BMJ Open

Triple Cardiovascular Disease Detection with an Artificial Intelligence-enabled Stethoscope (TRICORDER): design and rationale for a decentralised, real-world cluster randomised controlled trial and implementation study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-098030
Article Type:	Protocol
Date Submitted by the Author:	15-Dec-2024
Complete List of Authors:	Kelshiker, Mihir A; Imperial College London, National Heart and Lung Institute Bächtiger, Patrik; Imperial College London; Imperial College Healthcare NHS Trust Mansell, Josephine; Imperial College London; Imperial College Healthcare NHS Trust Kramer, Daniel B; Imperial College London National Heart and Lung Institute Petri, Camille F; Imperial College London, Almonte, Melanie T; Imperial College London, ; Imperial College Healthcare NHS Trust, Alrumayh, Abdullah; Imperial College London Peters, Alexei; Imperial College London, Costelloe, Ceire; Imperial College London, Global Digital Health Unit, Department of Primary Care and Public Health Falaschetti, Emanuela; University College London, Epidemiology and public health Barton, Carys; Imperial College Healthcare NHS Trust Al-Lamee, Rasha; Imperial College London, National Heart and Lung Institute Majeed, Azeem; Imperial College London, Primary Care Plymen, Carla M; Imperial College London; Imperial College Healthcare NHS Trust Peters, Nicholas S; St Marys Hospital and Imperial College,
Keywords:	Artificial Intelligence, CARDIOLOGY, HEALTH ECONOMICS, PUBLIC HEALTH, Heart failure < CARDIOLOGY, Primary Health Care

Triple Cardiovascular Disease Detection with an Artificial Intelligence-enabled Stethoscope (TRICORDER): design and rationale for a decentralised, real-world cluster randomised controlled trial and implementation study

Mihir A Kelshiker*[1,2], Patrik Bächtiger*[1,2], Josephine Mansell[1,2], Daniel B Kramer[1], Camille F Petri[1], Melanie T Almonte[1,2], Abdullah Alrumayh[1], Alexei Peters[1], Céire Costelloe[4], Emanuela Falaschetti[5], Carys Barton, Rasha Al-Lamee[1,2], Azeem Majeed[4], Carla M Plymen[1,2], Nicholas S Peters[1,2]**

1. National Heart and Lung Institute, Imperial College London, UK
2. Imperial College Healthcare NHS Trust, UK
3. School of Public Health, Imperial College London, UK
4. Imperial Clinical Trials Unit, School of Public Health, Imperial College London, UK

*contributed equally, **corresponding author.

Corresponding author:

Professor Nicholas S Peters

National Heart and Lung Institute, Imperial College London

Hammersmith Hospital

Du Cane Road

London, W12 0HS

n.peters@imperial.ac.uk

Funding: National Institute for Health Research Invention for Innovation (i4i) Challenge Award, Accelerated Access Collaborative and NHSX via the AI Award in Health and Social Care; British Heart Foundation; Imperial Health Charity.

Keywords: Cardiology/Artificial Intelligence/Digital Health/Primary Care/Public Health

ABSTRACT

Introduction

Early detection of cardiovascular disease in primary care is a public health priority, for which the clinical- and cost-effectiveness of an artificial intelligence-enabled stethoscope that detects left ventricular systolic dysfunction, atrial fibrillation, and cardiac murmurs is unproven but potentially transformative.

Methods and analysis

TRICORDER is a pragmatic, two-arm, multi-centre (decentralised), cluster randomised controlled trial and implementation study. Up to 200 primary care practices in urban North West London and rural North Wales, UK, will be randomised to usual care or to have artificial intelligence-enabled stethoscopes available for use. Primary care clinicians will use the artificial intelligence-enabled stethoscopes at their own discretion, without patient-level inclusion or exclusion criteria. They will be supported to do so by a clinical guideline developed and approved by the regional health system executive board. Patient and outcome data will be captured from pooled primary and secondary care records, supplemented by qualitative and quantitative clinician surveys. The co-primary endpoints are i. difference in coded incidence (detection) of heart failure and ii. difference in ratio of coded incidence of heart failure via hospital admission versus community-based diagnostic pathways. Secondary endpoints include difference in incidence of atrial fibrillation and valvular heart disease, cost-consequence differential, and prescription of guideline-directed medical therapy.

Ethics and dissemination

This trial has ethical approval from the UK Health Research Authority (23/LO/0051). Findings from this trial will be disseminated through publication of peer-reviewed manuscripts, presentations at scientific meetings and conferences with local and national stakeholders.

Trial registration number: NCT05987670

Summary

TRICORDER will examine the clinical- and cost-effectiveness of an artificial intelligence-enabled stethoscope for detection of priority cardiovascular diseases in real-world primary care. Findings will underpin implementation strategies for at-scale deployment and evaluation of artificial intelligence-enabled tools in routine clinical practice.

Strengths and limitations of this study

- This study is a large cluster randomised controlled trial in primary care, of the impact of an artificial intelligence (AI) -enabled, point-of-care diagnostic for three highly prevalent cardiovascular diseases. It will provide clinicians and policymakers with important clinical and health economic information to underpin the case for implementation across the NHS.
- By avoiding the need for patient-level research consent, workflows and data are "real-world", minimising impact on frontline primary care services and enhances inclusion of underserved groups and the generalisability of results.
- The parallel implementation study design will measure and address the social and behavioural barriers to uptake and sustained use of the technology, without impacting on its safety or statistical performance.
- Integration of AI predictions with routinely collected electronic medical record data allows sensitivity analyses at patient, site and regional level.
- Outcome assessment based on routine medical coding lacks the granularity and fidelity of primary research form collection.

INTRODUCTION

Early detection of cardiovascular disease is a global public health priority[1-3]. Heart failure (HF), atrial fibrillation (AF) and valvular heart disease (VHD) can have common pathophysiology and significant mortality and morbidity[4-7]. Given the reported adult prevalence of HF of 2%, the recognised incidence of HF in Europe of 5 per 1000 person-years in adults[8,9]is likely to be an underestimate[10-13]. HF is deadlier than common, serious cancers[14], profoundly impacting quality of life and costing the UK National Health Service (NHS) over £2 billion per year – or over 2% of its annual budget[15,16]. Unacceptably, 70-80% of all new HF diagnoses are only made after an emergency hospital admission, despite most patients having been seen in primary care with opportunities for detection and therefore earlier initiation of disease-modifying treatment[17,18] ,expected to translate to improved survival, quality of life and health service utilisation costs[18].

Notwithstanding the mixed evidence of effectiveness of screening programmes to improve mortality and stroke associated with AF, the substantial clinical and health economic costs of AF and VHD are similarly driven by late diagnosis[19,20].

Across these cardiac conditions, early detection reduces the risk of disease progression and improves survival, through initiation of inexpensive, guideline-directed therapies[21-23]. However, primary care providers face the challenge that symptoms (if present) are frequently non-specific[24,25], and there are no universally effective point-of-care screening tools to identify the diagnosis or diagnostic pathways.

We have recently shown that artificial intelligence (AI) applied to electrocardiogram (ECG) and phonocardiogram (PCG) waveforms captured by a "smart" stethoscope can detect left ventricular ejection fraction ≤40%[26], the cause of heart failure with reduced ejection fraction (HFrEF, the most common form of HF); AF, and cardiac murmurs – a cardinal examination finding in valvular heart disease[27]. The statistical performance of these three AI algorithms has been shown to be high and consistent against international external validation studies[28]. This takes a familiar clinical tool with an established workflow and augments examination findings with actionable additional insights, taking only 15 seconds[26]. The clinical and health economic impact of deploying such a triple point-of-care AI diagnostic in primary care practices is unknown. Furthermore, the determinants of successfully implementing[29-31] such an innovation in routine clinical practice are also not known, given the historical failure of AI technology adoption in healthcare systems that is partly attributable to inattention to implementation science, and a failure to overcome bias[32,33].

The objective of the TRICORDER study is therefore to conduct a real-world, decentralised cluster randomised controlled trial and implementation study to determine if providing primary care teams with an AI-stethoscope improves early diagnosis of HFrEF, AF and VHD. Secondary objectives include measurement of cost effectiveness, examination of implementation strategies and identifying the determinants of uptake and utilisation of the AI-stethoscope in routine clinical practice. This report describes the clinical and policy context for this publicly funded clinical trial, details of the protocol and plans for implementation, and expected findings with a focus on implications for widespread adoption in health systems.

METHODS AND ANALYSIS

Setting

TRICORDER is set within the primary care system of the NHS. General Practitioners (GPs) within primary care teams are the gatekeepers to most specialist diagnostic and management services in the NHS through nationally agreed pathways recommended by the National Institute for Health and Care Excellence (NICE), Figure 1. Each primary care practice (>8,000 in UK) serves an average list of 8,500 patients in the community[34]. Typically, patients in whom the GP suspects HF are referred for natriuretic peptide (NP) testing[35], which may be via in-practice phlebotomy or via hospital appointment. The NP level determines the urgency of referral for echocardiography and specialist review. There is substantial geographical variation in access to community-based echocardiography services, and lead times to hospital outpatient clinic review[36].

