

# Identification of undiagnosed atrial fibrillation using a machine learning risk-prediction algorithm and diagnostic testing (PULsE-AI) in primary care: a multi-centre randomized controlled trial in England

Nathan R. Hill <sup>1</sup>, Lara Groves <sup>2</sup>, Carissa Dickerson<sup>2</sup>, Andreas Ochs<sup>2</sup>, Dong Pang<sup>2</sup>, Sarah Lawton<sup>3</sup>, Michael Hurst<sup>2</sup>, Kevin G. Pollock<sup>1</sup>, Daniel M. Sugrue<sup>2</sup>, Carmen Tsang <sup>2</sup>, Chris Arden<sup>4</sup>, David Wyn Davies<sup>5</sup>, Anne-Celine Martin<sup>6,7</sup>, Belinda Sandler<sup>1</sup>, Jason Gordon <sup>2\*</sup>, Usman Farooqui<sup>1</sup>, David Clifton<sup>8</sup>, Christian Mallen <sup>3</sup>, Jennifer Rogers<sup>9</sup>, Alan John Camm<sup>10</sup>, and Alexander T. Cohen<sup>11</sup>

<sup>1</sup>Bristol Myers Squibb Pharmaceutical Ltd, Uxbridge, UK; <sup>2</sup>Health Economics and Outcomes Research Ltd, Cardiff, UK; <sup>3</sup>School of Medicine, Keele University, Staffordshire, UK; <sup>4</sup>University Hospital Southampton, Southampton, UK; <sup>5</sup>St Mary's Hospital, London, UK; <sup>6</sup>Université de Paris, INSERM, Innovative Therapies in Haemostasis, F-75006 Paris, France; <sup>7</sup>Service de Cardiologie, AP-HP, Hôpital Européen Georges Pompidou, F-75015 Paris, France; <sup>8</sup>Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK; <sup>9</sup>PHASTAR, London, UK; <sup>10</sup>Cardiology Clinical Academic Group, Molecular & Clinical Sciences Research Institute, St George's University of London, London, UK; and <sup>11</sup>Department of Haematological Medicine, Guys and St Thomas' NHS Foundation Trust, King's College London, London, UK

Received 13 January 2022; revised 24 February 2022; accepted 22 March 2022; online publish-ahead-of-print 23 March 2022

## Aims

The aim of the PULsE-AI trial was to assess the effectiveness of a machine learning risk-prediction algorithm in conjunction with diagnostic testing for identifying undiagnosed atrial fibrillation (AF) in primary care in England.

## Methods and results

Eligible participants (aged  $\geq 30$  years without AF diagnosis;  $n = 23\,745$ ) from six general practices in England were randomized into intervention and control arms. Intervention arm participants, identified by the algorithm as high risk of undiagnosed AF ( $n = 944$ ), were invited for diagnostic testing ( $n = 256$  consented); those who did not accept the invitation, and all control arm participants, were managed routinely. The primary endpoint was the proportion of AF, atrial flutter, and fast atrial tachycardia diagnoses during the trial (June 2019–February 2021) in high-risk participants. Atrial fibrillation and related arrhythmias were diagnosed in 5.63% and 4.93% of high-risk participants in intervention and control arms, respectively {odds ratio (OR) [95% confidence interval (CI)]: 1.15 (0.77–1.73),  $P = 0.486$ }. Among intervention arm participants who underwent diagnostic testing (28.1%), 9.41% received AF and related arrhythmia diagnoses [vs. 4.93% (control); OR (95% CI): 2.24 (1.31–3.73),  $P = 0.003$ ].

## Conclusion

The AF risk-prediction algorithm accurately identified high-risk participants in both arms. While the proportions of AF and related arrhythmia diagnoses were not significantly different between high-risk arms, intervention arm participants who underwent diagnostic testing were twice as likely to receive arrhythmia diagnoses compared with routine care. The

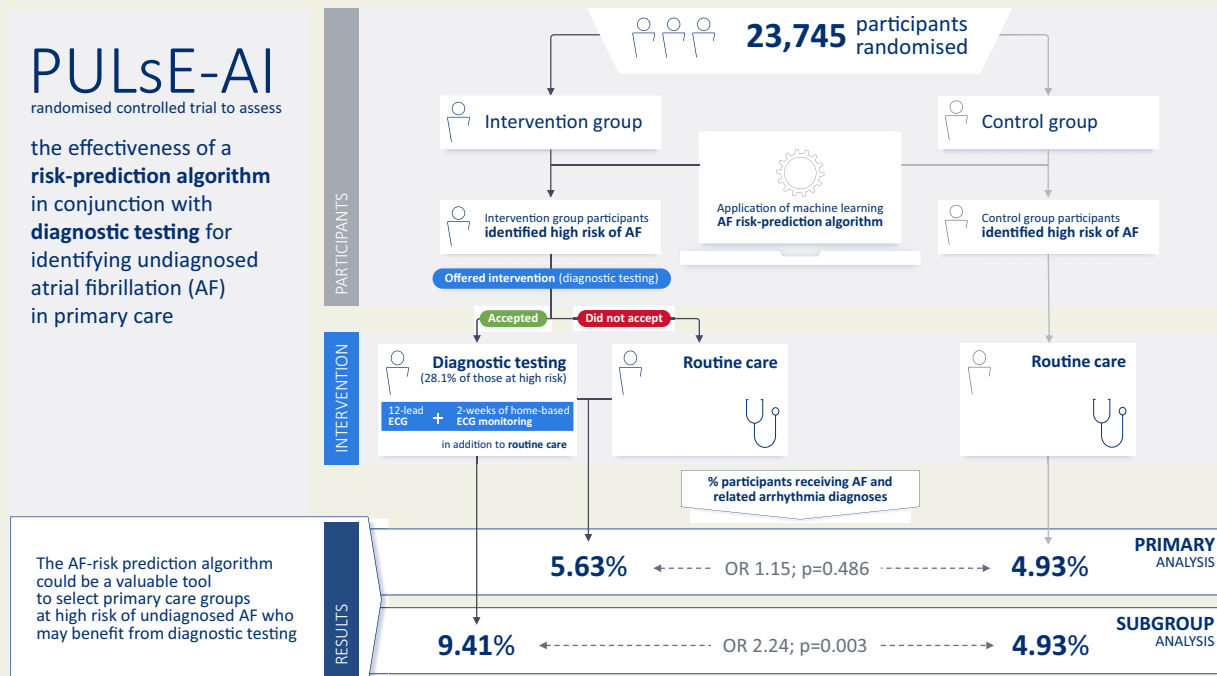
\* Corresponding author. Tel: +44 2920 399146, Email: [jason.gordon@heor.co.uk](mailto:jason.gordon@heor.co.uk)

© The Author(s) 2022. Published by Oxford University Press on behalf of European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

algorithm could be a valuable tool to select primary care groups at high risk of undiagnosed AF who may benefit from diagnostic testing.

## Graphical Abstract



## Keywords

Atrial fibrillation • Machine learning • Risk prediction • Primary care • Screening

## Introduction

Atrial fibrillation (AF) is a common cardiovascular health problem affecting ~3% of the general population, and is associated with a five-fold increase in the risk of stroke.<sup>1</sup> Cardioembolic strokes associated with AF are particularly severe, highly recurrent, and often fatal.<sup>2</sup> Atrial fibrillation-related stroke risk can be reduced by the early diagnosis of AF, risk stratification, and preventive oral anticoagulation treatment, but detection is challenging because AF may be both paroxysmal and asymptomatic.<sup>3</sup> Consequently, an estimated 300 000 people in the UK<sup>4</sup> and 700 000 people in the USA<sup>5</sup> are living with undiagnosed AF.

