2022;12(9):e065066.
35. Janssen Scientific Affairs LLC. HEARTLINE - A Heart Health Study Using Digital Technology to Investigate if Early AF Diagnosis Reduces the Risk of Thromboembolic Events Like Stroke IN the Real-world Environment (ClinicalTrials.gov Identifier: NCT04276441) 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT04276441.
36. Wachter R. Intensive Rhythm Monitoring to Decrease Ischemic Stroke and Systemic Embolism - the Find-AF 2 Study (Find-AF2) (ClinicalTrials.gov Identifier: NCT04371055) 2020 [Available from: https://clinicaltrials.gov/ct2/show/NCT04371055.
37. Uittenbogaart SB, Lucassen WAM, van Etten-Jamaludin FS, de Groot JR, van Weert H. Burden of atrial high-rate episodes and risk of stroke: a systematic review. Europace. 2017.
38. Chen LY, Chung MK, Allen LA, Ezekowitz M, Furie KL, McCabe P, et al. Atrial Fibrillation Burden: Moving Beyond Atrial Fibrillation as a Binary Entity: A Scientific Statement From the American Heart Association. Circulation. 2018;137(20):e623-e44.
39. Harrison SL, Buckley BJR, Proietti R, Lip GYH. Atrial Fibrillation Diagnosed following Stroke: Dealing with a New Clinical Entity or Just a Matter of Definition? Cerebrovasc Dis. 2022;51(2):149-51.
40. Meng Y, Zhang Y, Zhu C, Nie C, Liu P, Chang S, et al. Atrial high-rate episode burden and stroke risks for patients with device-detected subclinical atrial fibrillation: A systematic review and meta-analysis. Int J Cardiol. 2023;371:211-20.
41. Wineinger NE, Barrett PM, Zhang Y, Irfanullah I, Muse ED, Steinhubl SR, et al. Identification of paroxysmal atrial fibrillation subtypes in over 13,000 individuals. Heart Rhythm. 2019;16(1):26-30.
42. Potpara TS, Lip GYH, Blomstrom-Lundqvist C, Boriani G, Van Gelder IC, Heidbuchel H, et al. The 4S-AF Scheme (Stroke Risk; Symptoms; Severity of Burden; Substrate): A Novel Approach to In-Depth Characterization (Rather than Classification) of Atrial Fibrillation. Thromb Haemost. 2021;121(3):270-8.

43. Schnabel RB, Marinelli EA, Arbelo E, Boriani G, Boveda S, Buckley CM, et al. Early diagnosis and better rhythm management to improve outcomes in patients with atrial fibrillation: the 8th AFNET/EHRA consensus conference. Europace. 2022.
44. Zwartkruis VW, Geelhoed B, Suthahar N, Bakker SJL, Gansevoort RT, van Gelder IC, et al. Atrial fibrillation detected at screening is not a benign condition: outcomes in screen-detected versus clinically detected atrial fibrillation. Results from the Prevention of Renal and Vascular End-stage Disease (PREVEND) study. Open Heart. 2021;8(2).
45. Lip GYH. The ABC pathway: an integrated approach to improve AF management. Nature Reviews Cardiology. 2017;14(11):627-+.
46. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: Antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med. 2007;146(12):857-67.
47. Gibbs H, Freedman B, Rosenqvist M, Virdone S, Mahmeed WA, Ambrosio G, et al. Clinical Outcomes in Asymptomatic and Symptomatic Atrial Fibrillation Presentations in GARFIELD-AF: Implications for AF Screening. Am J Med. 2021.
48. Kirchhof P, Camm AJ, Goette A, Brandes A, Eckardt L, Elvan A, et al. Early Rhythm-Control Therapy in Patients with Atrial Fibrillation. N Engl J Med. 2020;383(14):1305-16.
49. Proietti M, Romiti GF, Olshansky B, Lane DA, Lip GYH. Improved Outcomes by Integrated Care of Anticoagulated Patients with Atrial Fibrillation Using the Simple ABC (Atrial Fibrillation Better Care) Pathway. Am J Med. 2018;131(11):1359-66 e6.
50. Yoon M, Yang PS, Jang E, Yu HT, Kim TH, Uhm JS, et al. Improved Population-Based Clinical Outcomes of Patients with Atrial Fibrillation by Compliance with the Simple ABC (Atrial Fibrillation Better Care) Pathway for Integrated Care Management: A Nationwide Cohort Study. Thromb Haemost. 2019;119(10):1695-703.
51. Romiti GF, Proietti M, Vitolo M, Bonini N, Fawzy AM, Ding WY, et al. Clinical complexity and impact of the ABC (Atrial fibrillation Better Care) pathway in patients with atrial fibrillation: a report from the ESC-EHRA EURObservational Research Programme in AF General Long-Term Registry. BMC Med. 2022;20(1):326.
52. Romiti GF, Proietti M, Bonini N, Ding WY, Boriani G, Huisman MV, et al. Adherence to the Atrial Fibrillation Better Care (ABC) pathway and the risk of major outcomes in patients with atrial fibrillation: A post-hoc analysis from the prospective GLORIA-AF Registry. EClinicalMedicine. 2023;55:101757.
53. Pastori D, Menichelli D, Violi F, Pignatelli P, Lip GYH, group A-As. The Atrial fibrillation Better Care (ABC) pathway and cardiac complications in atrial fibrillation: a potential sex-based difference. The ATHERO-AF study. Eur J Intern Med. 2021;85:80-5.
54. Proietti M, Romiti GF, Olshansky B, Lane DA, Lip GYH. Comprehensive Management With the ABC (Atrial Fibrillation Better Care) Pathway in Clinically Complex Patients With Atrial Fibrillation: A Post Hoc Ancillary Analysis From the AFFIRM Trial. J Am Heart Assoc. 2020;9(10):e014932.
55. Proietti M, Lip GYH, Laroche C, Fauchier L, Marin F, Nabauer M, et al. Relation of outcomes to ABC (Atrial Fibrillation Better Care) pathway adherent care in European patients with atrial fibrillation: an analysis from the ESC-EHRA EORP Atrial Fibrillation General Long-Term (AFGen LT) Registry. Europace. 2021;23(2):174-83.
56. Guo YT, Lane DA, Wang LM, Zhang H, Wang H, Zhang W, et al. Mobile Health Technology to Improve Care for Patients With Atrial Fibrillation. J Am Coll Cardiol. 2020;75(13):1523-34.

57. van den Dries CJ, van Doorn S, Rutten FH, Oudega R, van de Leur SJCM, Elvan A, et al. Integrated management of atrial fibrillation in primary care: results of the ALL-IN cluster randomized trial. Eur Heart J. 2020;41(30):2836-44.
58. Koetsier DW, De Jong JSSG. Zorgprogramma integrale chronische zorg bij atriumfibrilleren (AF) - Amsterdamse Huisartsenalliantie 2022 [Available from: https://www.amsterdamse-huisartsen.nl/onze-projecten/atriumfibrilleren.
59. Huisartsenzorg Drenthe. Zorgprotocol Atriumfibrilleren: Integrale Eerstelijnszorg in Drenthe 2021 [Available from: https://www.hzd.nu/zorgprogrammas/atriumfibrilleren/het-protocol/$18711/$18712.
60. Chu G, Seelig J, Trinks-Roerdink EM, Geersing GJ, Rutten FH, de Groot JR, et al. Antithrombotic management of patients with atrial fibrillation-Dutch anticoagulant initiatives anno 2020. Neth Heart J. 2020;28(Suppl 1):19-24.
61. Han Y, Zhang Y, Vermund SH. Blockchain Technology for Electronic Health Records. Int J Environ Res Public Health. 2022;19(23).
62. De Clercq L, Schut MC, Bossuyt PMM, van Weert H, Handoko ML, Harskamp RE. TARGET-HF: developing a model for detecting incident heart failure among symptomatic patients in general practice using routine health care data. Fam Pract. 2023;40(1):188-94.
63. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. N Engl J Med. 2023;388(13):1201-8.
64. Zuo L, Du X, Zhao W, Jiang C, Xia S, He L, et al. Improving Anticoagulant Treatment Strategies of Atrial Fibrillation Using Reinforcement Learning. AMIA Annu Symp Proc. 2020;2020:1431-40.
65. Huang TC, Lee PT, Huang MS, Su PF, Liu PY. Higher premature atrial complex burden from the Holter examination predicts poor cardiovascular outcome. Sci Rep. 2021;11(1).
66. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet. 2019.
67. Hygrell T, Viberg F, Dahlberg E, Charlton PH, Kemp Gudmundsdottir K, Mant J, et al. An artificial intelligence-based model for prediction of atrial fibrillation from single-lead sinus rhythm electrocardiograms facilitating screening. Europace. 2023;25(4):1332-8.
68. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). Eur Heart J. 2016;37(29):2315-81.
69. Harris RP, Sheridan SL, Lewis CL, Barclay C, Vu MB, Kistler CE, et al. The Harms of Screening A Proposed Taxonomy and Application to Lung Cancer Screening. Jama Intern Med. 2014;174(2):281-5.