Study dates – October 2023 to October 2025

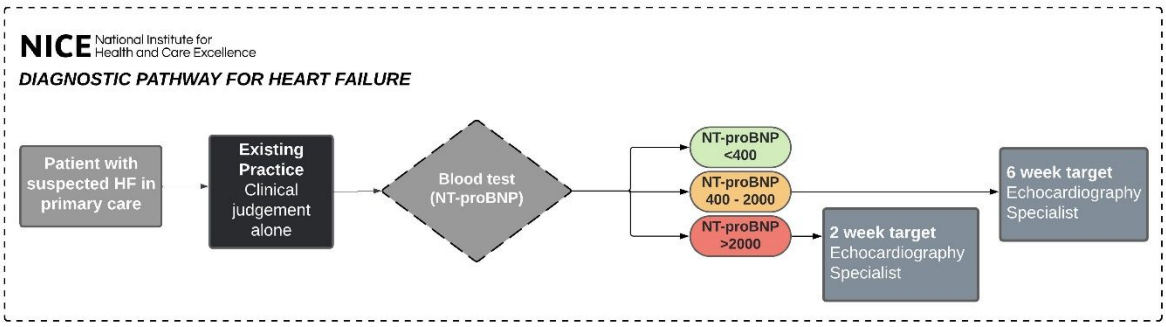


Figure 1. NICE diagnostic pathway for HF. *There are substantial delays to confirmatory echocardiographic testing and specialist review for those are referred for suspected HF through this pathway, such that only 10% of patients complete the pathway to time and target##.*

Funding

TRICORDER is funded by the UK National Institute for Health and Care Research (NIHR) Invention for Innovation Challenge Award. This is a competitively awarded public funding source, following a multi-stage process involving internal, external clinical and academic peer review, in addition to patient and public scrutiny. Eko Health are not involved in the design, conduct, analysis, interpretation of data, or reporting of TRICORDER.

Study design

TRICORDER is an open-label, two-arm cluster randomised controlled trial. Taking a decentralised approach, primary care practices in the NHS North West London Integrated Care System (NWL ICS, 2.8 million residents, Figure 2) and the Betsi Cadwaladr University Health Board (BCUHB, 670,000 residents) in North Wales, UK, will be enrolled for 18 months.

Each individual primary care practice employing GPs serves as a cluster – the unit of participation and randomisation. Each participating cluster will variously employ GPs, Practice Nurses, Physician Associates and other healthcare professionals who constitute the care team. This unit of recruitment was chosen because the NHS clinical, financial and information governance processes are standardised at this level.

Sample size

Interrogation of the Discover dataset[37] (UK Secure Data Environment for TRICORDER endpoint analysis) suggests an incidence rate of HF of 0.62/1000 patients per year. Combining NWL (n=350) and North Wales (n=>50), there are at least 400 potentially eligible study sites. We estimate that each GP practice will have at least 60 eligible patients per year who will receive a stethoscope examination as part of their routine clinical care. Estimating a conservative recruitment ratio of only 50%, we anticipate

200 GP practices being recruited. This provides an eligible study population of 6,000 patients per study arm. A review of prior literature in the field indicates an intracluster correlation coefficient of 0.01[38].

Utilising these estimates and assumptions, we would have over 80% power to detect a statistically and clinically significant mean difference if HF is diagnosed at a rate of 0.79/1000 patients in the intervention group, assuming a two-tailed test with a type I error of 0.05. This corresponds to a relative increase of 22%, which is deemed to be clinically meaningful (~17 to 22 HF diagnoses/GP practice).

The power calculation refers to overall HF, which includes all relevant ICD-10 and SNOMED-CT codes (used in UK primary care), rather than the codes relating to HFrEF specifically. There are three reasons for such an inclusive approach.

Firstly, we aim to mitigate an artificially low incidence of HFrEF due to coding heterogeneity in primary care. Preliminary analysis of the Discover dataset revealed substantial heterogeneity in clinical coding of HF subtypes. This is well recognised in recent studies[39].

Secondly, this approach captures the effect on clinician behaviour from participation in the study – it is feasible that more of all HF subtypes may be detected during the study period.

Finally, this approach enables rapid clinical and health economic endpoint analysis, avoiding the time and cost burden of manual data entry/data collection from chart arbitration, which would otherwise risk undermining the scalability of TRICORDER and set an impractical precedent for longer-term outcomes analyses.

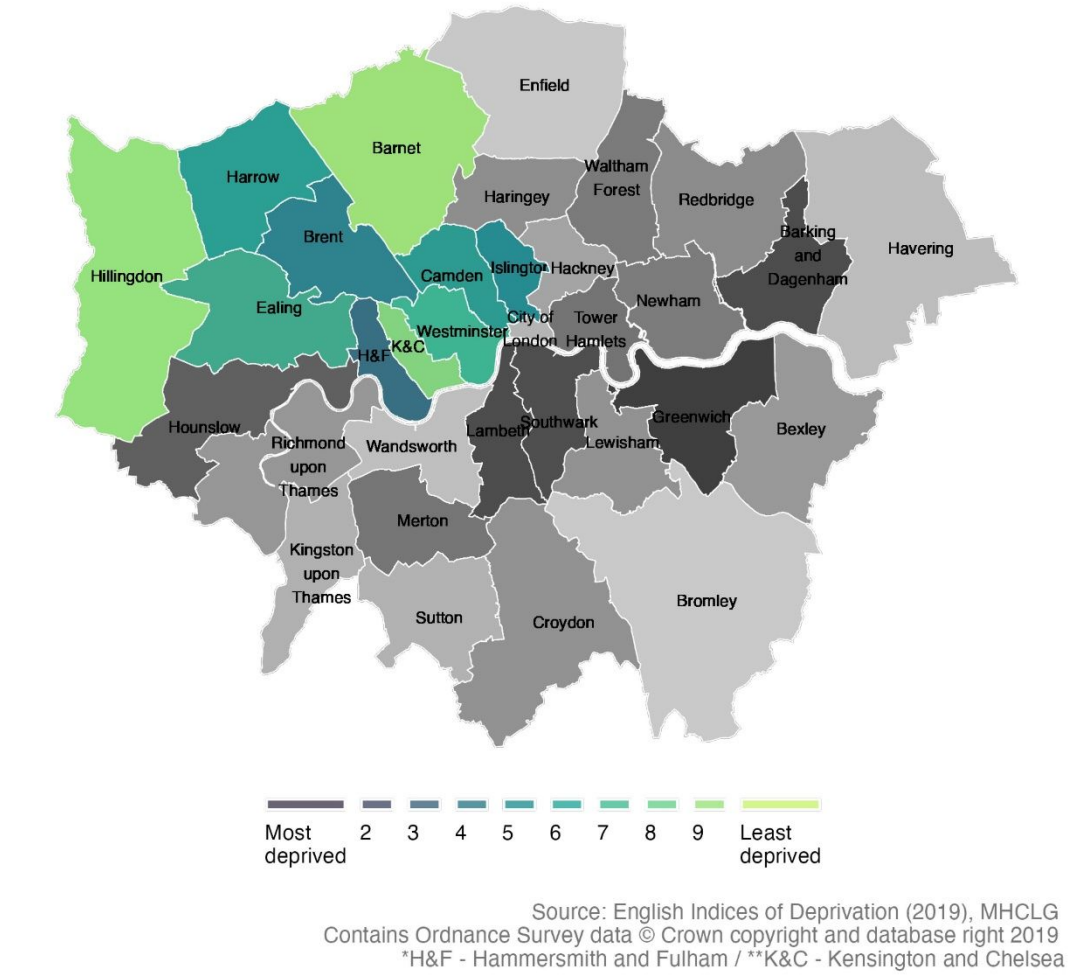


Figure 2: *Map of London, with the boroughs of North West London stratified by Index of Multiple Deprivation deciles*

Inclusion and exclusion criteria

Eligible primary care practices must meet the following inclusion criteria:

- Caring for adult patients ≥18 years old.
- Able to request natriuretic peptide blood testing (a standard screening test for HF in patients with symptoms); and initiating the HF diagnostic pathway recommended by the National Institute for Health and Care Excellence (NICE)[35]

Exclusion criteria:

- Poor WiFi and/or mobile data connectivity within primary care consulting rooms, prohibiting use of the AI-enabled stethoscope.

- Not providing face-to-face patient consultations.

Patient and public involvement

Patients were involved in the conception and trial design via a patient steering group, and through online survey of over 10,000 patients at Imperial College Healthcare NHS Trust, UK. The study design was further developed in collaboration with the Pumping Marvellous Foundation, the UK's largest patient-led heart failure charity, and their Patient Educator group.

Enrolment

Potential participating primary care practices will be approached by their parent NHS organisations: the North West London (NWL) Integrated Care System, or Betsi Cadwaladr University Health Board (BCUHB), in England and Wales, respectively. Additionally, in NWL, primary care practices will be approached by the National Institute for Health and Care Research Clinical Research Network, whose mandate is to widen access to and diversify research participation. Practices expressing an interest will be contacted by the study team with up to two further follow-up emails, and a subsequent phone call if necessary. Written informed consent will be recorded from leadership at each participating practice. Clinicians at all practices will be asked to complete a baseline questionnaire measuring their confidence in detecting and managing cardiovascular disease. The study team will have no direct contact with patients.