Screening for AF can be opportunistic [e.g. screening of patients attending their general practitioner (GP) for another reason], targeted (e.g. screening of higher risk patients), or systematic (e.g. screening all patients aged >65 years), and can include simple pulse checking or more resource-intensive electrocardiogram (ECG) assessment.<sup>3</sup> European Society of Cardiology (ESC) guidelines currently recommend opportunistic screening for AF by pulse check or ECG rhythm strip in patients aged >65 years (IB recommendation) to prevent AF-related complications.<sup>3</sup> This screening approach requires 70

people to be screened to identify one person with AF [number needed to screen (NNS) of 1 in 70] and consequently lacks cost-effectiveness.<sup>3</sup> Data to confirm the benefits of screening are scarce, and therefore, questions such as whom to screen, how to screen, and where to screen for the most effective and cost-effective screening strategy currently remain unanswered.<sup>6</sup>

People with undiagnosed AF comprise ~5% of all strokes in the UK.<sup>7</sup> Many of these strokes are potentially avoidable with early diagnosis and effective anticoagulation. Thus, there is an urgent need for an easily implementable, effective, and cost-effective screening strategy that can identify people at increased risk of undiagnosed AF who require further assessment via ECG or other diagnostic testing.<sup>8</sup> This is even more pertinent during (and beyond) the COVID-19 pandemic, with increased use of online and/or telephone primary care consultations reducing the opportunities to screen for AF using pulse check. The role of augmented/artificial intelligence (AI) in cardiovascular medicine is considered the next frontier in cardiovascular diagnostics, paving the way to the implementation of personalized strategies in cardiovascular therapeutics.<sup>9</sup>

Previously, we reported on the development and validation of an AF risk-prediction algorithm that utilized machine learning

techniques and did not require ECG data as an input.<sup>10,11</sup> The algorithm was trained, tested, and validated on retrospective data of patients ( $\geq 30$  years) in the UK Clinical Practice Research Datalink (CPRD) GOLD database and further validated against overfitting using the DISCOVER database for North West London.<sup>11</sup> The machine learning algorithm was effective at identifying patients with AF with an NNS of 9 (at 75% sensitivity and 99% specificity).<sup>10</sup> Here, we present findings from the Prediction of Undiagnosed atrial fibrillation using a machine learning Algorithm (PULSe-AI) randomized controlled trial to assess the real-world ability of a strategy including the machine learning-based AF risk-prediction algorithm coupled with diagnostic testing (ECG  $\pm$  KardiaMobile) to identify cases of AF compared with routine clinical care.

## Methods

### Study design and participants

The PULSe-AI trial is a prospective, randomized controlled trial conducted across six general practices within the National Institute for Health Research Clinical Research Network: West Midlands, England, UK. Adult patients registered at participating general practices were identified from medical records by staff at the participating practices. Patients were eligible for participation in the study if they were aged  $\geq 30$  years and without a prior diagnosis of AF, atrial flutter, or fast atrial tachycardia, and had a complete set of key clinical measurements [height, weight, body mass index (BMI), systolic blood pressure (SBP), and diastolic blood pressure (DBP); i.e. they had a valid index date] recorded during a rolling 12-month window ending at any time during the 11 years prior to 6 June 2019 (date of first data extraction from practice records). Trial eligibility was limited to patients aged  $\geq 30$  years to align with the algorithm development population.<sup>10</sup>

In compliance with data protection regulations,<sup>12</sup> consent for the extraction of pseudonymized patient data from medical records was provided by each general practice participating in the study; written informed consent was obtained from the subgroup of participants in the intervention arm at high risk of undiagnosed AF who attended the research clinic for diagnostic testing. The trial complies with the Declaration of Helsinki and ethical approval was granted by the Wales Research Ethics Committee 5—Bangor, and the study was approved by the Health Research Authority and Health and Care Research Wales (IRAS project ID: 252934). The trial protocol has been published previously.<sup>13</sup>

### Randomization and masking

All eligible participants across the six sites were pooled and individually randomized using simple randomization into intervention and control arms in a 1:1 ratio to receive routine care plus the screening strategy (intervention arm) or routine care only (control arm). Randomization was performed centrally by the trial statistician using R (version 3.5.3) code. Following randomization, the machine learning AF risk-prediction algorithm was run, and a risk score was generated for each individual. Participants and local study staff could not be masked to the allocated arms during the trial, but blinded analyses were undertaken by the trial team.

### Procedures

Prior to the invitation, GPs reviewed the lists of participants randomized to the intervention arm at high risk of AF and excluded those deemed

clinically unsuitable to participate in the study (including those recently deceased or receiving end-of-life care, those receiving treatment for active cancer, those housebound or residing in a care facility and unable to attend the research clinic, those with solid organ transplant, and those with anxiety- or mental health-related illness whose mental health may be negatively affected by being identified as at high risk of undiagnosed AF). Participants in the intervention arm at high risk of AF and suitable for participation in the study were invited via letter by their general practice to attend a research clinic for diagnostic testing. Participants who did not respond to the first letter of invitation were sent a reminder letter and practices were able to follow-up with the remaining non-responders once via phone. Those who accepted the invitation had updated baseline characteristics, lifestyle-related variables, basic cardiorespiratory variables, recent clinical history, and presence of symptoms recorded by the research nurse and received a 12-lead ECG. Intervention arm participants with a negative or an unconfirmed diagnosis following the 12-lead ECG were provided with a single-lead ECG KardiaMobile portable device (AliveCor Inc., CA, USA). Participants without access to a compatible smartphone or tablet were offered a loan smartphone provided by the trial team to use alongside the KardiaMobile for the study period. If declined, participants were invited for up to two further 12-lead ECGs. Participants using the KardiaMobile were asked to record their ECG twice daily (morning and evening), in addition to any time they felt unwell or at their discretion, for 2 weeks. All participants allocated to the control arm, low-risk participants in the intervention arm, and high-risk participants in the intervention arm who did not accept the invitation to participate in the study had no direct contact with the investigators.

Pseudonymized patient-level data from participating study sites were extracted from the Egton Medical Information System (EMIS) for all eligible participants using the EMIS search and report function. Participants who attended the research clinic had additional data collected during and after the research clinic appointment; these variables were entered into EMIS by practice staff for each participant through a bespoke designed EMIS template. All ECG data were pseudonymized and sent via encrypted email for cardiologist review. All KardiaMobile data were uploaded to secure servers in Europe based on unique study identifiers.

### Outcomes

The primary endpoint was the proportion of diagnoses of AF and related arrhythmias including atrial flutter and fast atrial tachycardia in the intervention and control cohorts throughout the trial. The primary analysis considered participants deemed to be at high risk of undiagnosed AF according to the algorithm (high-risk population) in both arms; analyses were also undertaken in the subgroup of participants who attended the research clinic (research clinic population, regardless of whether they completed the entire intervention), in the subgroup of those in the research clinic population who completed the intervention as per the protocol (12-lead ECG followed by KardiaMobile monitoring; per-protocol population), and in the full population (all randomized participants irrespective of AF risk-prediction score) (see [Supplementary material online, Table S1](#) for descriptions of analysis populations).