10

Summary

‘Early identification of atrial fibrillation in primary care and post-stroke patients’

SUMMARY

Atrial fibrillation (AF) is a common arrhythmia that can lead to complications, such as stroke and heart failure. The prevailing hypothesis is that timely detection and treatment of AF in the community, including population-based screening efforts, will reduce these risks. A key question is which patients should be screened for AF. As AF is mostly found in older people, screening individuals aged 65 or older could be a starting point. However, in countries with strong community-based healthcare systems and high AF awareness, there is already a high AF yield through routine care in older patients. Opportunistic AF screening in otherwise unselected older Dutch patients at the general practitioner's office has been shown not to lead to more AF cases. Such broad screening was thus not cost-effective. This does not mean that all AF cases will be found through routine care before AF-related complications occur. On the contrary: we still observe debilitating strokes as the first manifestation of AF. This means that we have to rethink how to design effective AF screening programs that provide additional benefit to an already strong usual care. The aim of this thesis was to find ways to more accurately predict a person's risk of AF than using only one's age. For this purpose, we explored risk prediction models and electrocardiographic (ECG) signatures of AF. Moreover, we aimed to provide more user-friendly ways to capture AF, by evaluating an unobtrusive, point-of-care diagnostic instrument. The subsequent paragraphs will contain a breakdown of these findings, organized in a chapter-by-chapter manner.

First in our investigation on how to better select patients for future community AF screening efforts we turned to clinical prediction models as a potential triage test. In **Chapter 2** we performed a systematic review and meta-analysis on risk models that had been developed and/or validated for AF in the community or primary care setting. We saw that, among the many models employed for this purpose, one model was likely the most accurate candidate for future patient selection in the community. This model, CHARGE-AF (Cohorts for Heart and Aging Research in Genomic Epidemiology model for atrial fibrillation), combined clinical and demographic patient information to estimate 5-year risk of AF among patients without an AF history.

In **Chapter 3** we performed our own validation of a number of the more promising multivariable risk models featured in **Chapter 2**, using a national routine primary care electronic health record (EHR) database in the Netherlands. We compared risk model performance against age alone for predicting 5-year risk of AF, in an attempt to test our hypothesis that risk models would be more efficient at selecting for 'high risk of future AF detection' than using age alone. We saw that this was the case in the overall sample, as well as many of the subgroups analysed in our validation. Impor-

tantly, however, we also exposed one of the vulnerabilities of CHARGE-AF for use in remote AF risk stratification. This was the model's reliance on a number of variables that – though easily identifiable – were often not systematically documented in Dutch routine primary care. As a secondary candidate, we therefore also validated CHA_2DS_2-VASc (Cardiac failure or dysfunction, Hypertension, Age >=75 [Doubled], Diabetes, Stroke [Doubled]-Vascular disease, Age 65-74, and Sex category [Female]) for prediction of future AF risk. Though originally developed for predicting risk of ischemic stroke in AF patients, and though showing more modest predictive ability for incident AF than CHARGE-AF, CHA_2DS_2-VASc had the advantage of consisting of parameters that are all readily available in routine primary care data.

We then zoomed in on an individual risk factor with potential predictive abilities for AF and other adverse outcomes: premature atrial contractions (PACs) on ECG. In **Chapter 4** we performed a systematic review and meta-analysis on the association between PACs on standard ECG or frequent PACs on continuous ECG and the risk of AF, brain ischemia or all-cause mortality. We found that after taking into account the significant heterogeneity among included studies, frequent PACs on Holter (a form of continuous ECG) were associated with AF, first stroke as well as all-cause mortality. There was a trend towards an association between PACs on 12-lead ECG and future AF detection, however due to the amount of heterogeneity in this analysis these findings were not conclusive. We concluded that PACs should not be regarded as the benign finding that they were traditionally held to be.

In **Chapter 5** we attempted to further assess the association between PACs on 12-lead ECG with adverse outcomes in people with type 2 diabetes (T2D) in Dutch primary care. In this clinically highly relevant group of patients due to increased risk of stroke in presence of AF and concomitant T2D, we saw that observing one or more PACs on 12-lead ECG was indeed associated with future AF detection. However, the association between PAC on ECG and ischemic events remained neutral in our analyses. Given the nature of the dataset with annual visits that included 12-lead ECG, we were also able to show that AF incidence was not associated with a higher risk of future ischemic events among this cohort of T2D patients – potentially testament to the quality of care received once enrolled in the program and with AF as a known, important comorbidity.

We then went on to investigating how best to screen for AF, with the aid of artificial intelligence (AI) algorithms applied to ECG signal. In **Chapter 6** we presented the results of a diagnostic accuracy study validating a hand-held, AI-enabled single-lead ECG device for detection of rhythm and conduction abnormalities including AF in prospectively enrolled primary care patients who underwent routine care indicated

12-lead ECG. We saw that this device, the AliveCor Kardia, had perfect properties for detecting and ruling out AF when the ECG signal was assessed by cardiologists. The in-built AI algorithm showed high specificity and sensitivity for AF, however slightly less accurate than the expert reader's assessment. Along with its high ease of use, these findings indicate that the AliveCor Kardia could be a viable candidate for use in AF screening and/or assessment for AF during home visits, provided that a positive algorithm reading is followed by visual assessment of the ECG signal by an experienced reader.

In **Chapter 7** we laid the groundwork for our final analysis, in which we aimed to validate an AI algorithm for assessing risk of paroxysmal AF (pAF) on Holter in high-risk patients (see **Chapter 8**). This validation required a new, prospective database of Holter recordings from patients screened for new AF in a high-risk cohort, in addition to the existing cohort of patients who had undergone 14-day Holter in the D_2AF trial (Detecting and Diagnosing AF). In **Chapter 7** we described the details of our prospectively enrolled cohort of consecutive patients presenting to the Amsterdam UMC, location AMC, for transient ischemic attack or ischemic stroke who underwent 14-day Holter for AF. In our final cohort of 379 patients, we saw that AF yield of 14-day Holter monitoring was much lower than expected compared to international literature despite good Holter adherence among participants. This was potentially due to sampling bias, as a comparison with a random sample of eligible non-included post-stroke patients showed that we included mainly younger patients with lower stroke severity and lower cardiovascular comorbidity. Still, our results were a valuable indication that a personalised approach within the current general recommendations for post-stroke cardiac monitoring could be considered for those at lowest risk of AF.

In **Chapter 8** we presented the results of our diagnostic accuracy study validating the AI algorithm that assesses the risk of underlying pAF during non-AF rhythm on the first 24 hours of Holter monitoring, with the outcome of total Holter (AF or no AF) as reference. We validated the algorithm in the post-stroke cohort featured in **Chapter 7** as well as in patients from the intervention arm of the cluster-randomised, controlled, D_2AF trial on primary AF screening, all of whom underwent 14-day Holter for AF. In both cohorts the rate of incident AF cases was low, with a majority also detected within the first 24 hours, considerably limiting our validation efforts. Still, if in these relatively low-risk samples the algorithm would have shown high sensitivity, the algorithm could be useful to safely rule out patients for further prolonged rhythm monitoring. The validation, however, showed that the AI algorithm's ability to further separate those at higher from those at lower risk of AF detecting during 14-day Holter monitoring was among included patients. We therefore concluded that the algorithm

was of insufficient diagnostic value to be used as a triage test for further monitoring up to 14 days in older primary care or post-stroke patients at lower risk of AF.