Randomisation

Practices will be randomised 1:1 to the intervention or control arms using a validated, automated and audited randomisation tool (Sealed Envelope, London, UK)[40] with allocation concealment. Practices will be notified of their treatment allocation by email.

Intervention arm: Artificial intelligence (AI)-enabled stethoscope

This study will investigate the impact of an AI-enabled stethoscope (henceforth referred to as AI-stethoscope) with integrated sensor technology – electrodes and a microphone – for recording digital single-lead ECG and PCG (Eko DUO, Eko Health Inc, California, US). The AI-stethoscope works as a conventional stethoscope with diaphragm and detachable tubing that enables conventional auscultation, with other features including noise filtration and cancellation. The device connects to an app on the user's smartphone (Eko: Digital Stethoscope + ECG, Eko Health, US) using Bluetooth connectivity for waveform visualisation. Connectivity via cellular/Wi-Fi allows access to cloud-based AI algorithms for analysis of waveforms (no data is stored in the AI-stethoscope or user's smartphone).

The AI-stethoscope and associated AI algorithms are regulated by the UK Medicines and Healthcare Products Regulatory Agency (MHRA) and are UKCA-marked as Class IIa medical devices, respectively. This means that they are authorised for use in clinical care in the UK, in accordance with their regulated intended purpose.

The AI-stethoscope is provided as a replacement stethoscope for routine patient clinical examination. Use of the AI-stethoscope will be at the discretion of the responsible clinician, with no patient-level inclusion or exclusion criteria. The additional functionality in the form of AI algorithms for HFrEF, AF and VHD detection will be used within its regulated intended purpose. As established by previous studies[26], the optimum unit of examination to derive AI insights entails placing the device – "listening" – over the pulmonary position (upper left sternal border) and taking a fifteen second recording.

In an NHS first, the clinical guidelines for actioning AI outputs have been agreed by the Cardiovascular Executive Groups of the NHS North West London Integrated Care System and Betsi Cadwaladr University Health Board (Appendix). Patients will be examined with the AI-stethoscope in accordance with these guidelines, and/or where stethoscope examination is deemed clinically appropriate by the primary care clinician. Patients will provide verbal consent for examination with the AI-stethoscope as per any physical examination performed by healthcare professionals for direct care purposes, in accordance with UK law and General Medical Council guidelines[41].

Primary care clinicians at practices in the intervention arm will be provided with one session of in-person training in use of the AI-stethoscope. The training includes:

1. Delivery and setup of the AI-stethoscope in each consultation room
2. Smartphone application installation and login (for either iOS or Android systems) on each clinician's smartphone for use of the AI-stethoscope
3. Pairing of all clinician smartphones with all AI-stethoscopes in the same practice
4. Demonstration of patient examination, single-lead ECG capture from the 'spot check' auscultation position and AI recommendation
5. Troubleshooting signal capture
6. Visual summary sheet for patient examination and troubleshooting affixed next to AI-stethoscope in each consultation room.
7. Update on the local clinical guideline and summary sheet affixed next to the AI-stethoscope in each consultation room (Appendix).

If any clinicians (who can initiate diagnostic pathways for HF, AF or VHD) decline to install the application on their smartphone, this will be logged by the study team.

GP practices in the intervention arm will receive six-monthly updates (three total) from the clinical research team. These will be communicated via direct email to a nominated clinician from each practice team and include relevant data on numbers of positive screening results and subsequent clinician actions such as specialist referrals and implemented medical therapies. The reports will include monthly results of:

1. Number of patients with positive screening results
2. Proportion of patients with positive screening results referred for specialist review
3. Proportion of patients with positive screening results receiving guideline-direct medical therapies

At the end of the study, clinicians in the intervention arm will be asked to complete a validated usability questionnaire for the AI-stethoscope (System Usability Scale[42]).

Control arm

Primary care practices in the control arm will continue with usual clinical care, with decision-making for consideration of HF, and initiation of the NICE diagnostic pathway based on clinical judgement alone.

Participating practices can withdraw at any time. If a primary care practice ceases face-to-face consultations during the study, they will be withdrawn from the study by the research team.

Data sources

In the NWL ICS, patient-level complete primary and secondary care clinical and cost data is pseudonymised and pooled within Discover-NOW, a UK Trusted Research Environment[43]. AI-stethoscope usage statistics and AI predictions will be integrated into the Discover-NOW and tagged to appropriate patient records. The study team will access this real-world data platform to measure population-level clinical and health economic outcomes, in addition to patient-level sensitivity analyses for NWL ICS.

At BCUH, only the direct care team will access medical records and record population-level outcomes as part of a service evaluation (co-primary endpoints), using epidemiological methods only.

Study Outcomes

Outcomes are summarised in Table 1.

Study outcomes comparing intervention and control groups	
Co-primary endpoints	
i.	Incidence of coded new diagnoses of HF
ii.	Ratio of coded diagnoses of HF via hospital admission-based versus community-based pathways
Secondary endpoints	
iii.	New coded diagnosis of atrial fibrillation (AF)
iv.	New coded diagnosis of valvular heart disease (VHD)
v.	Cost-consequence differential (HF, AF, VHD)
vi.	Health service utilisation for diagnostics
vii.	Prescription of guideline-directed medical therapy for HF, AF, VHD
viii.	New implantation of cardiac resynchronisation therapy (CRT) and/or implantable cardiac defibrillator (ICD) devices
ix.	Differential rates of uptake and utilisation of AI-stethoscope in primary care
x.	Determinants of utilisation of AI-stethoscope in primary care
xi.	Quality of life – healthy days at home (HDAH)[44,45]

Independent variables for sensitivity and subgroup analyses	
xii.	GP practice population social deprivation
xiii.	GP practice national target performance
xiv.	GP practice clinical staffing model

Table 1. Study outcomes

Statistical analysis

Baseline patient-level data in NWL will be extracted at the index date, defined as the date of first examination using the AI-stethoscope at each primary care practice. Data will be summarised as mean (SD) or median (interquartile range) for skewed data. Comparisons between groups will be undertaken on an intention-to-treat basis, using two-sample t-tests for continuous variables, and $\chi 2$ tests for binary and categorical variables. For each outcome, we will perform generalised mixed-effect logistic regression to compare the study arms, with the GP practice as a random effect.

Subgroup analyses for patient-level outcomes will be performed, stratified by age, sex, ethnicity and geography, patient comorbidities and GP practice characteristics (mean age of clinicians, size of clinical team, number of registered patients).

Patient-level sensitivity analyses will be performed for patients with abnormal AI-stethoscope predictions for HF, to identify direct associations between AI-stethoscope predictions and specific diagnostic codes for HF, AF and VHD.

We will measure relationships between items in the clinician questionnaire, utilisation rates of the AI-stethoscope, compliance with recommendations and the primary endpoints. These analyses will identify domains that are most strongly related to utilisation and compliance, which will help improve the technology, and guide future implementation and clinical pathways.

Interim Analysis

An interim analysis will be performed 6 months after the first site has been randomised. The interim analysis will report the following selected outcomes and will be reviewed by the Trial Steering Committee, who will remain blinded to the treatment allocation of each group until the study end:

Co-primary endpoints:

a. Difference in incidence of coded HF between groups;
b. Difference in ratio of coded HF incidence via hospital admission-based versus community-based diagnostic pathways between groups

Secondary endpoints

a. Healthcare diagnostics utilisation
b. Primary care appointments
c. Emergency department presentations
d. Non-elective hospital admissions

e. AI-stethoscope utilisation rates
f. Completion of site setup for all practices in intervention group
g. Data fidelity - ascertainment

The following interim outcomes will be considered significant and warranting intervention:

a. Mean AI-stethoscope utilisation rate of less than five recordings per month per practice in the intervention arm
b. Healthcare diagnostics utilisation differential affecting patient safety
c. Incomplete site setup in intervention group

In keeping with formative evaluations in implementation research[29-31] and behavioural change frameworks[46], adaptive interventions will be considered prospectively to maximise the success of the implementation of AI-stethoscope (given the established confidence for the technology "working" on the basis of widely validated statistical accuracy). In addition, pre-specified interventions will be available to the Steering Committee following the interim analysis.

- Semi-structured interviews of high and low-use participants
- Provision of repeat training in use of the AI-stethoscope
- Redistribution of the AI-stethoscope to new, prospectively enrolled GP practices
- Provision of latest generation of the AI-stethoscope hardware and software to intervention group
- Integration of the AI-stethoscope system with the electronic health record system

To mitigate lead-time effects, interim analysis of the co-primary endpoints will not incur a stopping rule at 6 months. If at 6 months there is no statistically significant difference between treatment arms for the co-primary endpoints, then a further interim review will be performed at 12 months. If there remains no clinically significant difference in the co-primary endpoint at 12 months, the Trial Steering Group will have the option to stop the trial on the basis of futility.