Participants were diagnosed with AF by the trial cardiologist if they had  $\geq 30$  s of arrhythmia, as per clinical guidelines.<sup>3</sup> Other sustained supraventricular arrhythmia, i.e. atrial flutter (atrial rhythm above 280 b.p.m.) and fast atrial tachycardia (rapid atrial rhythm  $> 200$  b.p.m.) were diagnosed based on ECG findings. Any uncertainty in diagnosis was reviewed by a second trial cardiologist, and a decision was made by consensus. Participants with uninterpretable 12-lead ECGs were invited to return to the research clinic for a second 12-lead ECG. All intervention arm participants with a cardiologist-confirmed diagnosis of AF,

atrial flutter, or fast atrial tachycardia (either from the 12-lead ECG or KardiaMobile data) were considered in the primary endpoint analyses. The NNS was calculated as  $1/(\text{diagnosis rate among those screened} - \text{diagnosis rate among those not screened})$ .

All AF and related arrhythmia diagnoses were reported to GPs for dedicated management and anticoagulation therapy as per routine clinical practice. Participants diagnosed with other cardiac abnormalities following ECG review were referred back to their GP and treated according to routine practice, if required. Intervention arm participants were also able to receive diagnoses of AF, atrial flutter, or fast atrial tachycardia through routine clinical practice. Control arm participants were only diagnosed with AF and related arrhythmias through routine clinical practice. Any diagnoses outside of the trial intervention during the study period were made according to local policies.

Nine months into the study, COVID-19 cases reached epidemic levels and a 'stay-at-home' message was disseminated by the UK government. Following reported associations between COVID-19 and AF,<sup>14</sup> a protocol amendment was submitted to include an exploratory objective to explore the impact of COVID-19 on: rates of COVID-19 diagnosis across trial arms; rates of AF and related arrhythmia diagnosis after COVID-19 diagnosis; and the impact of AF and related arrhythmia diagnosis after COVID-19 diagnosis on background AF diagnoses.

## AF risk-prediction algorithm

The AF risk-prediction algorithm has been described in detail elsewhere.<sup>10</sup> In short, the algorithm was developed using machine learning techniques and electronic medical record data from a retrospective cohort of almost 3 000 000 adult participants (aged  $\geq 30$  years) without a prior history of AF and listed on the CPRD GOLD between January 2006 and December 2016.<sup>10</sup> During the study period, 3.2% of the cohort were diagnosed with AF. Both baseline [patient demographics (age, sex, race, smoking status), history of anti-hypertensive use, Type 1 or 2 diabetes, and cardiovascular comorbidities] and time-varying (recent cardiovascular event(s), recent BMI and change in BMI, recent pulse pressure, change in SBP and DBP, and recent frequency of SBP, DBP, and BMI recordings) patient data were incorporated into the machine learning algorithm to generate a risk score for AF.<sup>10</sup> Participants with a risk-prediction score of  $\geq 7.4\%$ —the threshold determined during algorithm validation corresponding with 50% sensitivity and 90% specificity—were considered to be at high risk of undiagnosed AF.<sup>10, 11</sup> The AF risk-prediction algorithm (v 4.0) evaluated in this trial was unchanged from that validated in prior publications.<sup>10,11</sup>

## Statistical analyses

Sample size assumptions were based on data from the AF risk-prediction algorithm development.<sup>10</sup> Across six study sites, and  $\sim 24\,000$  participants, it was assumed that  $\sim 1000$  participants per arm would be at high risk of AF. We also estimated that  $\sim 30\%$  of high-risk participants in the intervention arm would accept the invitation to attend the research clinic, and of those,  $\sim 1$  in 16 would be diagnosed with AF (6.5%).<sup>13</sup> Thus, over a 6-month study period,  $\sim 2.4$  and 0.7% of high-risk participants in the intervention and control arms, respectively, would have been diagnosed with AF, yielding a statistical power of 88.5%. Power calculations were conducted using the 'pow' package in R (version 3.4.2).

Demographic and clinical characteristics of the high-risk and research clinic populations were summarized using descriptive statistics. Following adjustment for baseline characteristics (age, sex, and history of hypertension, heart failure, Type 1 or 2 diabetes mellitus), penalized logistic regression analyses were undertaken using the logistf function in R software to compare the proportion of diagnoses of AF and related

arrhythmias in participants at high risk of undiagnosed AF throughout the study in the intervention and control arms. Furthermore, a sensitivity analysis was undertaken on the proportions of AF diagnoses (not including related arrhythmias) across trial arms in the analysis populations but without prior exclusion of participants with atrial flutter or fast atrial tachycardia at baseline. All analyses were undertaken with R software (version 4.0.2), and all statistical tests were conducted at the 5% significance level. This study was registered with ClinicalTrials.gov (NCT04045639).

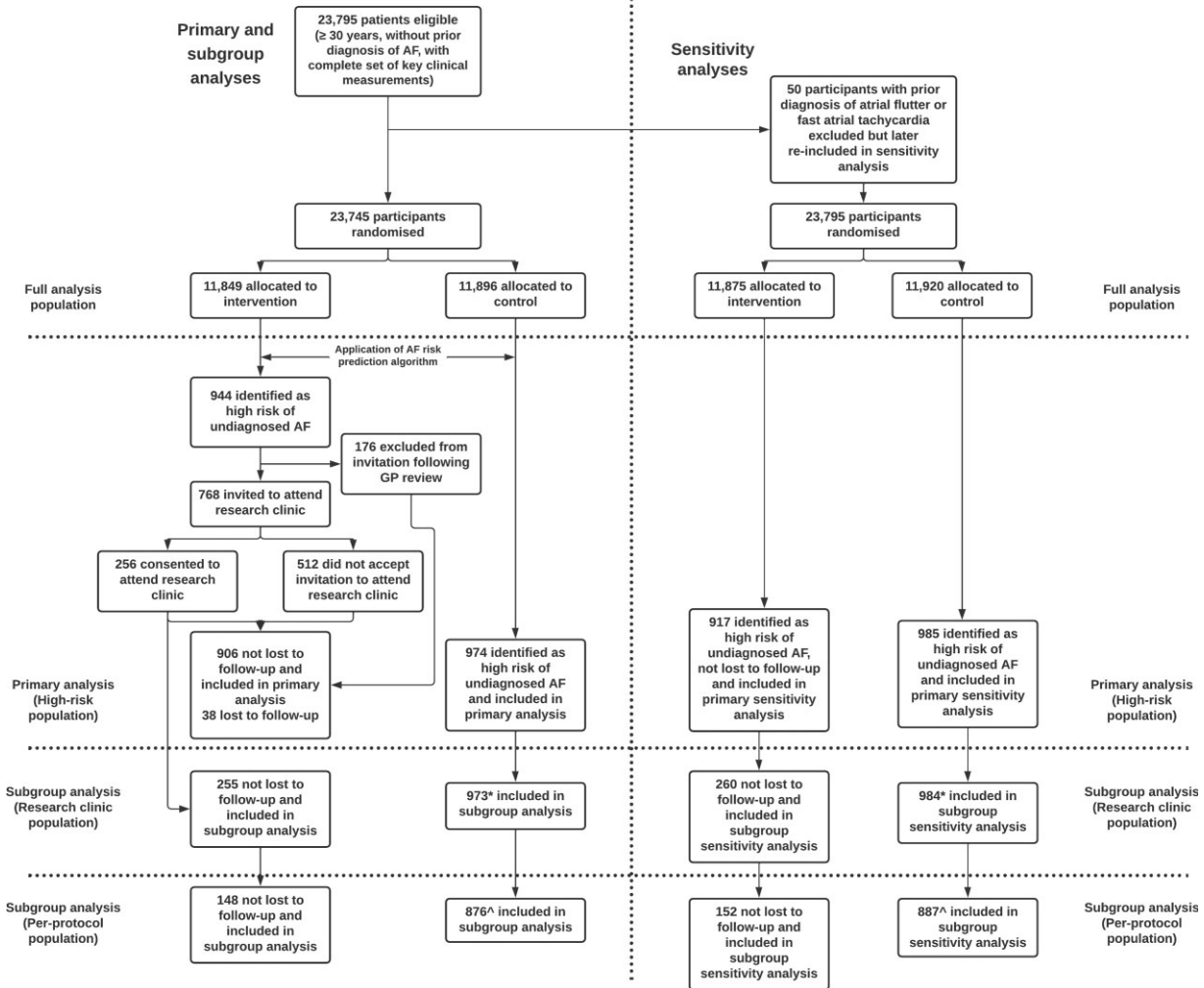
## Role of the funding source

The funder of the study was involved in the study design, data interpretation, and review of the report but had no role in data collection or data analysis. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. The funder had no influence on treatment decisions made by GPs following diagnosis of AF or related arrhythmia.