Finally, in **Chapter 9** we provided a discussion of the findings of this thesis within the context of the evolving research on risk-stratified screening for AF. The discussion was followed by a projection on the potential outline of future research that combines risk-stratified screening with an integrated AF care intervention, complemented by an advanced digital health environment, in order to optimise outcomes in all AF patients – whether detected through routine care or through an AF screening intervention.

Samenvatting

‘Early identification of atrial fibrillation in primary care and post-stroke patients’

SAMENVATTING

Atriumfibrilleren (AF) is een veelvoorkomende ritmestoornis die kan leiden tot complicaties zoals een herseninfarct of hartfalen. Algemeen wordt aangenomen dat vroegtijdig opsporen en behandelen van AF in de algehele bevolking, bijvoorbeeld door populatiegericht screenen, deze risico's kan verlagen. Een cruciale vraag hierbij is welke patiënten moeten worden gescreend op AF. Aangezien AF vooral wordt aangetroffen in ouderen, ligt screenen van mensen van 65 jaar of ouder voor de hand. Echter, in landen waarin reeds een sterk gezondheidssysteem aanwezig is met ruime aandacht voor AF wordt er onder ouderen reeds een groot aantal mensen met AF gevonden door middel van de gangbare zorg. Eerder is aangetoond dat opportunistische screening voor AF in anderszins ongeselecteerde patiënten van 65 jaar of ouder in Nederlandse huisartspraktijken weinig additionele AF-gevallen opspoort. Zulke brede screening was daarom niet kosteneffectief. Dit betekent niet dat alle AF-gevallen door gebruikelijke zorg worden opgespoord voordat deze tot complicaties kunnen leiden. In tegendeel: we zien nog steeds invaliderende herseninfarcten als eerste uiting van AF. Dit betekent dat we het ontwerp van effectieve screeningsprogramma's voor AF moeten herzien, om deze van toegevoegde waarde te laten zijn bovenop de reeds sterke gebruikelijke zorg. Het doel van dit proefschrift was om manieren te vinden waarmee we preciezer iemands risico op AF kunnen schatten dan alleen op basis van diens leeftijd. Voor dit doel hebben wij onderzoek gedaan naar risicomodellen voor en elektrocardiografische (ECG) vingerafdrukken van AF. Daarnaast wilden wij een meer gebruiksvriendelijke manier voor vaststellen van AF onderzoeken, middels evaluatie van een handzaam apparaat voor gebruik in de spreekkamer of aan het bed. Hieronder volgt een samenvatting van deze bevindingen, georganiseerd per hoofdstuk uit het proefschrift.

Als eerste in ons onderzoek naar hoe we beter patiënten konden selecteren voor toekomstige AF-screeningsonderzoek, keken we naar klinische predictiemodellen als mogelijk hulpmiddel voor triage. In **Hoofdstuk 2** verrichtten wij een systematische review en meta-analyse naar risicomodellen die waren opgesteld en/of gevalideerd voor voorspellen van AF in de algemene bevolking of eerstelijns zorgsetting. We zagen dat er, onder de vele modellen die voor dit doel bleken te zijn ingezet, één model was dat waarschijnlijk de meest precieze kandidaat was voor toekomstige patiëntselectie in populatiegerichte screening. Dit model, CHARGE-AF (*Cohorts for Heart and Aging Research in Genomic Epidemiology model for atrial fibrillation*), combineerde klinische en demografische patiëntinformatie om 5-jaars risico op AF te voorspellen in patiënten zonder AF in de voorgeschiedenis.

In **Hoofdstuk 3** valideerden wij een aantal van de in **Hoofdstuk 2** gevonden multivariabele risicomodellen binnen een landelijke database met elektronische patiëntdata (EPD) uit gebruikelijke eerstelijnszorg in Nederland. We vergeleken de prestaties van risicomodellen met die van alleen leeftijd voor voorspellen van 5-jaars risico op AF, met oog op testen van onze hypothese dat risicomodellen efficiënter zouden moeten zijn dan enkel gebruik van leeftijd voor het selecteren van patiënten voor 'hoog risico op toekomstige AF-detectie'. We zagen dat dit in het totale geselecteerde cohort het geval was, net als in vele van de uitgevoerde subgroepanalyses. Belangrijk was echter dat we ook een zwakte van CHARGE-AF blootlegden voor op afstand voorspellen van iemands AF-risico. Dit betrof de afhankelijkheid van het model van een aantal variabelen die – hoewel afzonderlijk gemakkelijk te bepalen – vaak niet systematisch werden vastgelegd in gangbare Nederlandse eerstelijnszorg. Als tweede kandidaat voor AF-risicovoorspelling valideerden we daarom ook de CHA_2DS_2-VASc score (*Cardiac failure or dysfunction, Hypertension, Age >=75 [Doubled], Diabetes, Stroke [Doubled]-Vascular disease, Age 65-74, and Sex category [Female]*). Hoewel deze oorspronkelijk was ontwikkeld voor voorspellen van risico op een herseninfarct in patiënten bij wie AF reeds was vastgesteld, en hoewel het model bescheidener voorspelkracht voor AF had dan CHARGE-AF, had CHA_2DS_2-VASc het voordeel dat het volledig bestond uit variabelen die gemakkelijk en uniform uit een eerstelijns EPD voor gebruikelijke zorg zijn op te maken.

Vervolgens zoomden wij in op een enkele risicofactor met potentiële voorspelkracht voor AF en andere negatieve uitkomsten: premature atriale complexen (PAC's) op ECG. In **Hoofdstuk 4** voerden wij een systematische review en meta-analyse uit naar het verband tussen PAC's op standaard-ECG of frequente PAC's op continu ECG en het risico op AF, ischemische beroerte of algehele sterfte. We zagen dat, ondanks de grote verscheidenheid aan geïncludeerde studies, frequente PAC's op 'Holter' (een vorm van continu ECG) geassocieerd waren met AF, een eerste herseninfarct, en algehele sterfte. Er werd een neiging gezien naar een verband tussen één of meer PAC's op 12-kanaals ECG en toekomstig AF, maar deze analyse was niet geheel sluitend door de grote verscheidenheid aan meegenomen studies. We concludeerden dat PAC's niet moeten worden gezien als de goedaardige bevinding waar ze traditioneel voor werden aangezien.

In **Hoofdstuk 5** probeerden we de associatie tussen PAC's op 12-kanaals ECG en negatieve gezondheidsuitkomsten verder in kaart te brengen in mensen met type 2 diabetes (T2D) in Nederlands eerstelijnszorg. In deze klinisch zeer relevante groep patiënten, vanwege het verhoogde beroerterisico in aanwezigheid van zowel AF als T2D, zagen we dat het zien van één of meer PAC's op 12-kanaals ECG inderdaad

geassocieerd was met kans op AF-detectie in de toekomst. Echter, het verband tussen PAC's op ECG en ischemische beroerte bleef neutraal in onze analyses. Gegeven de aard van de dataset met jaarlijkse controle met 12-kanaals ECG, waren we ook in staat om te laten zien dat het optreden van nieuw AF niet geassocieerd was met een hoger risico op latere ischemische beroertes in dit cohort van T2D-patiënten. Mogelijk is dit te danken aan de kwaliteit van zorg die men kreeg zodra in een T2D-patiënt ook AF werd vastgesteld als belangrijke nevendiagnose.

We verlegden ons toen naar het onderzoeken hoe we het beste konden screenen op AF met de hulp van kunstmatige intelligentie (AI; *artificial intelligence*) toegepast op ECG-signaal. In **Hoofdstuk 6** presenteerden we de resultaten van een diagnostische accuratessestudie met validatie van een handzaam, AI-gestuurd één-afleiding ECG-apparaat voor vaststellen van ritme- en geleidingsstoornissen, waaronder AF. Dit deden wij in prospectief geïncludeerde patiënten die 12-afleiding ECG hadden ondergaan tijdens gebruikelijke eerstelijnszorg, voor welke reden dan ook. We zagen dat dit apparaat, de AliveCor Kardia, perfecte eigenschappen had voor aantonen en uitsluiten van AF wanneer het ECG-signaal werd beoordeeld door cardiologen. Het ingebouwde AI-algoritme had weliswaar hoge specificiteit en sensitiviteit voor AF, maar was minder accuraat dan wanneer experts het ECG-signaal zelf lazen. Samen met het hoge gebruiksgemak, wezen deze bevindingen erop dat de AliveCor Kardia een mogelijke kandidaat was voor toepassing in AF-screening en/of diagnostiek naar AF tijdens huisbezoek, mits een positieve algoritme-uitslag gevolgd wordt door beoordeling van het ECG-signaal door een ervaren beoordelaar.