ETHICS AND DISSEMINATION

TRICORDER will not require written/signed consent from individual patients since the AI-stethoscope has full regulatory (MHRA) approval for use in direct clinical care, and will be used within its regulatory approved intended purpose. The NHS Integrated Care Systems involved in TRICORDER have developed and approved a clinical guideline for use of the AI-stethoscope in direct care, through agreement between executive primary and secondary care stakeholders (Appendix), which are anchored in best practice national guidelines from the UK National Institute for Health and Care Excellence (NICE). As with any clinical test, the decision to perform it and act on results in line with guidelines is ultimately the underpinned by the clinical judgement of the responsible medical professional, and involves informed consent of the patient. This protocol has been

reviewed and received favourable opinion from a UK Health Research Authority (HRA) Research Ethics Committee (reference: 23/LO/0051), including its cluster design and confirmation that patient-level research consent is not indicated. For use of Discover-NOW, there is existing UK HRA approval for access procedures, pseudonymisation and use of this platform for outcomes research (reference:18/WM/0323). A scientific manuscript with the primary outcomes of the study will be published in a peer-reviewed journal. Further manuscripts reporting secondary outcomes will also be published in peer-reviewed journals. Results will be presented at scientific meetings, and conferences with local and national stakeholders, including the national funder (National Institute for Health and Care Research), NHS organisations, and patient and public groups such as the Pumping Marvellous Foundation.

Trial registration

TRICORDER is registered with the NIH National Library of Medicine (NCT05987670). The study protocol has been reported in accordance with the SPIRIT-AI Extension[47] (see Appendix) and outcomes will be reported in accordance with the CONSORT-AI Extension[48].

Discussion

TRICORDER is the first cluster-randomised implementation trial to evaluate a real-time, AI-dependent point-of-care diagnostic in primary care. The study is designed to meet the specifications of the NICE Evidence Standards Framework for Digital Health Technologies[49] and therefore underpin recommendations for subsequent NHS-wide commissioning. This protocol – and lessons from execution of the trial – will inform a replicable blueprint for evaluation of other similar digital health technologies.

The study is designed to address the unacceptable reality that cardiovascular disease, heart failure particularly, is most frequently detected at a late stage, after disease progression precipitates a hospital admission[17,18]. The NHS Long Term plan prioritises the need to reverse this trend, emphasising the need for increased rates of diagnosis through community (primary care initiated) pathways. [1]. We have recently highlighted the £2,500 saving unlocked by every patient diagnosed with heart failure through such community pathways[18]. Similar models inform a compelling health economic case for community-based detection of AF and VHD.[19,20]

By spanning primary care practices across urban and rural geographies and serving populations with varied sociodemographic and ethnic backgrounds, this study will encompass a uniquely representative patient population. This is intended to consider the well-established concerns around AI bias[50], the digital maturity for technology adoption[51], and to demonstrate the AI-stethoscope's suitability for deployment across diverse patient and clinician populations. The research team will collate qualitative data via clinician surveys, ensuring the AI-stethoscope's efficient integration into primary care, preservation of the traditional clinician-patient interaction and high-level usability of the device.

This study is both decentralised and principally uses "real-world" data; there is no lead/central study site or logistics/travel outside each GP practice. Each will contribute data equally, but without the need to complete any study-specific (cumbersome) data collection instruments. Instead, our use of Discover-NOW enables outcomes measurement with routinely recorded (real world) data. The platform's offer of comprehensive linked (primary and secondary care) clinical and cost data affords patient-level sensitivity analyses. This will allow robust measurement of associations between diagnostic outcomes and use of the AI-stethoscope. Separate to such novel elements of this study, TRICORDER has set several other important precedents for translational AI research, including approval of the first regional NHS clinical guideline for use of a primary care focused AI technology; and sector-wide data governance approval covering use of an AI technology across hundreds of disparate primary care sites.

This research protocol is best interpreted in the context of its limitations. The pragmatic design, which aims to have a minimal impact on GP workflows, may in some cases limit sustained use of the technology by not setting any expectations or requirement for use – though this may somewhat mitigate any Hawthorne effect[52]. This study is the first to accompany a novel AI technology with an NHS sector-approved clinical guideline for use, but variable adherence may limit impact attributable to the technology. This will be addressed systematically, taking an implementation science approach to maximise uptake of the intervention. Finally, the examination of real-world data is universally limited by the inconsistency and variable fidelity of medical coding in capturing specific variables of interest. For example, for HF, this is rarely coded with a granularity that describes preserved, moderately reduced, or reduced ejection fraction. However, the otherwise comprehensive dataflows associated with this study will allow holistic scrutiny of a broad selection of outcomes that need to be understood to underpin any recommendations for system-wide uptake. Ultimately, this will serve to provide a realistic estimate of the potential clinical and health economic impact of the AI-stethoscope technology in NHS primary care.

AUTHORS CONTRIBUTIONS

MAK, PB and NSP were involved in conception and trial design. CFP, CC and EF provided statistical expertise. RAL provided expertise on patient and public involvement. MAK and PB were involved in drafting this manuscript. JM, DBK, MTA, AP, AA, CB, AM, CMP and NSP provided critical revision of the manuscript and intellectual content. All authors were involved in the final approval of the article.

ACKNOWLEDGEMENTS

The authors are grateful for the infrastructural support from the National Institute of Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. AM is supported by the NIHR Applied Research Collaboration NW London. The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. We are indebted to the co-development with and feedback from patients from the Pumping Marvellous Foundation heart failure charity.

FUNDING

This work was supported by the National Institute for Health and Care Research Invention for Innovation (i4i) Challenge Award, Accelerated Access Collaborative and NHSX via the AI Award in Health and Social Care; Imperial Health Charity; British Heart Foundation.

Funders were not involved in the study design, writing of the report or decision to submit the article for publication. All authors had full access to the protocol and take responsibility for the integrity of the manuscript.

COMPETING INTERESTS

M.A.K. has nothing to declare; P.B. has nothing to declare; J.M. has nothing to declare; C.F.P. has nothing to declare; M.T.A. has nothing to declare; A.A. has nothing to declare; D.B.K. has nothing to declare; A.P.. has nothing to declare; C.C. has nothing to declare; E.F. has nothing to declare; R.A.L. has nothing to declare; A.M. has nothing to declare; C.M.P. has nothing to declare; N.S.P. is an advisor to Eko Health Inc.

REFERENCES

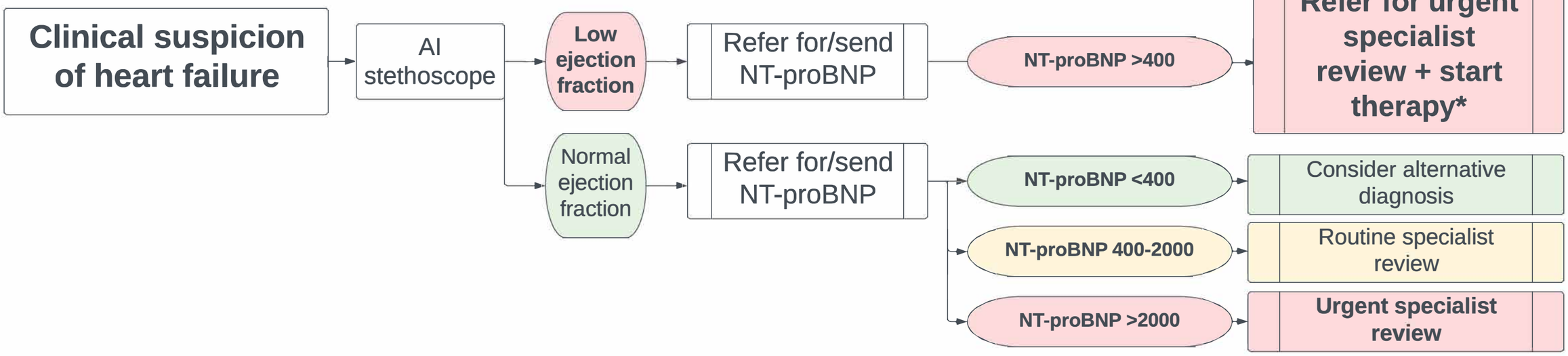
1. NHS. NHS Long Term Plan. 2019 (https://www.longtermplan.nhs.uk/).
2. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 2020;396(10258):1204-1222. DOI: 10.1016/S0140-6736(20)30925-9.
3. Organisation WH. Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. World Health Organisation, 2020.
4. Kannel WB, Abbott RD, Savage DD, McNamara PM. Epidemiologic Features of Chronic Atrial Fibrillation. New England Journal of Medicine 1982;306(17):1018-1022. DOI: 10.1056/nejm198204293061703.
5. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA. Independent Risk Factors for Atrial Fibrillation in a Population-Based Cohort: The Framingham Heart Study. JAMA 1994;271(11):840-844. DOI: 10.1001/jama.1994.03510350050036.
6. Krum H, Gilbert RE. Demographics and concomitant disorders in heart failure. The Lancet 2003;362(9378):147-158. DOI: https://doi.org/10.1016/S0140-6736(03)13869-X.
7. EHRLICH JR, NATTEL S, HOHNLOSER SH. Atrial Fibrillation and Congestive Heart Failure: Specific Considerations at the Intersection of Two Common and Important Cardiac Disease Sets. Journal of Cardiovascular Electrophysiology 2002;13(4):399-405. DOI: https://doi.org/10.1046/j.1540-8167.2002.00399.x.
8. Meyer S, Brouwers FP, Voors AA, et al. Sex differences in new-onset heart failure. Clin Res Cardiol 2015;104(4):342-50. (In eng). DOI: 10.1007/s00392-014-0788-x.
9. Brouwers FP, de Boer RA, van der Harst P, et al. Incidence and epidemiology of new onset heart failure with preserved vs. reduced ejection fraction in a community-based cohort: 11-year follow-up of PREVEND. Eur Heart J 2013;34(19):1424-31. (In eng). DOI: 10.1093/eurheartj/eht066.
10. Conrad N, Judge A, Tran J, et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. Lancet (London, England) 2018;391(10120):572-580. (In eng). DOI: 10.1016/s0140-6736(17)32520-5.
11. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet (London, England) 2018;392(10159):1789-1858. (In eng). DOI: 10.1016/s0140-6736(18)32279-7.
12. Roger VL. Epidemiology of heart failure. Circ Res 2013;113(6):646-59. (In eng). DOI: 10.1161/circresaha.113.300268.
13. Mosterd A, Hoes AW. Clinical epidemiology of heart failure. Heart 2007;93(9):1137-46. (In eng). DOI: 10.1136/hrt.2003.025270.