## Results

A total of 23 745 participants from six general practices were randomized into intervention and control arms on 6 June 2019 (the first participant attended the research clinic on 5 August 2019). Of the 11 849 participants allocated to the intervention arm, 944 were identified as high risk of undiagnosed AF by the algorithm. One hundred and seventy-six participants were excluded following GP review, and 768 participants were invited to attend the research clinic for diagnostic testing; 256 (33.3%) of whom consented to attend the research clinical for diagnostic testing (Figure 1). Of the 944 participants initially identified as high risk, 38 were lost to follow-up, leaving 906 included in the final analysis. Furthermore, of the 11 896 participants allocated to the control arm, 974 were identified as high risk of undiagnosed AF by the algorithm and included in the final analysis. The trial ran for 20 months, until 18 February 2021, but the intervention was paused from 16 March 2020 to 21 July 2020 due to the COVID-19 pandemic. During the pause, no trial interventions took place, but trial participants in both arms could still be diagnosed with AF and related arrhythmias via routine clinical care. During the trial, 1.3 and 0.9% of participants in the full analysis population in the intervention and control arms, respectively, were lost to follow-up.

High-risk participants in the intervention and control arms had mean [standard deviation (SD)] ages of 78.4 (8.9) and 78.7 (8.6) years, were more likely to be male [54.9% (497/906) and 54.6% (532/974)], and had mean (SD) BMI of 30.7 (7.2) and 30.4 (7.6) kg/m<sup>2</sup>, respectively (Table 1). The majority were hypertensive [77.9% (706/906) and 78.6% (766/974)] and many had histories of coronary heart disease [34.1% (309/906) and 32.3% (315/974)], Type 2 diabetes mellitus [21.3% (193/906) and 19.8% (193/974)], or prior myocardial infarction [9.4% (85/906) and 11.4% (111/974)] in intervention and control arms, respectively. Only 1.2% (11/906) and 2.0% (19/974) of participants in intervention and control arms, respectively, had primary care codes for prior echocardiography, ECG, or Holter monitoring. There was little difference in the baseline characteristics of high-risk participants in the intervention arm who accepted the invitation to attend the research clinic ( $n = 255$ ) compared with those who did not ( $n = 651$ ) (Supplementary material online, Table S3). Baseline characteristics of the full randomized



**Figure 1** Trial profile summarising number of participants eligible at each trial stage and included in each analysis population. \*One participant randomized to the control arm who attended the research clinic in error was removed from the subgroup analysis population. ^See [Supplementary material online, Table S1](#) for descriptions of analysis populations.

population, stratified by risk score are provided in [Supplementary material online, Table S2](#).

Overall, 51 of 906 (5.63%) and 48 of 974 (4.93%) of high-risk participants in the intervention and control arms, respectively, were diagnosed with AF, atrial flutter, or fast atrial tachycardia {odds ratio (OR) [95% confidence interval (CI)]: 1.15 (0.77–1.73),  $P = 0.486$ ; [Figure 2](#) and [Table 2](#)} during the trial. There was a statistically significant difference in the intervention arm participants who attended the research clinic (research clinic population;  $n = 255$ ; 28.1%) for diagnostic testing with, 24 of 255 (9.41%) diagnosed with AF and related arrhythmias compared with 4.93% of high-risk participants in the control arm [OR (95% CI): 2.24 (1.31–3.73),  $P = 0.003$ ]. The outcome was more favourable among participants who completed the intervention as per the protocol (12-lead ECG followed by KardiaMobile monitoring or up to two additional 12-lead ECGs) (per-protocol population;  $n = 148$ ; 58.0% of participants who attended the research clinic for diagnostic testing) with, 15 of 148

(10.14%) diagnosed with AF or related arrhythmias compared with 4.45% (39/876) high-risk participants in the control arm who were alive at the end of the trial [OR (95% CI): 3.07 (1.57–5.81),  $P = 0.001$ ]. Results from the sensitivity analysis [the proportions of AF diagnoses (not including related arrhythmias) across trial arms in the analysis populations but without prior exclusion of participants with atrial flutter or fast atrial tachycardia at baseline] are presented in [Supplementary material online, Table S4](#).

Across all diagnoses in the high-risk population during the trial period, 93.9% (93/99) were for AF and 7.0% (7/99) for atrial flutter and/or fast atrial tachycardia. Of the 24 diagnoses of AF and related arrhythmia made in the research clinic population, 13 diagnoses were a direct result of the trial intervention, and the remaining 11 diagnoses were made through routine care during the trial period. Six participants (46.2%) had arrhythmia present on the 12-lead ECG, and the remaining seven participants (53.8%) were diagnosed following KardiaMobile home-based ECG recording. In the research



**Table 1** Baseline characteristics of the high-risk population

	Intervention (n = 906)	Control (n = 974)	P-value <sup>a</sup>
Age, years	78.4 (8.9)	78.7 (8.6)	0.442
Sex			
Female	409 (45.1)	442 (45.4)	0.918
Male	497 (54.9)	532 (54.6)	0.918
Ethnicity			
White	194 (21.4)	189 (19.4)	0.362
Black, Asian, and minority ethnic	1 (0.1)	5 (0.5)	0.606
Unknown	711 (78.7)	585 (80.1)	0.391
Weight, kg	89.3 (22.8)	88.2 (22.4)	0.312
Height, cm	170.9 (11.5)	170.2 (11.0)	0.293
BMI, kg m <sup>2</sup>	30.7 (7.2)	30.4 (7.6)	0.449
Systolic blood pressure, mmHg	138.9 (18.6)	133.2 (19.0)	0.030
Diastolic blood pressure, mmHg	76.6 (10.2)	74.2 (10.4)	0.103
Hypertension	618 (68.2)	685 (70.3)	0.320
Smoking status			
Non-smoker	482 (53.2)	501 (51.4)	0.444
Current smoker	55 (6.1)	62 (6.4)	0.791
Former smoker	368 (40.6)	411 (42.2)	0.487
Passive smoker	1 (0.1)	0 (0.0)	0.300
Type 1 diabetes mellitus	31 (3.4)	29 (3.0)	0.584
Type 2 diabetes mellitus	193 (21.3)	193 (19.8)	0.425
Coronary heart disease	309 (34.1)	315 (32.3)	0.417
Prior myocardial infarction	85 (9.4)	111 (11.4)	0.153
Prior cardiac arrest	2 (0.2)	1 (0.1)	0.522
Heart failure	64 (7.1)	82 (8.4)	0.273
Congestive heart failure	49 (5.4)	64 (6.6)	0.289
Prior transient ischaemic attack	70 (7.7)	82 (8.4)	0.582
Left ventricular dysfunction	20 (2.2)	28 (2.9)	0.359
Left ventricular hypertrophy	13 (1.4)	15 (1.5)	0.851
Prior echocardiography, electrocardiogram, Holter monitoring	11 (1.2)	19 (2.0)	0.203

Age is as recorded at baseline, and data for continuous variables are based on the most recent record within a 5-year look back period; comorbidities have no time restrictions. For continuous variables, numbers represent mean (SD) and for categorical variables numbers represent *n* (%).