In **Hoofdstuk 7** legden we de basis voor onze laatste analyse, waarin we als doel hadden om een AI-algoritme te valideren voor schatten van het risico op onderliggend paroxismaal (aanvalsgewijs) AF (pAF) op Holter in hoog-risicopatiënten (zie **Hoofdstuk 8**). Deze validatie vereiste aanleg van een nieuwe, prospectieve dataset met Holteropnames gemaakt bij screenen op nieuw AF in een hoog-risicocohort – naast een reeds aanwezig cohort van patiënten die 14-daagse Holter hadden ondergaan in kader van de D_2AF-studie (*Detecting and Diagnosing AF*). In **Hoofdstuk 7** beschreven we de details van ons prospectief verzamelde cohort van achtereenvolgende patiënten die zich hadden gepresenteerd bij het Amsterdam UMC, locatie AMC, wegens een TIA (*transient ischemic attack*) of herseninfarct en die daarna 14-daagse Holter hadden ondergaan op zoek naar AF. In ons uiteindelijke cohort van 379 patiënten zagen we dat de AF-opbrengst uit 14-daagse Holter veel lager was dan was te verwachten op basis van internationale literatuur, ondanks dat deelnemende patiënten zich grotendeels goed aan het Holter-protocol hadden gehouden. Dit kwam mogelijk door sampling bias, aangezien deelnemende patiënten veelal jonger waren met lager beroerterisico

en lagere cardiovasculaire comorbiditeit in vergelijking met een willekeurige selectie van niet-deelnemende patiënten die wel in aanmerking zouden zijn gekomen voor de studie. Desondanks waren onze resultaten een waardevolle aanwijzing dat een persoonlijke aanpak binnen de huidige richtlijnen voor ritmemonitoring na een herseninfarct of TIA overwogen kan worden voor patiënten met een relatief laag risico op AF.

In **Hoofdstuk 8** presenteerden we de resultaten van onze diagnostische accuratessestudie waarin we het AI-algoritme voor voorspellen van onderliggend pAF valideerden op de eerste 24 uur van een 14-daagse Holter-opname, tegen de uitkomst van de totale Holter (wel of geen AF) als referentie. We valideerden het algoritme zowel in het cohort na herseninfarct of TIA uit **Hoofdstuk 7**, als in de patiënten uit de interventie-arm van de cluster-gerandomiseerde, gecontroleerde D_2AF-studie naar eerstelijns AF-screening. Alle deelnemers hadden 14-daagse Holter voor AF ondergaan. In beide cohorten was het aantal nieuwe AF-gevallen laag, met een meerderheid van de AF-gevallen ook nog vastgesteld binnen de eerste 24 uur, hetgeen onze validatie sterk belemmerde. Als het AI-algoritme in deze relatieve laag-risicocohorten een hoge sensitiviteit had laten zien, dan zou het algoritme nog nuttig kunnen zijn geweest om veilig patiënten te ontslaan van verdere langdurige ritmemonitoring. De validatie liet echter zien dat de mogelijkheden van het AI-algoritme om verder laag van hoog risico te onderscheiden erg bescheiden waren onder de deelnemende patiënten. We concludeerden daarom dat het algoritme onvoldoende diagnostische waarde had om te worden gebruikt als triagetest voor verdere monitoring tot 14 dagen in oudere eerstelijnspatiënten en patiënten na herseninfarct of TIA met een relatief laag AF-risico.

Tot slot boden we in Hoofdstuk 9 een discussie over de bevindingen uit dit proefschrift binnen de context van het zich ontwikkelende onderzoeksveld naar risicogestratificeerd screenen op AF. De discussie werd gevolgd door een vergezicht op mogelijke contouren van toekomstig onderzoek dat risico-gestratificeerde AF-screening combineert met een geïntegreerde interventie voor AF-zorg, aangevuld met een geavanceerde digitale zorg-omgeving, met als doel op uitkomsten te verbeteren voor alle AF-patiënten – of deze nu zijn vastgesteld door gebruikelijke zorg of door AF-screening.

A

Contributing authors
PhD portfolio
Acknowledgements - Dankwoord
Curriculum vitae

CONTRIBUTING AUTHORS

Claire Aussems
Netherlands Institute for Health Services Research, Utrecht, The Netherlands
Contribution:
- Chapter 3: statistical analysis, valuable input to & final approval of the manuscript

Joline W.J. Beulens
Amsterdam University Medical Centers, Location Vrije Universiteit Amsterdam, Department of Epidemiology and Data Science, Amsterdam, The Netherlands
Amsterdam Cardiovascular Sciences, Diabetes & Metabolism, Amsterdam, The Netherlands
Contribution:
- Chapter 4: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

Marieke T. Blom
Amsterdam University Medical Centers, Location Vrije Universiteit Amsterdam, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 4: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

Patrick M.M. Bossuyt
Amsterdam University Medical Centers, Location Academic Medical Center, Department of Clinical Epidemiology, Amsterdam, The Netherlands
Amsterdam Public Health, Methodology, Amsterdam, The Netherlands
Contribution:
- Chapter 4: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

Wim B. Busschers
Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 8: statistical analysis, data interpretation, valuable input to & final approval of the manuscript

Lukas de Clercq
Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 8: data interpretation, valuable input to & final approval of the manuscript

Jonathan M. Coutinho
Amsterdam University Medical Centers, Location Academic Medical Center, Department of Neurology, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 7: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 8: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

Petra Elders
Amsterdam University Medical Centers, Location Vrije Universiteit Amsterdam, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 5: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

Faridi S. van Etten-Jamaludin
Amsterdam University Medical Centers, Location Academic Medical Center, Medical Library, Amsterdam, The Netherlands
Contribution:
- Chapter 4: co-design of the work, data search, valuable input to & final approval of the manuscript

Joris R. de Groot
Amsterdam University Medical Centers, Location Academic Medical Center, Department of Cardiology, Amsterdam, The Netherlands
Amsterdam Cardiovascular Sciences, Heart failure & arrhythmias, Amsterdam, The Netherlands
Contribution:
- Chapter 6: data interpretation, valuable input to & final approval of the manuscript

- Chapter 7: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

M. Louis Handoko
Amsterdam University Medical Centers, Location Vrije Universiteit Amsterdam, Department of Cardiology, Amsterdam, The Netherlands
Amsterdam Cardiovascular Sciences, Heart failure & arrhythmias, Amsterdam, The Netherlands
Contribution:
- Chapter 6: data interpretation, valuable input to & final approval of the manuscript

Peter P. Harms
Amsterdam University Medical Centers, Location Vrije Universiteit Amsterdam, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 5: co-design of the work, data collection, statistical analysis, data interpretation, valuable input to & final approval of the manuscript

Ralf E. Harskamp
Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 2: valuable input to & final approval of the manuscript
- Chapter 3: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 4: valuable input to & final approval of the manuscript
- Chapter 6: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 7: valuable input to & final approval of the manuscript
- Chapter 8: data interpretation, valuable input to & final approval of the manuscript

Martijn Heugen
Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:

- Chapter 4: co-design of the work, data collection, data interpretation, valuable input to & final approval of the manuscript

Evert P.M. Karregat
Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 6: data collection, data interpretation, valuable input to & final approval of the manuscript

Wim A.M. Lucassen
Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:
- Chapter 2: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 3: valuable input to & final approval of the manuscript
- Chapter 4: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 5: co-design of the work, valuable input to & final approval of the manuscript
- Chapter 6: valuable input to & final approval of the manuscript
- Chapter 7: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 8: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

Mark M.J. Nielen
Netherlands Institute for Health Services Research, Utrecht, The Netherlands
Contribution:
- Chapter 3: co-design of the work, statistical analysis, data interpretation, valuable input to & final approval of the manuscript

Giel Nijpels
Amsterdam University Medical Centers, Location Vrije Universiteit Amsterdam, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands

Contribution:

- Chapter 5: co-design of the work, data collection, valuable input to & final approval of the manuscript

Robin Nijveldt

Radboud University Medical Center, Department of Cardiology, Nijmegen, The Netherlands

Contribution:

- Chapter 6: data interpretation, valuable input to & final approval of the manuscript

Michiel Rienstra

Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Contribution:

- Chapter 2: valuable input to & final approval of the manuscript

Renate B. Schnabel

Department of General and Interventional Cardiology, University Heart Center Hamburg/German Center for Cardiovascular Research, Hamburg, Germany

Contribution:

- Chapter 2: valuable input to & final approval of the manuscript

Hanno L. Tan

Amsterdam University Medical Centers, Location Academic Medical Center, Department of Cardiology, Amsterdam, The Netherlands

Amsterdam Cardiovascular Sciences, Heart failure & arrhythmias, Amsterdam, The Netherlands

Contribution:

- Chapter 4: valuable input to & final approval of the manuscript

Lieke Veelers

Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands

Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands

Contribution:

- Chapter 2: co-design of the work, data collection, data interpretation, valuable input to & final approval of the manuscript

Henk C.P.M. van Weert
Amsterdam University Medical Centers, Location Academic Medical Center, Department of General Practice, Amsterdam, The Netherlands
Amsterdam Public Health, Personalized Medicine, Amsterdam, The Netherlands
Contribution:

- Chapter 2: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 3: valuable input to & final approval of the manuscript
- Chapter 4: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 6: valuable input to & final approval of the manuscript
- Chapter 7: co-design of the work, data interpretation, valuable input to & final approval of the manuscript
- Chapter 8: co-design of the work, data interpretation, valuable input to & final approval of the manuscript

PHD PORTFOLIO & LIST OF PUBLICATIONS

Name PhD student:	Jelle C.L. Himmelreich
PhD period:	2016-2023
Names of PhD supervisor(s) & co-supervisor(s):	prof. dr. Henk C.P.M. van Weert (*supervisor*), dr. Jonathan M. Coutinho (*co-supervisor*), dr. Ralf E. Harskamp (*co-supervisor*)

1. PhD training

PhD training	Year	ECTS
General courses		
- AMC world of science	2016	0.5
- Practical biostatistics	2016	0.5
- E-science	2016	0.3
- Basiscursus Regelgeving en Organisatie voor Klinisch onderzoekers (bROK)	2016	1.0
- Clinical data management	2017	0.3
- Clinical epidemiology: Systematic reviews	2017	0.5
- Qualitative health research	2017	0.5
- Clinical epidemiology: Evaluation of medical tests	2017	0.5
Specific courses		
- MRC Prognosis summer course, Keele University, Keele, United Kingdom	2017	1.0
- IPD meta-analysis, Keele University, Keele, United Kingdom	2019	0.8
Seminars, workshops and master classes		
- LOVAH wetenschapsdag	2016	0.3
- ZonMW/HGOG annual meeting, Utrecht	2017	0.3
- HBCD workgroup meeting, Amsterdam	2020	0.3
- Medical coding training, Thrombosis Research Institute London (*online seminar*)	2020	0.2
- Safety monitor training, Thrombosis Research Institute London (*online seminar*)	2020	0.2
- ZonMW/HGOG annual meeting, Utrecht	2022	0.3
Presentations		
Oral presentations:		
- European Society of Cardiology Congress, Munich	2018	0.5
- Amsterdam Public Health Annual Meeting, Amsterdam	2018	0.5
- North American Primary Care Research Group Annual meeting, Toronto	2019	0.5
- European Society of Cardiology Congress, Amsterdam	2023	0.5
Poster presentations:		
- HartVaatHAG, Utrecht	2017	0.4
- HartVaatHAG, Utrecht	2018	0.4
- European Society of Cardiology Congress, Paris	2019	0.4
- European Society of Cardiology Congress, online (*2x*)	2021	0.6
Webinar presentation:		

- Powerful Medical webinar, online	2023	0.3
(Inter)national conferences		
- HartVaatHAG conference, Utrecht (*presenter*)	2016	0.3
- HartVaatHAG conference, Utrecht (*presenter*)	2017	0.3
- Nederlands Huisartsen Genootschap Congres, Amsterdam (*participant*)	2017	0.3
- Nederlands Huisartsen Genootschap Wetenschapsdag, Utrecht (*participant*)	2017	0.3
- World Conferences on Research Integrity, Amsterdam (*participant*)	2017	0.3
- European Society of Cardiology Congress, Munich (*presenter*)	2018	1.0
- Amsterdam Public Health Annual Meeting, Amsterdam (*presenter*)	2018	0.3
- European Society of Cardiology Congress, Paris (*presenter*)	2019	1.0
- North American Primary Care Research Group Annual meeting, Toronto (*presenter*)	2019	1.0
- Amsterdam Public Health Annual Meeting, Amsterdam (*participant*)	2019	0.3
- European Society of Cardiology Congress, online (*participant*)	2020	0.4
- European Society of Cardiology Congress, online (*presenter*)	2021	0.4
- World Organization of Family Doctors, London (*participant*)	2022	0.8
- European Society of Cardiology Congress, Amsterdam (*presenter*)	2023	1.0
Other		
- Research fellowship, Nivel, Utrecht (Supervisor: Mark M.J. Nielen; part-time, 12 months)	2019	10
- Clinical research fellowship, Thrombosis Research Institute, London (supervisor: Karen S. Pieper; full-time, 9 months)	2020-2021	30
- Monthly journal club, department of General Practice, Amsterdam UMC, Amsterdam	2016-2023	1.0
- Monthly research meeting, department of General Practice, Amsterdam UMC, Amsterdam	2016-2023	1.0

2. Teaching

Teaching	Year	ECTS
Lecturing		
- None		
Tutoring, Mentoring		
- Tutor, Bachelor of Medicine, Amsterdam UMC, (courses: *"Schrijven van een PICO"*; *"Abstract voor wetenschappelijk onderzoek schrijven"*)	2016-2017	0.6
Supervising		
- Bachelor thesis student co-supervisor (2x)	2019	2.0
- Master thesis student co-supervisor	2020	0.5
Other		
- None		

3. Parameters of Esteem

Parameters of Esteem	Year

Grants	
- Stoffels-Hornstra Stichting grant	2017
Awards and Prizes	
- APH Personalized Medicine Travel Grant, Amsterdam Public Health Annual Meeting, Amsterdam, Netherlands (*won*)	2018
- NAPCRG Pearl Award, North American Primary Care Research Group Annual meeting, Toronto (*won*)	2019
- Wetenschapsprijs, Nederlands Huisartsen Genootschap (*nomination*)	2019
- Wetenschapsprijs, Nederlands Huisartsen Genootschap (*nomination*)	2020

4. Publications

Publications	Year
Peer reviewed	
Included in this thesis:	
- Himmelreich JCL, Veelers L, Lucassen WAM, Schnabel RB, Rienstra M, van Weert H, et al. Prediction models for atrial fibrillation applicable in the community: a systematic review and meta-analysis. Europace. 2020;22(5):684–94.	2020
- Himmelreich JCL, Lucassen WAM, Harskamp RE, Aussems C, van Weert H, Nielen MMJ. CHARGE-AF in a national routine primary care electronic health records database in the Netherlands: validation for 5-year risk of atrial fibrillation and implications for patient selection in atrial fibrillation screening. Open Heart. 2021;8(1):1-11.	2021
- Himmelreich JCL, Lucassen WAM, Heugen M, Bossuyt PMM, Tan HL, Harskamp RE, et al. Frequent premature atrial contractions are associated with atrial fibrillation, brain ischaemia, and mortality: a systematic review and meta-analysis. Europace. 2019;21(5):698-707	2019
- Himmelreich JCL, Karregat EPM, Lucassen WAM, van Weert H, de Groot JR, Handoko ML, et al. Diagnostic Accuracy of a Smartphone-Operated, Single-Lead Electrocardiography Device for Detection of Rhythm and Conduction Abnormalities in Primary Care. Ann Fam Med. 2019;17(5):403-11.	2019
- Himmelreich JCL, Lucassen WAM, Coutinho JM, Harskamp RE, de Groot JR, van Weert HCPM. 14-day Holter monitoring for atrial fibrillation after ischemic stroke: The yield of guideline-recommended monitoring duration. European Stroke Journal. 2022;00(0):1-11.	2022
Not included in this thesis:	
- Lucassen WAM, van Peet PG, Himmelreich JCL, Uittenbogaart SB. [Elderly patients and atrial fibrillation: age alone should not be a contra-indication for anticoagulant treatment]. Ned Tijdschr Geneeskd. 2019;163.	2019
- Harskamp RE, Laeven SC, Himmelreich JC, Lucassen WAM, van Weert H. Chest pain in general practice: a systematic review of prediction rules. BMJ Open. 2019;9(2):e027081.	2019
- Manten A, Cuijpers CJJ, Rietveld R, Groot E, van de Graaf F, Voerman S, et al. Rationale and design of a cohort study evaluating triage of acute chest pain in out-of-hours primary care in the Netherlands (TRACE). Prim Health Care Res. 2020;21.	2020