14. Mamas MA, Sperrin M, Watson MC, et al. Do patients have worse outcomes in heart failure than in cancer? A primary care-based cohort study with 10-year follow-up in Scotland. European Journal of Heart Failure 2017;19(9):1095-1104. (https://doi.org/10.1002/ejhf.822). DOI: https://doi.org/10.1002/ejhf.822.
15. Braunschweig F, Cowie MR, Auricchio A. What are the costs of heart failure? Europace 2011;13 Suppl 2:ii13-7. (In eng). DOI: 10.1093/europace/eur081.
16. Excellence NIfHaC. Resource impact report: Chronic heart failure in adults: diagnosis and management (NG106). https://www.nice.org.uk/guidance/ng106/resources/resource-impact-report-pdf-6537494413#:~:text=2.3%20It%20is%20estimated%20that,into%20Living%20with%20Heart%20failure: National Institute for Health and Care Excellence, 2018.
17. Bottle A, Kim D, Aylin P, Cowie MR, Majeed A, Hayhoe B. Routes to diagnosis of heart failure: observational study using linked data in England. Heart 2018;104(7):600-605. (In eng). DOI: 10.1136/heartjnl-2017-312183.
18. Bachtiger P, Kelshiker MA, Petri CF, et al. Survival and health economic outcomes in heart failure diagnosed at hospital admission versus community settings: a propensity-matched analysis. BMJ Health & Care Informatics 2023;30(1):e100718. DOI: 10.1136/bmjhci-2022-100718.
19. Burdett P, Lip GYH. Atrial fibrillation in the UK: predicting costs of an emerging epidemic recognizing and forecasting the cost drivers of atrial fibrillation-related costs. European Heart Journal - Quality of Care and Clinical Outcomes 2020;8(2):187-194. DOI: 10.1093/ehjqcco/qcaa093.
20. Moore M, Chen J, Mallow PJ, Rizzo JA. The direct health-care burden of valvular heart disease: evidence from US national survey data. Clinicoecon Outcomes Res 2016;8:613-627. (In eng). DOI: 10.2147/ceor.S112691.
21. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021;42(36):3599-3726. (In eng). DOI: 10.1093/eurheartj/ehab368.
22. Hindricks G, Potpara T, Dagres N, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. European Heart Journal 2021;42(5):373-498. DOI: 10.1093/eurheartj/ehaa612.
23. Vahanian A, Beyersdorf F, Praz F, et al. 2021 ESC/EACTS Guidelines for the management of valvular heart disease: Developed by the Task Force for the management of valvular heart disease of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS). European Heart Journal 2021;43(7):561-632. DOI: 10.1093/eurheartj/ehab395.
24. Fonseca C, Morais H, Mota T, et al. The diagnosis of heart failure in primary care: value of symptoms and signs. Eur J Heart Fail 2004;6(6):795-800, 821-2. (In eng). DOI: 10.1016/j.ejheart.2004.08.002.

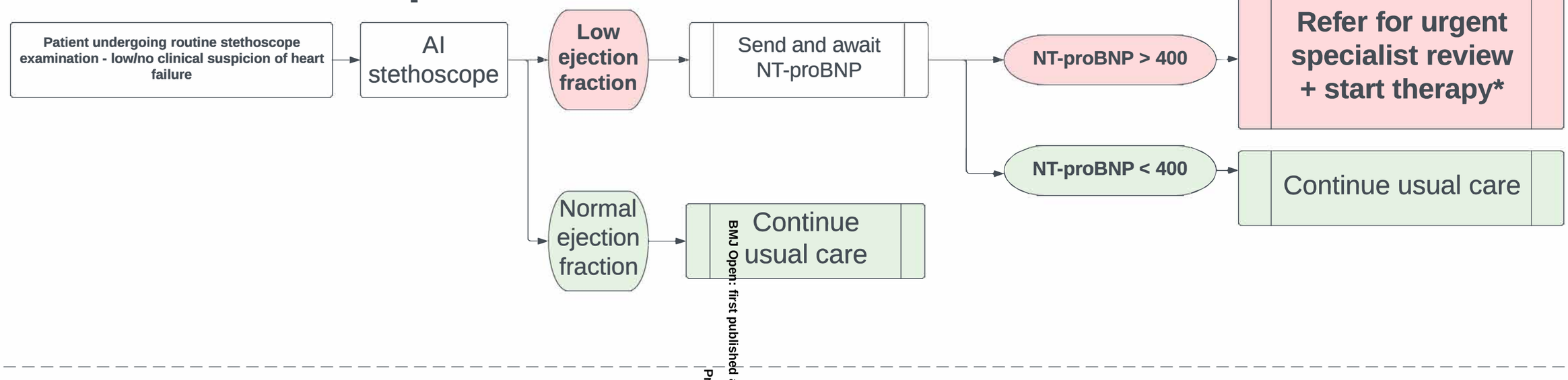
25. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2013;62(16):e147-239. (In eng). DOI: 10.1016/j.jacc.2013.05.019.
26. Bachtiger P, Petri CF, Scott FE, et al. Point-of-care screening for heart failure with reduced ejection fraction using artificial intelligence during ECG-enabled stethoscope examination in London, UK: a prospective, observational, multicentre study. Lancet Digit Health 2022 (In eng). DOI: 10.1016/s2589-7500(21)00256-9.
27. Davidsen AH, Andersen S, Halvorsen PA, Schirmer H, Reierth E, Melbye H. Diagnostic accuracy of heart auscultation for detecting valve disease: a systematic review. BMJ Open 2023;13(3):e068121. (In eng). DOI: 10.1136/bmjopen-2022-068121.
28. Chorba JS, Shapiro AM, Le L, et al. Deep Learning Algorithm for Automated Cardiac Murmur Detection via a Digital Stethoscope Platform. Journal of the American Heart Association 2021;10(9):e019905. DOI: 10.1161/JAHA.120.019905.
29. Bauer MS, Kirchner J. Implementation science: What is it and why should I care? Psychiatry Res 2020;283:112376. (In eng). DOI: 10.1016/j.psychres.2019.04.025.
30. Elwy AR, Wasan AD, Gillman AG, et al. Using formative evaluation methods to improve clinical implementation efforts: Description and an example. Psychiatry Res 2020;283:112532. (In eng). DOI: 10.1016/j.psychres.2019.112532.
31. Margaret AH, Anuradha G, Adithya C. Strategies for implementing implementation science: a methodological overview. Emergency Medicine Journal 2016;33(9):660. DOI: 10.1136/emermed-2015-205461.
32. Parikh RB, Teeple S, Navathe AS. Addressing Bias in Artificial Intelligence in Health Care. Jama 2019;322(24):2377-2378. (In eng). DOI: 10.1001/jama.2019.18058.
33. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019;28(3):231-237. (In eng). DOI: 10.1136/bmjqs-2018-008370.
34. Digital N. Patients Registered at a GP Practice. March 2023 2023.
35. Excellence NIfHaC. Chronic heart failure in adults: diagnosis and management. National Institute for Health and Care Excellence, 2018. (NG106) (https://www.nice.org.uk/guidance/ng106/resources/chronic-heart-failure-in-adults-diagnosis-and-management-pdf-66141541311685).
36. Cowie MR. The heart failure epidemic: a UK perspective. Echo Res Pract 2017;4(1):R15-r20. (In eng). DOI: 10.1530/erp-16-0043.
37. Bottle A, Cohen C, Lucas A, et al. How an electronic health record became a real-world research resource: comparison between London's Whole Systems Integrated Care database and the Clinical Practice Research Datalink. BMC Med Inform Decis Mak 2020;20(1):71. (In eng). DOI: 10.1186/s12911-020-1082-7.
38. Yao X, Rushlow DR, Inselman JW, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. Nat Med 2021;27(5):815-819. DOI: 10.1038/s41591-021-01335-4.