<sup>a</sup>For continuous variables, *P* value is based on a *t*-test for comparison of means; for categorical values, *P* value is based on a  $\chi^2$  two-sample test for equality of proportions (without continuity correction).

clinic population, i.e. participants at high risk according to the AF risk-prediction algorithm who consented to screening, the NNS for every diagnosis of AF or related arrhythmia was 12. Compliance with KardiaMobile ECG recording was high; only 7% of participants who received a device recorded <10 ECGs during the 2-week monitoring period. Furthermore, because the KardiaMobile technology

will only record an ECG once a stable signal has been detected, the number of unreadable ECGs was small.

Analysis of the temporal trends in the diagnosis of AF and related arrhythmia revealed a slightly steeper slope in the cumulative sum of diagnoses in the intervention arm during the first nine months of the trial (June 2019–March 2020) prior to the COVID-19 pandemic. This was followed by a relative plateau in diagnoses during the following 4–6 months—aligned with the national lockdown in the UK (Figure 3). Diagnoses in both arms increased again from September 2020 to the end of the trial during the relaxation of restrictions. Exploratory analyses are reported in the [Supplementary material](#) online.

No adverse events relating to the trial intervention were reported to investigators, the study sponsor, or the MHRA.

## Discussion

This is the first multi-centre randomized controlled trial to evaluate the performance of a machine learning-based AF risk-prediction algorithm in primary care. The AF risk-prediction algorithm was effective at identifying patients at high risk of undiagnosed AF in both arms of the trial, as evidenced by 5.3% of high-risk participants across intervention and control arms diagnosed with AF or related arrhythmia during the 20-month trial, compared with 0.6% of participants classed as low risk of undiagnosed AF. The proportions of AF and related arrhythmia diagnoses were not significantly different in the two arms of the high-risk populations. However, due to the COVID-19 pandemic, the trial data collection period was extended (from estimated 6 months to actual 20 months). This extension may have contributed to the higher-than-expected number of background AF diagnoses observed during the trial and may have reduced participants' willingness to participate in screening interventions, thereby impacting assumptions made to inform sample size and power calculations.

In the subgroup of participants in the intervention arm who attended the research clinic, the AF risk-prediction algorithm followed by screening had a significant impact on diagnoses and doubled the likelihood of identifying AF and related arrhythmia compared with routine care in ~30% of participants who accepted the screening invitation. Current systematic screening methods for AF require ~70 people aged 65 years or older to be screened to identify one person with AF.<sup>3</sup> Results from this study indicate that application of the algorithm in primary care can successfully narrow the population who should be invited for diagnostic testing such that only 12 patients need to be screened to detect one case of AF or related arrhythmia. The NNS of 12 observed in this trial is not dissimilar to the NNS of 9 reported in the retrospective algorithm validation studies<sup>10,11</sup> and indicates that more widespread adoption of the algorithm may significantly reduce the resource burden associated with current systematic approaches for AF diagnosis.

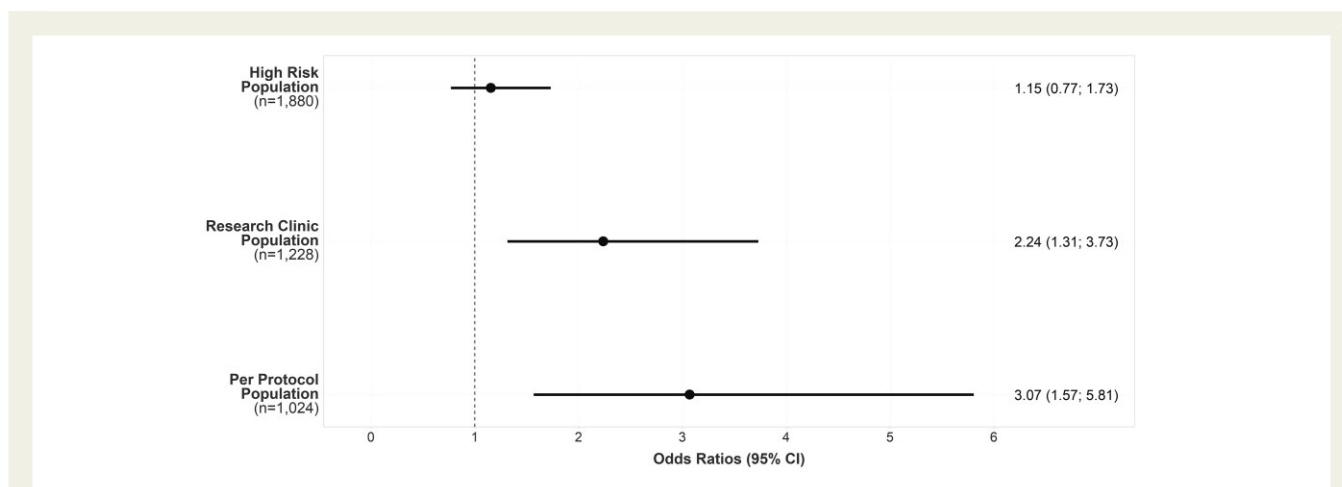
In participants who attended the research clinic and received a diagnosis of AF, atrial flutter, or fast atrial tachycardia, 56% had the arrhythmia present during the 12-lead ECG, indicating the disease presented intermittently in around half of the participants who were diagnosed via the research clinic. The finding highlights the value of screening beyond a single 12-lead ECG for the detection of AF.

The advent of new technologies such as the KardiaMobile portable monitor and wearable technologies such as smartwatches have increased the accessibility of home-based ECG monitoring and may play an important role in the future of cardiovascular disease management. Indeed, studies such as the Apple Heart Study<sup>15</sup> and Huawei Heart Study<sup>16</sup> have demonstrated promising results following evaluation of the feasibility of wearable technologies for AF screening. However, as evidenced by the observation that only 58% of participants who attended the research clinic completed the trial per protocol (i.e. took part in home-based ECG monitoring), there are significant barriers—such as cost and technological confidence—to more widespread adoption of these technologies for first-line screening, especially in older populations who are at higher risk of undiagnosed AF. Furthermore, the clinical relevance of very short episodes of AF detected with continuous or near-continuous monitoring is not yet understood.

Existing risk-prediction models for AF include CHARGE-AF,<sup>17</sup> ARIC,<sup>18</sup> Framingham AF,<sup>19</sup> SAAFE,<sup>20</sup> and C2HEST<sup>21</sup> models. ARIC and Framingham AF both require ECG data as an input, limiting their applicability in a primary care population, and none are routinely implemented in clinical practice because none are automated. Artificial intelligence techniques have been shown to be effective in the identification of undiagnosed AF.<sup>22,23</sup> The majority of these studies have

involved the application of machine learning algorithms to ECG traces to detect small changes in ECG activity that are associated with or may precede AF. Since the publication of our algorithm,<sup>10</sup> we are aware of three further studies that have reported on the development and validation of AF risk-prediction models based on machine learning techniques and utilizing data contained within electronic medical records.<sup>24–26</sup> However, all these models were developed based on more complex data sets sourced from secondary care records, reducing their applicability for use in a primary care setting. The novelty of this study lies in the prospective evaluation of the performance of an AF risk-prediction algorithm developed using machine learning techniques and inputs routinely collected in a real-world primary care setting.