-	Karregat EPM, Himmelreich JCL, Lucassen WAM, Busschers WB, van Weert H, Harskamp RE. Evaluation of general practitioners' single-lead electrocardiogram interpretation skills: a case-vignette study. Fam Pract. 2020.	2020
-	Beers L, van Adrichem LP, Himmelreich JCL, Karregat EPM, de Jong JSSG, Postema PG, et al. Manual QT interval measurement with a smartphone-operated single-lead ECG versus 12-lead ECG: a within-patient diagnostic validation study in primary care. Bmj Open. 2021;11(11).	2021
-	Harskamp RE, Bekker L, Himmelreich JCL, De Clercq L, Karregat EPM, Sleeswijk ME, et al. Performance of popular pulse oximeters compared with simultaneous arterial oxygen saturation or clinical-grade pulse oximetry: a cross-sectional validation study in intensive care patients. Bmj Open Respir Res. 2021;8(1).	2021
-	Karregat EPM, Gurp NV, Bouwman AC, Uittenbogaart SB, Himmelreich JCL, Lucassen WAM, et al. Screening for paroxysmal atrial fibrillation in primary care using Holter monitoring and intermittent, ambulatory single-lead electrocardiography. Int J Cardiol. 2021;345:41-6.	2021
-	Witvliet MP, Karregat EPM, Himmelreich JCL, de Jong J, Lucassen WAM, Harskamp RE. Usefulness, pitfalls and interpretation of handheld singlelead electrocardiograms. J Electrocardiol. 2021;66:33-7.	2021
-	Ten Broeke CEM, Himmelreich JCL, Cals JWL, Lucassen WAM, Harskamp RE. The Roth score as a triage tool for detecting hypoxaemia in general practice: a diagnostic validation study in patients with possible COVID-19. Prim Health Care Res Dev. 2021;22:e56.	2021
-	Harskamp RE, Himmelreich JCL, Wong GWM, Teichert M. Prescription patterns of direct oral anticoagulants and concomitant use of interacting medications in the Netherlands. Neth Heart J. 2021;29(9):451-9.	2021
-	Harskamp RE, Kleton M, Smits IH, Manten A, Himmelreich JCL, van Weert H, et al. Performance of a simplified HEART score and HEART-GP score for evaluating chest pain in urgent primary care. Neth Heart J. 2021.	2021
-	Himmelreich JCL, Harskamp RE, Geelhoed B, Virdone S, Lucassen WAM, Gansevoort RT, et al. Validating risk models versus age alone for atrial fibrillation in a young Dutch population cohort: should atrial fibrillation risk prediction be expanded to younger community members? BMJ Open. 2022;12(2):e057476.	2022
-	Harskamp RE, Lucassen WAM, Lopes RD, Himmelreich JCL, Parati G, Weert H. Risk of stroke and bleeding in relation to hypertension in anticoagulated patients with atrial fibrillation: a meta-analysis of randomised controlled trials. Acta Cardiol. 2022;77(3):191-5.	2022
-	Cools F, Virdone S, Sawhney J, Lopes RD, Jacobson B, Arcelus JI, et al. Thromboprophylactic low-molecular-weight heparin versus standard of care in unvaccinated, at-risk outpatients with COVID-19 (ETHIC): an open-label, multicentre, randomised, controlled, phase 3b trial. Lancet Haematol. 2022;9(8):e594-e604.	2022
-	Brik T, Lucassen WAM, Harskamp RE, Karregat EPM, Himmelreich JCL, Busschers WB, et al. Personalized approach using wearable technology for early detection of atrial fibrillation in high-risk primary care patients (PATCH-AF): Study protocol for a cluster randomized controlled trial. Am Heart J. 2022;254:172-82.	2022

- Himmelreich JCL, Harskamp RE. Diagnostic accuracy of the PMcardio smartphone application for artificial intelligence–based interpretation of electrocardiograms in primary care (AMSTELHEART-1). Cardiovascular Digital Health Journal. 2023;0(0):1-11.	2023
Under review	
Included in this thesis:	
- Harms PP, Himmelreich JCL, Blom MT, Beulens JWJ, Nijpels G, Elders P, Lucassen WAM. Longitudinal association of premature atrial contractions with atrial fibrillation and brain ischemia in people with type 2 diabetes: The Hoorn Diabetes Care System cohort.	
- Himmelreich JCL, de Clercq L, Coutinho JM, Harskamp RE, Busschers WB, van Weert HCPM, Lucassen WAM. Validation of an algorithm for assessing risk of paroxysmal atrial fibrillation on continuous ECG versus 14-day Holter in primary care and post-stroke patients.	
Not included in this thesis:	
- Himmelreich JCL, Virdone S, Camm AJ, Pieper KS, Harskamp RE, Lucassen WAM, Kakkar AK, for the GARFIELD-AF investigators. Safety and efficacy of apixaban and rivaroxaban versus warfarin in real-world atrial fibrillation patients depend on selection criteria used by the original randomized trials: insights from the GARFIELD-AF registry.	
- Himmelreich JCL, Virdone S, Camm AJ, Pieper KS, Harskamp RE, Lucassen WAM, Oto A, Jacobson B, Sawhney JPS, Lim TW, Gibbs H, Goto S, Haas S, Fox KAA, Jansky P, Verheugt FWA, Kakkar AK, for the GARFIELD-AF investigators. Comparing Rivaroxaban and Apixaban in GARFIELD-AF according to ROCKET AF and ARISTOTLE trial selection criteria.	
- Karregat EPM, de Koning M, Himmelreich JCL, Koetsier D, de Jong JS, Harskamp RE, Lucassen WAM. Evaluation of the introduction of a single-lead ECG device and digital cardiologist consultation platform among general practitioners in The Netherlands.	
- Barco S, Virdone S, Götschi A, Ageno W, Arcelus JI, Bingisser R, Colucci G, Cools F, Duerschmied D, Gibbs H, Fumagalli R, Gerber B, Haas S, Himmelreich JCL, Hobbs FDR, Hobohm L, Jacobson B, Kayani G, Lopes RD, MacCallum P, Micieli E, Righini M, Robert-Ebadi H, Rocha AT, Rosemann T, Sawhney J, Schellong S, Sebastian T, Spirk D, Stortecky S, Turpie AGG, Voci D, Kucher N, Pieper KS, Held U, Kakkar AK; on behalf of the OVID and ETHIC Investigators. Enoxaparin for symptomatic COVID-19 managed in the ambulatory setting: an individual patient level analysis of the OVID and ETHIC trials.	
- Harskamp RE, De Clercq L, Er A, Handoko ML, van Weert HCPM, Schut M, Moll van Charante E. Characteristics of Heart Failure in the Amsterdam Metropolitan Area (AMSTERDAM-HF): Data from a Dynamic General Practice Cohort (2011-2021).	
Other	
- Himmelreich JCL, Harskamp RE. Cardiovasculair risicomanagement: indianenverhalen? Huisarts & Wetenschap. 2017;60(7):317.	2017
- Himmelreich JCL, Karregat EPM. App voor ecg-meting bij ritmestoornissen. Huisarts & Wetenschap. 2018;61(6):80.	2018
- Himmelreich JCL, Cuijpers CJJ. Diabeten in remissie na intensief dieet. Huisarts & Wetenschap. 2019;62(1):7.	2019

-	Himmelreich JCL. Veel diabeten blijven in remissie na intensief dieet. Huisarts & Wetenschap. 2019;62(9):6.	2019
-	Himmelreich JCL. A pint of milk a day keeps the doctor away. Huisarts & Wetenschap. URL: https://www.henw.org/artikelen/pint-milk-day-keeps-doctor-away. 2019.	2019
-	Himmelreich JCL, Harskamp RE. COVID-19 bij ouderen biedt beperkt bescherming tegen nieuwe besmetting. URL: https://www.henw.org/artikelen/covid-19-bij-ouderen-biedt-beperkt-bescherming-tegen-nieuwe-besmetting. 2021	2021

ACKNOWLEDGEMENTS - DANKWOORD

Dit proefschrift had niet tot stand kunnen komen zonder de hulp van velen die mij onderweg op diverse wijzen bijgestaan hebben.