39. Kahn M, Grayson AD, Chaggar PS, et al. Primary care heart failure service identifies a missed cohort of heart failure patients with reduced ejection fraction. Eur Heart J 2022;43(5):405-412. (In eng). DOI: 10.1093/eurheartj/ehab629.
40. Ltd SE. Simple randomisation service. https://www.sealedenvelope.com/simple-randomiser/v1/: Sealed Envelope Ltd; 2022.
41. Council GM. Good Medical Practice. General Medical Council; 2013.
42. Brooke J. SUS: A quick and dirty usability scale. Usability Eval Ind 1995;189.
43. Bottle A, Cohen C, Lucas A, et al. How an electronic health record became a real-world research resource: comparison between London's Whole Systems Integrated Care database and the Clinical Practice Research Datalink. BMC Medical Informatics and Decision Making 2020;20(1):71. DOI: 10.1186/s12911-020-1082-7.
44. Greene SJ, O'Brien EC, Mentz RJ, et al. Home-Time After Discharge Among Patients Hospitalized With Heart Failure. J Am Coll Cardiol 2018;71(23):2643-2652. (In eng). DOI: 10.1016/j.jacc.2018.03.517.
45. Lam MB, Riley KE, Zheng J, Orav EJ, Jha AK, Burke LG. Healthy days at home: A population-based quality measure for cancer patients at the end of life. Cancer 2021;127(22):4249-4257. (In eng). DOI: 10.1002/cncr.33817.
46. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implementation Science 2011;6(1):42. DOI: 10.1186/1748-5908-6-42.
47. Rivera SC, Liu X, Chan A-W, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. BMJ 2020;370:m3210. DOI: 10.1136/bmj.m3210.
48. Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nature Medicine 2020;26(9):1364-1374. DOI: 10.1038/s41591-020-1034-x.
49. Unsworth H, Dillon B, Collinson L, et al. The NICE Evidence Standards Framework for digital health and care technologies – Developing and maintaining an innovative evidence framework with global impact. DIGITAL HEALTH 2021;7:205520762110186. DOI: 10.1177/20552076211018617.
50. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366(6464):447-453. DOI: doi:10.1126/science.aax2342.
51. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017;19(11):e367. (In eng). DOI: 10.2196/jmir.8775.
52. Chiesa M, Hobbs S. Making sense of social research: how useful is the Hawthorne Effect? European Journal of Social Psychology 2008;38(1):67-74. DOI: https://doi.org/10.1002/ejsp.401.

Clinical suspicion of heart failure



Low/no clinical suspicion of heart failure



*Initiating Heart Failure prognostic therapy in Primary Care with AI stethoscope

NT-proBNP > 400

AND

AI stethoscope

eGFR >45 HR >70bpm in SR (>90bpm in AF) SBP > 120mmHg	Bisoprolol 1.25mg PO OD AND Ramipril 1.25mg PO OD
eGFR 20 - 45 HR >70bpm in SR (>90bpm in AF) SBP > 120mmHg Not on insulin	Bisoprolol 1.25mg PO OD AND SGLT2 inhibitor according to license*
eGFR 20 - 45 HR <70bpm in SR (<90bpm in AF) SBP > 120mmHg Not on insulin	Ramipril 1.25mg PO OD AND SGLT2 inhibitor according to license*
eGFR >20 HR <70bpm in SR (<90bpm in AF) SBP < 110mmHg Not on insulin	SGLT2 inhibitor according to license*

*e.g. Dapagliflozin / Empagliflozin 10mg PO OD

Check renal function within four weeks of starting medications

Titrate Furosemide as needed to symptoms: 20mg to 40mg PO OD

Approved by NWL ICB Cardiovascular Executive
Authorised by: NWL ICB Heart Failure Working Group

BMJ Open

Triple Cardiovascular Disease Detection with an Artificial Intelligence-enabled Stethoscope (TRICORDER): design and rationale for a decentralised, real-world cluster randomised controlled trial and implementation study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-098030.R1
Article Type:	Protocol
Date Submitted by the Author:	23-Mar-2025
Complete List of Authors:	Kelshiker, Mihir A; Imperial College London National Heart and Lung Institute; Imperial College Healthcare NHS Trust Bächtiger, Patrik; Imperial College London National Heart and Lung Institute; Imperial College Healthcare NHS Trust Mansell, Josephine; Imperial College London National Heart and Lung Institute; Imperial College Healthcare NHS Trust Kramer, Daniel B; Imperial College London National Heart and Lung Institute; Harvard Medical School Petri, Camille F; Imperial College London National Heart and Lung Institute Almonte, Melanie T; Imperial College London National Heart and Lung Institute; Imperial College Healthcare NHS Trust Alrumayh, Abdullah; Imperial College London National Heart and Lung Institute Peters, Alexei; Imperial College London National Heart and Lung Institute Costelloe, Ceire; Imperial College London, Global Digital Health Unit, Department of Primary Care and Public Health Falaschetti, Emanuela; Imperial College London, School of Public Health Barton, Carys; Imperial College Healthcare NHS Trust Al-Lamee, Rasha; Imperial College London, National Heart and Lung Institute Majeed, Azeem; Imperial College London, Primary Care Plymen, Carla M; Imperial College London; Imperial College Healthcare NHS Trust Peters, Nicholas S; Imperial College London National Heart and Lung Institute; Imperial College Healthcare NHS Trust
Primary Subject Heading:	Cardiovascular medicine
Secondary Subject Heading:	General practice / Family practice, Health economics, Health informatics, Public health
Keywords:	Artificial Intelligence, CARDIOLOGY, HEALTH ECONOMICS, PUBLIC HEALTH, Heart failure < CARDIOLOGY, Primary Health Care

Triple Cardiovascular Disease Detection with an Artificial Intelligence-enabled Stethoscope (TRICORDER): design and rationale for a decentralised, real-world cluster randomised controlled trial and implementation study

Mihir A Kelshiker*[1,2], Patrik Bächtiger*[1,2], Josephine Mansell[1,2], Daniel B Kramer[1], Camille F Petri[1], Melanie T Almonte[1,2], Abdullah Alrumayh[1], Alexei Peters[1], Céire Costelloe[4], Emanuela Falaschetti[5], Carys Barton, Rasha Al-Lamee[1,2], Azeem Majeed[4], Carla M Plymen[1,2], Nicholas S Peters[1,2]**

1. National Heart and Lung Institute, Imperial College London, UK
2. Imperial College Healthcare NHS Trust, UK
3. School of Public Health, Imperial College London, UK
4. Imperial Clinical Trials Unit, School of Public Health, Imperial College London, UK

*contributed equally, **corresponding author.

Corresponding author:

Professor Nicholas S Peters

National Heart and Lung Institute, Imperial College London

Hammersmith Hospital

Du Cane Road

London, W12 0HS

n.peters@imperial.ac.uk

Funding: National Institute for Health Research Invention for Innovation (i4i) Challenge Award, Accelerated Access Collaborative and NHSX via the AI Award in Health and Social Care; British Heart Foundation; Imperial Health Charity.

Keywords: Cardiology/Artificial Intelligence/Digital Health/Primary Care/Public Health

ABSTRACT

Introduction

Early detection of cardiovascular disease in primary care is a public health priority, for which the clinical- and cost-effectiveness of an artificial intelligence-enabled stethoscope that detects left ventricular systolic dysfunction, atrial fibrillation, and cardiac murmurs is unproven but potentially transformative.

Methods and analysis

TRICORDER is a pragmatic, two-arm, multi-centre (decentralised), cluster randomised controlled trial and implementation study. Up to 200 primary care practices in urban North West London and rural North Wales, UK, will be randomised to usual care or to have artificial intelligence-enabled stethoscopes available for use. Primary care clinicians will use the artificial intelligence-enabled stethoscopes at their own discretion, without patient-level inclusion or exclusion criteria. They will be supported to do so by a clinical guideline developed and approved by the regional health system executive board. Patient and outcome data will be captured from pooled primary and secondary care records, supplemented by qualitative and quantitative clinician surveys. The co-primary endpoints are i. difference in coded incidence (detection) of heart failure and ii. difference in ratio of coded incidence of heart failure via hospital admission versus community-based diagnostic pathways. Secondary endpoints include difference in incidence of atrial fibrillation and valvular heart disease, cost-consequence differential, and prescription of guideline-directed medical therapy.

Ethics and dissemination

This trial has ethical approval from the UK Health Research Authority (23/LO/0051). Findings from this trial will be disseminated through publication of peer-reviewed manuscripts, presentations at scientific meetings and conferences with local and national stakeholders.

Trial registration number: NCT05987670

Summary

TRICORDER will examine the clinical- and cost-effectiveness of an artificial intelligence-enabled stethoscope for detection of priority cardiovascular diseases in real-world primary care. Findings will underpin implementation strategies for at-scale deployment and evaluation of artificial intelligence-enabled tools in routine clinical practice.

Strengths and limitations of this study

- Use of real-world data for outcomes evaluation minimises the administrative impact on frontline primary care services, and enhances the generalisability of results.
- Outcome assessment from routine medical coding lacks the granularity and fidelity of primary research form collection.
- Integration of AI predictions with routinely collected electronic medical record data allows sensitivity analyses at the level of patient, site, and region.