Following the beginning of the COVID-19 pandemic, we sought to understand the impact of COVID-19 on AF and related arrhythmias within our trial. The pandemic itself did impact trial implementation, and in line with national guidance, we paused the trial intervention during the first lockdown in the UK. Only 11 participants received both AF or related arrhythmia and COVID-19 diagnoses and, therefore, findings of the exploratory objective should be interpreted with caution. While over twice as many participants in the intervention arm had both diagnoses compared with the control arm, there was no adjustment for COVID-19 in our analyses as the number



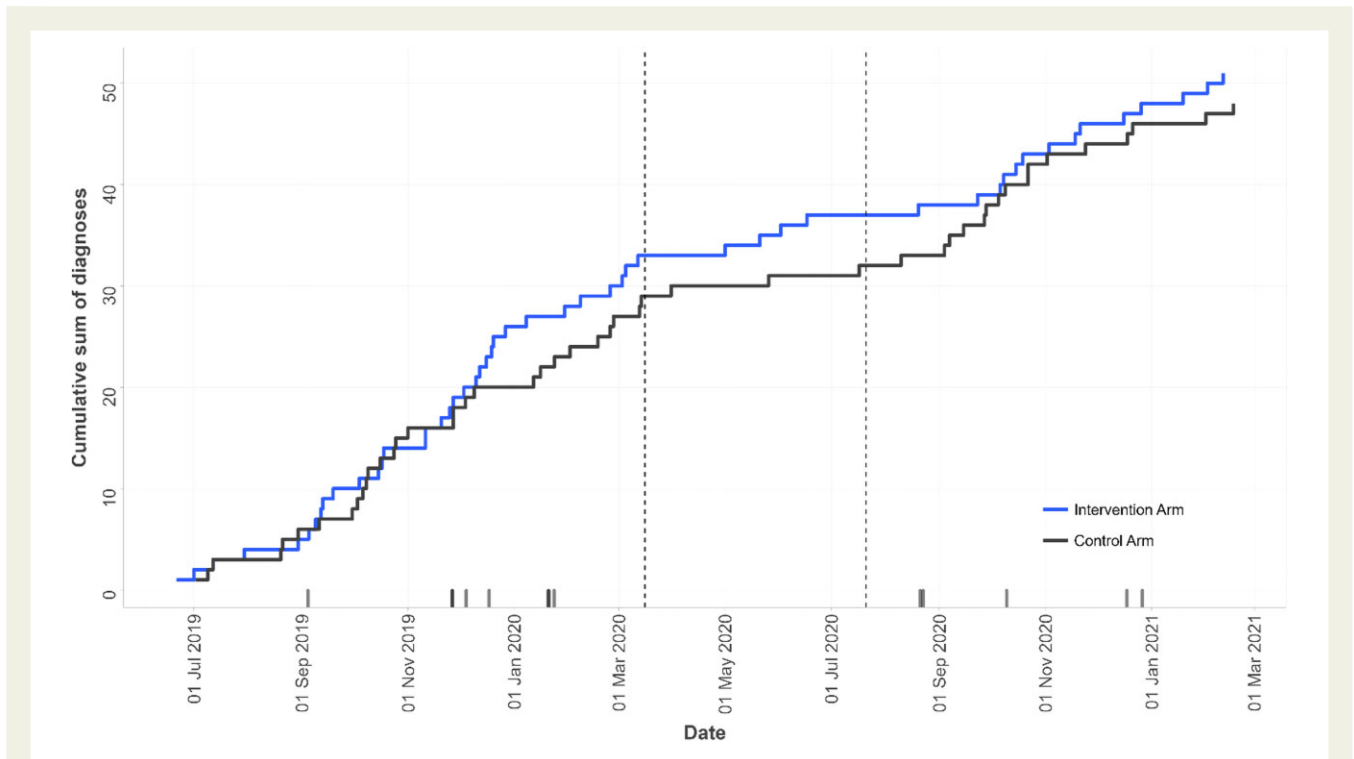
**Figure 2** Effect of intervention on diagnoses of atrial fibrillation, atrial flutter, and fast atrial tachycardia in the high-risk population, the research clinic population, and the per-protocol population.

**Table 2** Adjusted analysis for diagnoses of atrial fibrillation, atrial flutter, and fast atrial tachycardia in participants without atrial fibrillation, atrial flutter, or fast atrial tachycardia at baseline, stratified by analysis subgroup

Population	Intervention		Control		Difference			P-value
	N	Cases (%)	N	Cases (%)	OR	Lower CI	Upper CI	
Full analysis population	11 849	116 (0.98)	11 896	114 (0.96)	1.07	0.82	1.39	0.625
High-risk population	906	51 (5.63)	974	48 (4.93)	1.15	0.77	1.73	0.486
Research clinic population	255	24 <sup>a</sup> (9.41)	973	48 (4.93)	2.24	1.31	3.73	0.003
Per-protocol population	148	15 (10.14)	876	39 (4.45)	3.07	1.57	5.81	0.001

CI, 95% confidence interval; OR, odds ratio.

<sup>a</sup>Thirteen diagnoses were a direct result of the trial intervention, and the remaining 11 diagnoses were a result of routine care during the trial period.



**Figure 3** Temporal trends in diagnoses of atrial fibrillation (AF) and related arrhythmia during the trial period in intervention and control arms. Vertical dotted lines represent the pause in the trial intervention from 16 March 2020 to 21 July 2020 due to the COVID-19 pandemic. Marks above the x-axis represent diagnoses of AF or related arrhythmia in the research clinic population and reflect the date of diagnosis at the research clinic. Note: In one instance a participant was diagnosed through routine care prior to receiving a confirmation of diagnosis at the research clinic, therefore, not all marks correspond exactly with steps in the Kaplan Meier plot.

of participants affected was so low. However, the impact of COVID-19 on AF prevalence, diagnosis, and management are important topics for further investigation in other studies.

The strengths of this trial include its multi-centre design and completion during the COVID-19 pandemic despite significant logistical challenges. There are some limitations: first, the response rate was relatively poor at ~30%. Although we were unable to collect data on reasons for individuals not accepting the invitation to participate, the smartphone component of the trial may have been a barrier for some. Nevertheless, it is accepted that recruitment to trials in primary care is often challenging,<sup>27</sup> and the pandemic likely added additional barriers to the high-risk cohort in the intervention arm invited to attend the research clinic. Second, not all assumptions made during the sample size and statistical power calculations were correct. Our background diagnosis rate assumptions were based on those observed in the CPRD data used to develop the AF risk-prediction algorithm; however, the rates observed in the current trial—although not significantly dissimilar to other more recent studies<sup>8,28,29</sup>—were higher than predicted. This was, at least in part, due to the pausing, and therefore, extension of the trial due to the pandemic increasing the diagnostic rates in the high-risk cohorts. It is also possible that measures to increase the detection of AF<sup>30,31</sup> in recent years may have had an impact, or a result of performance bias due to trial participation. Third, a significant proportion of high-risk participants (18.6%) were excluded from the invitation to the trial following the GP decision. A number of these exclusions were

due to participants being housebound or residents in a care facility and therefore unable to attend the research clinic. Exclusion of these participants likely resulted in selection bias of a high-risk cohort at lower risk of AF; however, due to the pragmatic nature of the trial, this selection bias was unavoidable. Fourth, the majority of participants deemed at high risk of undiagnosed AF were older and many did not have access to smartphones compatible with the KardiaMobile device. We sought to address this through the provision of loan smartphones for participants; however, many were unfamiliar with how to use the technology and therefore did not complete the full 2 weeks of home-based ECG monitoring. Fifth, although the number of participants lost to follow-up during the trial was relatively small, we were unable to understand the reason(s) for this because data for these participants were absent in medical records at the end of the trial. We suspect this may have been due to participants leaving the practice but are unfortunately unable to provide this detail. Lastly, the generalizability of findings to patients and practices in the UK beyond those involved in the trial and the surrounding areas is uncertain.