Allereerst mijn dank aan alle deelnemers van de verschillende studies die deel uitmaken van dit proefschrift. Deelnemers aan de RAPID-AF studie op afdeling Neurologie, AMC, van wie ik velen zelf heb mogen includeren, dank voor uw grote inzet en voor de enthousiaste en hartelijke reacties op onze studie. Maar ook de vele duizenden anderen wier gegevens opgenomen zijn in database-onderzoeken waaruit de analyses in dit proefschrift zijn onderbouwd. Uw bereidheid om belangeloos uw ervaringen en medische gegevens te delen zijn van groot belang geweest voor mijzelf, ons onderzoeksteam, en de wetenschap als geheel. Deze bereidwilligheid is de basis voor het soort medisch onderzoek waarmee wij streven om de zorg beter te begrijpen – en waar mogelijk beter te maken. Moge dit elkaar gegund blijven worden.

Mijn promotieteam verdient bijzondere dank in totstandkoming van dit proefschrift. Henk, vanaf onze eerste ontmoeting heb ik mij gesteund gevoeld in mijn komst naar de afdeling voor zowel onderzoek als huisartsopleiding. Vooral in het begin van dit traject hebben we intensief alle lopende (D_2AF) en opstartende (RAPID-AF, systematic reviews) projecten doorgenomen, waar je met rake aanwijzingen alles in goede banen leidde. In latere jaren was de begeleiding minder intensief, maar steeds als we elkaar spraken, kwamen we tot mooie gesprekken met goede inzichten.

Jonathan, als stille kracht, maar met de sleutelrol in het tot stand komen van het post-stroke cohort, dankjewel voor alle hulp en gastvrijheid op afdeling Neurologie om RAPID-AF uit te voeren. Je bent een bijzonder goed wetenschapper met enorm breed overzicht waar ik veel van hoop mee te nemen naar nieuwe projecten. Wanneer de vuren soms heter werden en jij daarin iets kon betekenen, heb ik me steeds door jou gesteund gevoeld.

Ralf, sinds jij op een dag mijn kamer binnenstapte om je voor te stellen als nieuwe onderzoeker, hebben we elkaar gevonden. Als collega's, en als vrienden. Ik bewonder je creativiteit, je werkethos, je vermogen om contacten te maken en te onderhouden, en om alle belangen en afwegingen die bij onderzoek horen goed af te stemmen. Maar bovenal ben je heel sociaal en inclusief, waardoor werken met jou altijd prettig is. Door jouw visie en mentorschap hebben we al vele mooie plannen tot een mooi einde gebracht, en ik hoop nog vele jaren met je samen te werken.

Wim, jij verdient mijn bijzondere dank in mijn gehele AIOTO-traject. Op twee vlakken was jij mijn mentor, steun en toeverlaat. Als leider van onze onderzoeksgroep hebben we vele mooie projecten en ideeën uitgewerkt en tot mooie eindes weten te brengen. Als huisarts-in-opleiding heb ik het eerste jaar bij jou mogen groeien. Je deur stond altijd open, en je feedback was altijd opbouwend, waardoor je de veiligheid bood waarin ik het beste gedij. Jouw kracht is om ideeën van grote hoogte te aanschouwen en daarin ordening aan te brengen. Stukken werden zo altijd beter na jouw input. Je bent het gelukkigst als jouw pupillen de beste versie van zichzelf weten te vinden en blijft bescheiden over je eigen rol daarin, maar neem van mij aan dat die aanzienlijk was. Ook toen je besloot het onderzoeksveld als lijnhouder te zullen verlaten om verder te gaan als huisarts kwam je grootmoedige karakter naar voren door allereerst ruimte voor mij en anderen te bieden om jouw stokje binnen de groep over te nemen. Bedankt voor alles.

Eric, bedankt voor de mentorrol die ook jij de afgelopen jaren op je hebt genomen, voor mij en anderen binnen onze groep. Je positieve houding, overzicht en enorme ervaring hebben bijgedragen aan rust, en daarmee het werkplezier van ons allemaal. Mede dankzij jouw inspanningen kan ik de komende jaren verder als onderzoeker in onze prachtige groep.

Mijn dank ook aan de leden van de promotiecommissie, prof. dr. Yvo Roos, dr. Judy van Es, dr. Geert-Jan Geersing, prof. dr. Patrick Bindels, prof. dr. Patrick Bossuyt en prof. dr. Joris de Groot, voor de tijd en zorgvuldigheid waarmee u ook mijn proefschrift hebt beoordeeld. Het is mijn grote eer dit proefschrift ten overstaan van u te verdedigen.

In het bijzonder wil ik hieruit prof. dr. Joris de Groot bedanken voor de aanzienlijke steun die ik heb mogen ervaren op meerdere momenten tijdens mijn promotietraject. Je klinische en elektrofysiologische expertise zijn groots, en je inzichten zijn immer raak.

Mijn paranimfen, Misha en Wouter, wat fijn dat jullie mij willen bijstaan op deze bijzondere dag. Jullie vriendschap is mij dierbaar, jullie steun is mij vertrouwd. Ik had mij geen beter duo kunnen wensen.

Aan alle coauteurs binnen en buiten dit proefschrift mijn grote dank voor de samenwerking de afgelopen jaren. Van elk project heb ik geleerd, en met sommigen van jullie zijn er blijvende samenwerkingen ontstaan. De studenten met wie we tot mooie publicaties zijn gekomen, Martijn, Lieke, Lisa, Lisa, Luuk, Charlotte en Amine, succes met jullie carrières en wellicht tot ziens in de huisartsgeneeskunde!

De afgelopen jaren stonden niet alleen in het teken van onderzoek. In het AIOTO-traject werden onderzoek en onderwijs voor de huisartsopleiding door elkaar verweven. Graag wil ik alle stage-opleiders en collega's in de verschillende praktijken en ziekenhuisafdelingen bedanken voor de prettige samenwerking en de ruimte die ik kreeg om te groeien. Wim, Renée en alle medewerkers van Huisartsen Risdam, bedankt voor jullie inwijding in het huisartsenvak tijdens mijn eerste opleidingsjaar, en voor de ruimte om het onderzoek daarnaast aandacht te blijven kunnen geven. Jan, Jasper en Anna, bedankt voor jullie gastvrijheid om in snelle tijd ervaring op te doen met resp. gynaecologie, spoedeisende geneeskunde en acute psychiatrie. Anna, dankjewel voor de ruimte die je mij gaf om zoveel mogelijk bij mijn gezin te zijn toen ik in de nacht voor start van mijn stage voor het eerst vader werd én een week later de eerste COVID-19 lockdown inging. En aan alle collega's van Therapeuticum Naarden-Bussum, dank voor het warme bad waarin ik mijn laatste opleidingsjaar heb mogen volbrengen. Jullie hebben een prachtige combinatie van professionaliteit en betrokkenheid bij elkaar en jullie patiënten die jullie ongetwijfeld zullen vasthouden. Wouter, jouw rol was daarin essentieel. Je bent een geweldig mentor, bij wie ik me vanaf onze eerste ontmoeting veilig en betrokken heb gevoeld. Dankjewel voor de prachtige gesprekken en bespiegelingen – ook ver buiten de geneeskunde. Herinneringen die ik koester, en in komende jaren nog vaker hoop te mogen hernieuwen.