INTRODUCTION

Early detection of cardiovascular disease is a global public health priority[1-3]. Heart failure (HF), atrial fibrillation (AF) and valvular heart disease (VHD) can have common pathophysiology and significant mortality and morbidity[4-7]. Given the reported adult prevalence of HF of 2%, the recognised incidence of HF in Europe of 5 per 1000 person-years in adults[8,9]is likely to be an underestimate[10-13]. HF is deadlier than common, serious cancers[14], profoundly impacting quality of life and costing the UK National Health Service (NHS) over £2 billion per year – or over 2% of its annual budget[15,16]. Unacceptably, 70-80% of all new HF diagnoses are only made after an emergency hospital admission, despite most patients having been seen in primary care with opportunities for detection and therefore earlier initiation of disease-modifying treatment[17,18] ,expected to translate to improved survival, quality of life and health service utilisation costs[18].

Notwithstanding the mixed evidence of effectiveness of screening programmes to improve mortality and stroke associated with AF, the substantial clinical and health economic costs of AF and VHD are similarly driven by late diagnosis[19,20].

Across these cardiac conditions, early detection reduces the risk of disease progression and improves survival, through initiation of inexpensive, guideline-directed therapies[21-23]. However, primary care providers face the challenge that symptoms (if present) are frequently non-specific[24,25], and there are no universally effective point-of-care screening tools to identify the diagnosis or diagnostic pathways.

We have recently shown that artificial intelligence (AI) applied to electrocardiogram (ECG) and phonocardiogram (PCG) waveforms captured by a “smart” stethoscope can detect left ventricular ejection fraction ≤40%[26], the cause of heart failure with reduced ejection fraction (HFrEF, the most common form of HF); AF, and cardiac murmurs – a cardinal examination finding in valvular heart disease[27]. The statistical performance of these three AI algorithms has been shown to be high and consistent against international external validation studies[28]. This takes a familiar clinical tool with an established workflow and augments examination findings with actionable additional insights, taking only 15 seconds[26]. The clinical and health economic impact of deploying such a triple point-of-care AI diagnostic in primary care practices is unknown. Furthermore, the determinants of successfully implementing[29-31] such an innovation in routine clinical practice are also not known, given the historical failure of AI technology adoption in healthcare systems that is partly attributable to inattention to implementation science, and a failure to overcome bias[32,33].

The objective of the TRICORDER study is therefore to conduct a real-world, decentralised cluster randomised controlled trial and implementation study to determine if providing primary care teams with an AI-stethoscope improves early diagnosis of HFrEF, AF and VHD. Secondary objectives include measurement of cost effectiveness, examination of implementation strategies and identifying the determinants of uptake and utilisation of the AI-stethoscope in routine clinical practice. This report describes the clinical and policy context for this publicly funded clinical trial,

details of the protocol and plans for implementation, and expected findings with a focus on implications for widespread adoption in health systems.

METHODS AND ANALYSIS

Setting

TRICORDER is set within the primary care system of the NHS. General Practitioners (GPs) within primary care teams are the gatekeepers to most specialist diagnostic and management services in the NHS through nationally agreed pathways recommended by the National Institute for Health and Care Excellence (NICE), Figure 1. Each primary care practice (>8,000 in UK) serves an average list of 8,500 patients in the community[34]. Typically, patients in whom the GP suspects HF are referred for natriuretic peptide (NP) testing[35], which may be via in-practice phlebotomy or via hospital appointment. The NP level determines the urgency of referral for echocardiography and specialist review. There is substantial geographical variation in access to community-based echocardiography services, and lead times to hospital outpatient clinic review[36].

Study dates – October 2023 to October 2025

Funding

TRICORDER is funded by the UK National Institute for Health and Care Research (NIHR) Invention for Innovation Challenge Award. This is a competitively awarded public funding source, following a multi-stage process involving internal, external clinical and academic peer review, in addition to patient and public scrutiny. Eko Health are not involved in the design, conduct, analysis, interpretation of data, or reporting of TRICORDER.

Study design

TRICORDER is an open-label, two-arm cluster randomised controlled trial. Taking a decentralised approach, primary care practices in the NHS North West London Integrated Care System (NWL ICS, 2.8 million residents, Figure 2) and the Betsi Cadwaladr University Health Board (BCUHB, 670,000 residents) in North Wales, UK, will be enrolled for 18 months.

Each individual primary care practice employing GPs serves as a cluster – the unit of participation and randomisation. Each participating cluster will variously employ GPs, Practice Nurses, Physician Associates and other healthcare professionals who constitute the care team. This unit of recruitment was chosen because the NHS clinical, financial and information governance processes are standardised at this level.

Sample size

Interrogation of the Discover dataset[37] (UK Secure Data Environment for TRICORDER endpoint analysis) suggests an incidence rate of HF of 0.62/1000 patients per year. Combining NWL (n=350) and North Wales (n=>50), there are at least 400 potentially eligible study sites. We estimate that each GP practice will have at least 60 eligible patients per year who will receive a stethoscope examination as part of their routine clinical care. Estimating a conservative recruitment ratio of only 50%, we anticipate 200 GP practices being recruited. This provides an eligible study population of 6,000 patients per study arm. A review of prior literature in the field indicates an intracluster correlation coefficient of 0.01[38].

Utilising these estimates and assumptions, we would have over 80% power to detect a statistically and clinically significant mean difference if HF is diagnosed at a rate of 0.79/1000 patients in the intervention group, assuming a two-tailed test with a type I error of 0.05. This corresponds to a relative increase of 22%, which is deemed to be clinically meaningful (~17 to 22 HF diagnoses/GP practice).

The power calculation refers to overall HF, which includes all relevant ICD-10 and SNOMED-CT codes (used in UK primary care), rather than the codes relating to HFrEF specifically. There are three reasons for such an inclusive approach.

Firstly, we aim to mitigate an artificially low incidence of HFrEF due to coding heterogeneity in primary care. Preliminary analysis of the Discover dataset revealed substantial heterogeneity in clinical coding of HF subtypes. This is well recognised in recent studies[39].

Secondly, this approach captures the effect on clinician behaviour from participation in the study – it is feasible that more of all HF subtypes may be detected during the study period.

Finally, this approach enables rapid clinical and health economic endpoint analysis, avoiding the time and cost burden of manual data entry/data collection from chart arbitration, which would otherwise risk undermining the scalability of TRICORDER and set an impractical precedent for longer-term outcomes analyses.

Inclusion and exclusion criteria

Eligible primary care practices must meet the following inclusion criteria:

- Caring for adult patients ≥18 years old.
- Able to request natriuretic peptide blood testing (a standard screening test for HF in patients with symptoms); and initiating the HF diagnostic pathway recommended by the National Institute for Health and Care Excellence (NICE)[35]

Exclusion criteria:

- Poor WiFi and/or mobile data connectivity within primary care consulting rooms, prohibiting use of the AI-enabled stethoscope.
- Not providing face-to-face patient consultations.

Patient and public involvement

Patients were involved in the conception and trial design via a patient steering group, and through online survey of over 10,000 patients at Imperial College Healthcare NHS Trust, UK. The study design was further developed in collaboration with the Pumping Marvellous Foundation, the UK's largest patient-led heart failure charity, and their Patient Educator group.

Enrolment

Potential participating primary care practices will be approached by their parent NHS organisations: the North West London (NWL) Integrated Care System, or Betsi Cadwaladr University Health Board (BCUHB), in England and Wales, respectively. Additionally, in NWL, primary care practices will be approached by the National Institute for Health and Care Research Clinical Research Network, whose mandate is to widen access to and diversify research participation. Practices expressing an interest will be contacted by the study team with up to two further follow-up emails, and a subsequent phone call if necessary. Written informed consent will be recorded from leadership at each participating practice. Clinicians at all practices will be asked to complete a baseline questionnaire measuring their confidence in detecting and managing cardiovascular disease. The study team will have no direct contact with patients.

Randomisation

Practices will be randomised 1:1 to the intervention or control arms using a validated, automated and audited randomisation tool (Sealed Envelope, London, UK)[40] with allocation concealment. Practices will be notified of their treatment allocation by email.

Intervention arm: Artificial intelligence (AI)-enabled stethoscope

This study will investigate the impact of an AI-enabled stethoscope (henceforth referred to as AI-stethoscope) with integrated sensor technology – electrodes and a microphone – for recording digital single-lead ECG and PCG (Eko DUO, Eko Health Inc, California, US). The AI-stethoscope works as a conventional stethoscope with diaphragm and detachable tubing that enables conventional auscultation, with other features including noise filtration and cancellation. The device connects to an app on the user's smartphone (Eko: Digital Stethoscope + ECG, Eko Health, US) using Bluetooth connectivity for waveform visualisation. Connectivity via cellular/Wi-Fi allows access to cloud-based AI algorithms for analysis of waveforms (no data is stored in the AI-stethoscope or user's smartphone).

The AI-stethoscope and associated AI algorithms are regulated by the UK Medicines and Healthcare Products Regulatory Agency (MHRA) and are UKCA-marked as Class IIa medical devices, respectively. This means that they are authorised for use in clinical care in the UK, in accordance with their regulated intended purpose.