An important avenue for future research is to determine the health economic impact of the intervention in a real-world setting and the costs associated with the management of patients who are diagnosed with AF and related arrhythmia compared with those who remain undiagnosed. If the application of the algorithm in combination with diagnostic testing is found to be cost-effective, there may be value in its wider implementation across primary care in



the UK. The AF risk-prediction algorithm, developed by Hill *et al.*,<sup>10</sup> is unique in that it does not require ECG data as an input. Instead, risk-prediction scores are generated based on routinely collected clinical data, such as age, sex, BMI, SBP, DBP, comorbidities, and cardiovascular events. This feature of the algorithm enables the estimation of AF risk without prior ECG assessment, and as evidenced by the very small proportion of high-risk participants with prior primary care codes for echocardiograms, ECG, and Holter monitoring, this is important for widespread employment. Furthermore, given the ongoing impact of COVID-19 on in-person primary and secondary care consultations reducing the chance for opportunistic AF screening, there is additional value in a method that could be applied across medical records at the practice-level for GPs to identify patients at the highest risk of undiagnosed AF who should undergo further assessment via ECG.

## Conclusion

The AF risk-prediction algorithm was effective in identifying participants at high risk of undiagnosed AF. Although our trial did not demonstrate significant differences in the total number of AF and related arrhythmia diagnoses across all high-risk patients, in the subgroup of participants who attended the research clinic, the intervention was superior to routine clinical care for the detection of undiagnosed AF. The AF risk-prediction algorithm may be an effective tool in narrowing the population at high risk of undiagnosed AF who should undergo diagnostic testing.

## Supplementary material

[Supplementary material](#) is available at *European Heart Journal – Digital Health*.

## Acknowledgements

The authors wish to thank the following principal investigators and research nurses for their significant contributions to data collection: Dr Andrew Wallis and Terri-Anne Thomson (Alcester Health Centre); Dr Manjit Kainth and Jacqueline Woodward (Primrose Lane Surgery); Dr Paul Ainsworth and Oliva Neely (Sherbourne Medical Centre); Dr Caron Morton (Station Drive Surgery); Dr Claire Jones (Spring Gardens Group Medical Practice); Dr Stefan Lachowicz and Deborah Williams (The Caxton Surgery); and Sian Jones and Jonathan Davies (National Institute for Health Research). The authors also wish to thank Nia Jenkins of Health Economics and Outcomes Research Ltd for her contributions to data analysis and Michelle Robinson and Simon Wathall of the School of Medicine, Keele University, UK, for their logistical support with the study. Nia Jenkins, Michelle Robinson, and Simon Wathall were paid consultants to Bristol Myers Squibb Pharmaceutical Ltd and Pfizer in connection with the development of this manuscript.

## Funding

This work was supported by Bristol Myers Squibb Pharmaceuticals Ltd and Pfizer. D.C. is funded by the NIHR Biomedical Research Centre, Oxford.

**Conflict of interest:** A.J.C. has received institutional grants and personal fees from Bayer, Boehringer Ingelheim, Bristol Myers Squibb/Pfizer Alliance, and Daiichi Sankyo. A.-C.M. has received consulting fees and a research grant from Bristol Myers Squibb/Pfizer Alliance and consulting fees from Bayer-Healthcare and Boehringer Ingelheim. D.C. declares fees from Oxford University Innovation, Biobeats, Bristol Myers Squibb, Sensyne Health, and academic research grants from GlaxoSmithKline, AstraZeneca, Apple, RCUK, Wellcome Trust, EU Horizon2020, and NIHR. A.T.C. receives consulting fees from AbbVie, ACI Clinical, Alexion, Aspen, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, Boston Scientific, CSL Behring, Daiichi Sankyo, GlaxoSmithKline, GLG, Guidepoint Global, Johnson and Johnson, Leo Pharma, Medscape, McKinsey, Navigant, ONO, Pfizer, Portola, Sanofi, Temasek Capital, TRN; advisory board membership with Bayer, Bristol Myers Squibb, Daiichi Sankyo, Johnson and Johnson, ONO, Pfizer, Portola, Sanofi; payments for lectures including speakers bureau services, payments for preparation of reports and payment for development of educational presentations from, Aspen, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, Daiichi, GlaxoSmithKline, Johnson and Johnson, Medscape, Pfizer, and Portola. He has advised the UK Government Health Select Committee, the all-party working group on thrombosis, the Department of Health, and the NHS, on the prevention of VTE. He is also an advisor to Lifeblood: the thrombosis charity and is the founder of the European educational charity the Coalition to Prevent Venous Thromboembolism. C.A. and D.W.D. have received consulting fees from Bristol Myers Squibb Pharmaceutical Ltd and Pfizer in relation to this study. N.R.H., U.F., M.H., K.G.P., and B.S. are employees of Bristol Myers Squibb UK. M.H. was previously employed by Health Economics and Outcomes Research Ltd (HEOR). J.G., L.G., C.D., A.O., D.P., D.M.S., and C.T. are or were employees of HEOR. HEOR was a paid consultant to Bristol Myers Squibb Pharmaceutical Ltd and Pfizer in connection with the development of this manuscript. J.R. is an employee of PHASTAR. PHASTAR was a paid consultant to Bristol Myers Squibb Pharmaceutical Ltd & Pfizer in connection with the development of this manuscript. C.M. and S.L. are employees of Keele Clinical Trials Unit. Keele Clinical Trials Unit was a paid consultant to Bristol Myers Squibb Pharmaceutical Ltd & Pfizer in connection with the development of this manuscript. C.M. is also funded by the National Institute for Health Research (NIHR) Applied Research Collaboration West Midlands and the NIHR School for Primary Care Research.

## Data availability

Deidentified individual participant data that underlie the results reported in this article, and the study protocol and statistical analysis plan will be made available upon reasonable request to the corresponding author.