Op de afdeling Huisartsgeneeskunde zijn er velen die mij hebben geholpen. Een aantal daarvan wil ik graag noemen. Alice en Sylvia, bedankt voor jullie support in de eerste jaren van mijn onderzoek, en het mij wegwijs maken op de afdeling. Amber en Annelies, bij jullie kon ik altijd terecht met mijn vragen, waarvoor dank. En Eda, dankjewel voor je ondersteuning bij alle projecten in onze groep, en in afronden van mijn proefschrift in bijzonder. Tot slot ook dank aan alle medewerkers, docenten en studiebegeleiders van de huisartsopleiding en SBOH, in bijzonder Jeroen, Agnes, Chris en Marja – bedankt voor jullie navigatie in de wirwar van het AIOTO-schap, en het bieden van de flexibiliteit die dit vereist.

Mijn mede-onderzoekers op J2 hebben mijn dagen op het AMC vaak kleur gegeven. Rosalie, Amalia, Evert, Lucinda en Marianne, dank jullie wel voor de gezelligheid als mijn kamergenoten – toen we nog eigen kamers hadden..! *Amalia, thank you for your friendship and the kind presents from Indonesia whenever you arrived for a new stint of research in Amsterdam, we hope to see you again and to be able to show you our new home as well.* Wim, dankjewel voor alle goede gesprekken – vooral over statistiek maar toch ook vaak ver daarbuiten – en je hulp in het stap voor stap doorgronden van de analyses die nodig zijn in elke afzonderlijke situatie. Ik heb veel van je geleerd. Amy, Tessa, Lukas en Indra, wat fijn om zulke gedreven en talentvolle collega-onderzoekers

in onze groep te hebben, ik ben heel benieuwd naar wat er allemaal uit jullie projecten gaat komen. Steven en Nicole, tot slot, jullie hebben mij ingewijd in atriumfibrilleren-onderzoek. Van dichtbij heb ik jullie indrukwekkende D_2AF-studie uitgevoerd zien worden, waarbij ik veel heb geleerd, en ook zelf veel mooie D_2AF-praktijkbezoeken heb mogen afleggen. Dank dat ik van dit platform gebruik heb mogen maken, en er deels mijn eigen draai aan heb mogen geven.

Ook op de andere afdelingen waar mijn onderzoek heeft plaatsgevonden kon ik niet zonder de hulp van een groot aantal collega's. Harriët, dankjewel voor je hulpvaardigheid in de uitvoer van het onderzoek op afdeling Neurologie, AMC. Mijn kamergenoten, Guido, Rita en Tamar, bedankt voor de gezellige werkdagen op H2. De RAPID-AF student-assistenten, Erik, Ibtisam, Janice, Josje, Maxim en Nina, dank voor jullie enthousiasme en inzet bij includeren van alle patiënten. Mark, tijdens mijn stage bij Nivel was jij mijn steun en toeverlaat. Dankjewel voor je gastvrijheid en de vele uren die we samen hebben gezeten om de analyses te overdenken en te programmeren. Ik kijk terug op een mooi jaar in Utrecht! *Dear Karen and Saverio, thank you for the opportunity of being embedded in your amazing team at TRI, London. I have learned so much, and have been able to experience aspects of medical research that I would never have been able to otherwise. Despite not having physically met during the entirety of the fellowship due to COVID-19, I have felt your support and kindness as if I was there at Manresa Road.*

Mijn dank ook aan alle medewerkers buiten AMC die mijn onderzoek de afgelopen jaren hebben ondersteund. Esther, Dave en Robert, bedankt dat jullie met jullie team van Fysiologic ECG Services bereid zijn geweest na D_2AF ook mijn RAPID-onderzoek van jullie Holters te voorzien. Dank ook in bijzonder aan Dhr. John den Engelsman voor de grote inzet in het compatibel maken van het Fysiologic ECG format met dat van anderen. *Dear Christian, Helmut and dr. Hirtz at Apoplex Medical Technologies, thank you very much for your support in our project, and for the many hours spent at converting file formats in order to produce analysable data. Your kindness and willingness to offer help was exceptional.* Patricia en collega's bij Neurologie en Wetenschapsbureau, NWZ, bedankt voor jullie bereidheid om RAPID-AF ook te overwegen in Alkmaar. Toen vlak voor onze eerste inclusie COVID-19 uitbrak, was afstel de enig juiste keuze, maar weet dat ik jullie hulp en bereidwilligheid zeer heb gewaardeerd.

Mijn vrienden, familie en bandleden die mijn afwezigheid de afgelopen jaren hebben ervaren, met dit boekje heb ik hopelijk enige tastbare bewijslast aangeleverd in de categorie verzachtende omstandigheden. Hoe fijn is het dan om toch om af en

toe in jullie gezelschap te zijn en me direct als vanouds te voelen. Bedankt voor die onvoorwaardelijke steun.

Cas en Carla, bedankt voor de steun die ik steeds van jullie heb gevoeld. Jullie zijn geweldige opa en oma voor onze meiden, en zorgen voor rust door hoe jullie je over hen – en ons – te ontfermen. Ik weet dat ik altijd bij jullie terecht kan, en beseffende dat dit niet voor iedereen het geval is, tel ik deze zegening dubbel. Sanne en Max, mijn grote zus en kleine broer, we zien elkaar niet vaak, maar ik denk vaak met warmte aan jullie. Dank voor de steun en het vertrouwen die ik altijd en van overal van jullie ervaar. Paul en Conny, bedankt dat jullie mij vanaf dag één in jullie warme midden hebben opgenomen. In Middelburg komen we graag tot rust, ik hoop die rit nog vele jaren te mogen maken.

Tot slot mijn meiden, Anna, June en Isa. An, wat ben ik blij dat ik je 10 jaar geleden dat halfje wit aanbood. Ik vind bij jou de rust en het anker dat ik nodig heb, en geniet elke dag van het team dat ik mij met jou voel. Je bent een geweldige moeder, en je leert mij veel over mijzelf en onze meisies. Dankjewel voor je engelengeduld tijdens de vele overuren die in dit boekje zitten. June en Isa, wat is het een genot om jullie te zien opgroeien. Mijn blije, enthousiaste meiden met een plan, ik hoop nog vele jaren jullie ontdekkingen van dichtbij mee te kunnen beleven.

CURRICULUM VITAE

Jelle Himmelreich was born on 21 October 1986 in Heiloo, The Netherlands. In 1995 the family (parents Cas & Carla, and siblings Sanne & Max) moved to Oudorp. Jelle attended Openbaar Scholengemeenschap Willem Blaeu in Alkmaar where he graduated gymnasium *cum laude*. In 2005 he started his medical education at the Academic Medical Center (AMC; currently Amsterdam UMC, location AMC), Amsterdam.

During his teens and early academic career music was an important part of Jelle's life, playing bass and double bass in different blues, rock and folk formations, with appearances at *Paradiso*, *Melkweg*, *Bevrijdingsfestivals*, *Sziget Festival*, *De Wereld Draait Door* and many other venues. To satisfy his increasing interest in world affairs and global inequalities, Jelle also temporarily left medical education in 2009 to attend a pre-master and master's course in Political Sciences (subspecialty: International Relations) at University of Amsterdam, for which he attained his master's degree in 2011 with a thesis on access to essential medicines in the Global South.

In early 2012 Jelle returned to the AMC to start his medical internships ('*coschappen*'), and attained his medical degree in 2014. During his *coschappen*, with internships in Brussels as well as in a primary care clinic in London, and further aided by his experience working as a nurse assistant at an Amsterdam nursing home, he decided that he wished to become a general practitioner (GP). After his medical graduation, Jelle worked at UWV, Haarlem, as an occupational health physician, and as a resident at the department of Neurology at NWZ, Alkmaar. There he developed an interest for vascular neurology, and particularly for the association between cardiovascular risk and stroke. When he saw the opportunity for a combination of GP traineeship and PhD research on early identification of atrial fibrillation in primary care and post-stroke patients, he immediately applied, which resulted in the current thesis.

Jelle lives in Alkmaar with his wife Anna and daughters June and Isa, having moved there from their home in Amsterdam after graduating as a GP in 2022. He works as a GP locum in the Alkmaar region, and will continue his research at the Department of General Practice, Amsterdam UMC, as a postdoctoral researcher.