## References

1. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;**22**:983–988.
2. Pistoia F, Sacco S, Tiseo C, Degan D, Ornello R, Carolei A. The epidemiology of atrial fibrillation and stroke. *Cardiol Clin* 2016;**34**:255–268.
3. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, Boriani G, Castella M, Dan G-A, Dilaveris PE, Fauchier L, Filippatos G, Kalman JM, La Meir M, Lane DA, Lebeau J-P, Lettino M, Lip GYH, Pinto FJ, Thomas GN, Valgimigli M, Van Gelder IC, Van Putte BP, Watkins CL, ESC Scientific Document Group. 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**:373–498.
4. British Heart Foundation. Atrial fibrillation: finding the missing 300,000; 2019. <https://www.bhf.org.uk/for-professionals/healthcare-professionals/blog/2019/atrial-fibrillation-finding-the-missing-300000> (accessed 1 October 2020).
  5. Turakhia MP, Shafrin J, Bognar K, Trocio J, Abdulsattar Y, Wiederkehr D, Goldman DP. Estimated prevalence of undiagnosed atrial fibrillation in the United States. *PLoS One* 2018;**13**:e0195088.
  6. Engler D, Heidebuchel H, Schnabel RB. Digital, risk-based screening for atrial fibrillation in the European community—the AFFECT-EU project funded by the European Union. *Eur Heart J* 2021;**42**:2625–2627.
  7. Sentinel Stroke National Audit Programme (SSNAP). *Clinical Audit National Results, 2019–20*. <https://www.strokeaudit.org/results/Clinical-audit/National-Results.aspx> (accessed 12 July 2021).
  8. Holt A, Gislason GH, Schou M, Zareini B, Biering-Sørensen T, Phelps M, Kragholm K, Andersson C, Fosbøl EL, Hansen ML, Gerds TA, Køber L, Torp-Pedersen C, Lamberts M. New-onset atrial fibrillation: incidence, characteristics, and related events following a national COVID-19 lockdown of 5.6 million people. *Eur Heart J* 2020;**41**:3072–3079.
  9. Nicholls M. Machine Learning-state of the art. *Eur Heart J* 2019;**40**:3668–3669.
  10. Hill NR, Ayoubkhani D, McEwan P, Sugrue DM, Farooqui U, Lister S, Lumley M, Bakhai A, Cohen AT, O'Neill M, Clifton D. Predicting atrial fibrillation in primary care using machine learning. *PLoS One* 2019;**14**:e0224582.
  11. Sekelj S, Sandler B, Johnston E, et al. Detecting undiagnosed atrial fibrillation in UK primary care: validation of a machine learning prediction algorithm in a retrospective cohort study. *Eur J Prev Cardiol* 2020;**28**:598–605.
  12. General Data Protection Regulation (GDPR). 2018. <https://gdpr.eu/tag/gdpr/> (accessed 6 June 2019).
  13. Hill NR, Arden C, Beresford-Hulme L, Camm AJ, Clifton D, Davies DW, Farooqui U, Gordon J, Groves L, Hurst M, Lawton S, Lister S, Mallen C, Martin A-C, McEwan P, Pollock KG, Rogers J, Sandler B, Sugrue DM, Cohen AT. Identification of undiagnosed atrial fibrillation patients using a machine learning risk prediction algorithm and diagnostic testing (PULSe-A): study protocol for a randomised controlled trial. *Contemp Clin Trials* 2020;**99**:106191.
  14. Stone E, Kiat H, McLachlan CS. Atrial fibrillation in COVID-19: a review of possible mechanisms. *FASEB J* 2020;**34**:11347–11354.
  15. Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, Balasubramanian V, Russo AM, Rajmane A, Cheung L, Hung G, Lee J, Kowey P, Talati N, Nag D, Gummidipundi SE, Beatty A, Hills MT, Desai S, Granger CB, Desai M, Turakhia MP. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med* 2019;**381**:1909–1917.
  16. Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y, Yan L, Xing Y, Shi H, Li S, Liu Y, Liu F, Feng M, Chen Y, Lip GYH, Guo Y, Lip GYH, Lane DA, Chen Y, Wang L, Eckstein J, Thomas GN, Tong L, Mei F, Xuejun L, Xiaoming L, Zhaoliang S, Xiangming S, Wei Z, Yunli X, Jing W, Fan W, Sitong Y, Xiaoqing J, Bo Y, Xiaojuan B, Yuting J, Yangxia L, Yingying S, Zhongju T, Li Y, Tianzhu L, Chunfeng N, Lili Z, Shuyan L, Zulu W, Bing X, Liming L, Yuanzhe J, Yunlong X, Xiaohong C, Fang W, Lina Z, yihong S, shujie J, Jing L, Nan L, shijun L, huixia L, Rong L, Fan L, qingfeng G, tianyun G, Yuan W, Xin L, Yan R, xiaoping C, ronghua C, Yun S, yulan Z, haili S, yujie Z, quanchun W, weidong S, Lin W, Chan E, Guangliang S, Chen Y, Wei Z, Dandi C, Xiang H, Anding X, Xiaohan F, Ziqiang Y, Xiang G, Fulin G. Mobile photoplethysmographic technology to detect atrial fibrillation. *J Am Coll Cardiol* 2019;**74**:2365–2375.
  17. Alonso A, Krijthe BP, Aspelund T, Stepos KA, Pencina MJ, Moser CB, Sinner MF, Sotoodehnia N, Fontes JD, Janssens ACJW, Kronmal RA, Magnani JW, Witteman JC, Chamberlain AM, Lubitz SA, Schnabel RB, Agarwal SK, McManus DD, Ellinor PT, Larson MG, Burke GL, Launer LJ, Hofman A, Levy D, Gottdiener JS, Käb S, Couper D, Harris TB, Soliman EZ, Stricker BHC, Gudnason V, Heckbert SR, Benjamin EJ. 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.
  18. Chamberlain AM, Agarwal SK, Folsom AR, Soliman EZ, Chambless LE, Crow R, Ambrose M, Alonso A. 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.
  19. Schnabel RB, Sullivan LM, Levy D, Pencina MJ, Massaro JM, D'Agostino RB, Newton-Cheh C, Yamamoto JF, Magnani JW, Tadros TM, Kannel WB, Wang TJ, Ellinor PT, Wolf PA, Vasari RS, Benjamin EJ. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. *Lancet* 2009;**373**:739–745.
  20. Linker DT, Murphy TB, Mokdad AH. Selective screening for atrial fibrillation using multivariable risk models. *Heart* 2018;**104**:1492–1499.
  21. Li YG, Pastori D, Farcomeni A, Yang P-S, Jang E, Joung B, Wang Y-T, Guo Y-T, Lip GYH. 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**:510–518.
  22. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, Carter RE, Yao X, Rabinstein AA, Erickson BJ, Kapa S, Friedman PA. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet* 2019;**394**:861–867.
  23. Oster J, Hopewell JC, Zibera K, Wijesurendra R, Camm CF, Casadei B, Tarassenko L. Identification of patients with atrial fibrillation: a big data exploratory analysis of the UK Biobank. *Physiol Meas* 2020;**41**:025001.
  24. Tiwari P, Colborn KL, Smith DE, Xing F, Ghosh D, Rosenberg MA. Assessment of a machine learning model applied to harmonized electronic health record data for the prediction of incident atrial fibrillation. *JAMA Netw Open* 2020;**3**:e1919396.
  25. Suzuki R, Katada J, Ramagopalan S, McDonald L. Potential of machine learning methods to identify patients with nonvalvular atrial fibrillation. *Future Cardiol* 2020;**16**:43–52.
  26. Grout RW, Hui SL, Imler TD, El-Azab S, Baker J, Sands GH, Ateya M, Pike F. Development, validation, and proof-of-concept implementation of a two-year risk prediction model for undiagnosed atrial fibrillation using common electronic health data (UNAFIED). *BMC Med Inform Decis Mak* 2021;**21**:112.
  27. White D, Hind D. Projection of participant recruitment to primary care research: a qualitative study. *Trials* 2015;**16**:473.
  28. Mendonça SC, Saunders CL, Lund J, Mant J, Edwards D. Temporal trends in incidence of atrial fibrillation in primary care records: a population-based cohort study. *BMJ Open* 2020;**10**:e042518.
  29. Williams BA, Chamberlain AM, Blankenship JC, Hylek EM, Voyce S. Trends in atrial fibrillation incidence rates within an integrated health care delivery system, 2006 to 2018. *JAMA Netw Open* 2020;**3**:e2014874.
  30. Atrial Fibrillation (AF) Toolkit. Detect, protect and perfect. <https://www.stroke.org.uk/sites/default/files/detect-protect-perfect-london-af-toolkit-062017.pdf>
  31. National Health Service. The NHS Long Term Plan; 2019. <https://www.longtermplan.nhs.uk/>