The AI-stethoscope is provided as a replacement stethoscope for routine patient clinical examination. Use of the AI-stethoscope will be at the discretion of the responsible clinician, with no patient-level inclusion or exclusion criteria. The additional functionality in the form of AI algorithms for HFrEF, AF and VHD detection will be used within its regulated intended purpose. As established by previous studies[26], the optimum unit of examination to derive AI insights entails placing the device – "listening" – over the pulmonary position (upper left sternal border) and taking a fifteen second recording.

In an NHS first, the clinical guidelines for actioning AI outputs have been agreed by the Cardiovascular Executive Groups of the NHS North West London Integrated Care System and Betsi Cadwaladr University Health Board (Supplementary Figure 1). Patients will be examined with the AI-stethoscope in accordance with these guidelines, and/or where stethoscope examination is deemed clinically appropriate by the primary care clinician. Patients will provide verbal consent for examination with the AI-stethoscope as per any physical examination performed by healthcare professionals for direct care purposes, in accordance with UK law and General Medical Council guidelines[41].

Primary care clinicians at practices in the intervention arm will be provided with one session of in-person training in use of the AI-stethoscope. The training includes:

1. Delivery and setup of the AI-stethoscope in each consultation room
2. Smartphone application installation and login (for either iOS or Android systems) on each clinician's smartphone for use of the AI-stethoscope
3. Pairing of all clinician smartphones with all AI-stethoscopes in the same practice
4. Demonstration of patient examination, single-lead ECG capture from the 'spot check' auscultation position and AI recommendation
5. Troubleshooting signal capture
6. Visual summary sheet for patient examination and troubleshooting affixed next to AI-stethoscope in each consultation room.
7. Update on the local clinical guideline and summary sheet affixed next to the AI-stethoscope in each consultation room (Supplementary Figure 1).

If any clinicians (who can initiate diagnostic pathways for HF, AF or VHD) decline to install the application on their smartphone, this will be logged by the study team.

GP practices in the intervention arm will receive six-monthly updates (three total) from the clinical research team. These will be communicated via direct email to a nominated clinician from each practice team and include relevant data on numbers of positive screening results and subsequent clinician actions such as specialist referrals and implemented medical therapies. The reports will include monthly results of:

1. Number of patients with positive screening results
2. Proportion of patients with positive screening results referred for specialist review
3. Proportion of patients with positive screening results receiving guideline-direct medical therapies

At the end of the study, clinicians in the intervention arm will be asked to complete a validated usability questionnaire for the AI-stethoscope (System Usability Scale[42]).

Control arm

Primary care practices in the control arm will continue with usual clinical care, with decision-making for consideration of HF, and initiation of the NICE diagnostic pathway based on clinical judgement alone.

Participating practices can withdraw at any time. If a primary care practice ceases face-to-face consultations during the study, they will be withdrawn from the study by the research team.

Data sources

In the NWL ICS, patient-level complete primary and secondary care clinical and cost data is pseudonymised and pooled within Discover-NOW, a UK Trusted Research Environment[43]. AI-stethoscope usage statistics and AI predictions will be integrated into the Discover-NOW and tagged to appropriate patient records. The study team will access this real-world data platform to measure population-level clinical and health economic outcomes, in addition to patient-level sensitivity analyses for NWL ICS.

At BCUH, only the direct care team will access medical records and record population-level outcomes as part of a service evaluation (co-primary endpoints), using epidemiological methods only.

Study Outcomes

Outcomes are summarised in Table 1.

Study outcomes comparing intervention and control groups	
Co-primary endpoints	
i.	Incidence of coded new diagnoses of HF
ii.	Ratio of coded diagnoses of HF via hospital admission-based versus community-based pathways
Secondary endpoints	
iii.	New coded diagnosis of atrial fibrillation (AF)
iv.	New coded diagnosis of valvular heart disease (VHD)
v.	Cost-consequence differential (HF, AF, VHD)
vi.	Health service utilisation for diagnostics
vii.	Prescription of guideline-directed medical therapy for HF, AF, VHD
viii.	New implantation of cardiac resynchronisation therapy (CRT) and/or implantable cardiac defibrillator (ICD) devices
ix.	Differential rates of uptake and utilisation of AI-stethoscope in primary care
x.	Determinants of utilisation of AI-stethoscope in primary care
xi.	Quality of life – healthy days at home (HDAH)[44,45]

Independent variables for sensitivity and subgroup analyses	
xii.	GP practice population social deprivation
xiii.	GP practice national target performance
xiv.	GP practice clinical staffing model

Table 1. Study outcomes

Statistical analysis

Baseline patient-level data in NWL will be extracted at the index date, defined as the date of first examination using the AI-stethoscope at each primary care practice. Data will be summarised as mean (SD) or median (interquartile range) for skewed data. Comparisons between groups will be undertaken on an intention-to-treat basis, using two-sample t-tests for continuous variables, and $\chi 2$ tests for binary and categorical variables. For each outcome, we will perform generalised mixed-effect logistic regression to compare the study arms, with the GP practice as a random effect.

Subgroup analyses for patient-level outcomes will be performed, stratified by age, sex, ethnicity and geography, patient comorbidities and GP practice characteristics (mean age of clinicians, size of clinical team, number of registered patients).

Patient-level sensitivity analyses will be performed for patients with abnormal AI-stethoscope predictions for HF, to identify direct associations between AI-stethoscope predictions and specific diagnostic codes for HF, AF and VHD.

We will measure relationships between items in the clinician questionnaire, utilisation rates of the AI-stethoscope, compliance with recommendations and the primary endpoints. These analyses will identify domains that are most strongly related to utilisation and compliance, which will help improve the technology, and guide future implementation and clinical pathways.

Interim Analysis

An interim analysis will be performed 6 months after the first site has been randomised. The interim analysis will report the following selected outcomes and will be reviewed by the Trial Steering Committee, who will remain blinded to the treatment allocation of each group until the study end:

Co-primary endpoints:

a. Difference in incidence of coded HF between groups;
b. Difference in ratio of coded HF incidence via hospital admission-based versus community-based diagnostic pathways between groups

Secondary endpoints

a. Healthcare diagnostics utilisation
b. Primary care appointments
c. Emergency department presentations
d. Non-elective hospital admissions

e. AI-stethoscope utilisation rates
f. Completion of site setup for all practices in intervention group
g. Data fidelity - ascertainment

The following interim outcomes will be considered significant and warranting intervention:

a. Mean AI-stethoscope utilisation rate of less than five recordings per month per practice in the intervention arm
b. Healthcare diagnostics utilisation differential affecting patient safety
c. Incomplete site setup in intervention group

In keeping with formative evaluations in implementation research[29-31] and behavioural change frameworks[46], adaptive interventions will be considered prospectively to maximise the success of the implementation of AI-stethoscope (given the established confidence for the technology "working" on the basis of widely validated statistical accuracy). In addition, pre-specified interventions will be available to the Steering Committee following the interim analysis.

- Semi-structured interviews of high and low-use participants
- Provision of repeat training in use of the AI-stethoscope
- Redistribution of the AI-stethoscope to new, prospectively enrolled GP practices
- Provision of latest generation of the AI-stethoscope hardware and software to intervention group
- Integration of the AI-stethoscope system with the electronic health record system

To mitigate lead-time effects, interim analysis of the co-primary endpoints will not incur a stopping rule at 6 months. If at 6 months there is no statistically significant difference between treatment arms for the co-primary endpoints, then a further interim review will be performed at 12 months. If there remains no clinically significant difference in the co-primary endpoint at 12 months, the Trial Steering Group will have the option to stop the trial on the basis of futility.

ETHICS AND DISSEMINATION

TRICORDER will not require written/signed consent from individual patients since the AI-stethoscope has full regulatory (MHRA) approval for use in direct clinical care, and will be used within its regulatory approved intended purpose. The NHS Integrated Care Systems involved in TRICORDER have developed and approved a clinical guideline for use of the AI-stethoscope in direct care, through agreement between executive primary and secondary care stakeholders (Supplementary Figure 1), which are anchored in best practice national guidelines from the UK National Institute for Health and Care Excellence (NICE). As with any clinical test, the decision to perform it and act on results in line with guidelines is ultimately the underpinned by the clinical judgement of the responsible medical professional, and involves informed consent of the patient. This protocol has

been reviewed and received favourable opinion from a UK Health Research Authority (HRA) Research Ethics Committee (reference: 23/LO/0051), including its cluster design and confirmation that patient-level research consent is not indicated. For use of Discover-NOW, there is existing UK HRA approval for access procedures, pseudonymisation and use of this platform for outcomes research (reference:18/WM/0323). A scientific manuscript with the primary outcomes of the study will be published in a peer-reviewed journal. Further manuscripts reporting secondary outcomes will also be published in peer-reviewed journals. Results will be presented at scientific meetings, and conferences with local and national stakeholders, including the national funder (National Institute for Health and Care Research), NHS organisations, and patient and public groups such as the Pumping Marvellous Foundation.

Trial registration

TRICORDER is registered with the NIH National Library of Medicine (NCT05987670). The study protocol has been reported in accordance with the SPIRIT-AI Extension[47] (Supplementary Table 1) and outcomes will be reported in accordance with the CONSORT-AI Extension[48].

Discussion

TRICORDER is the first cluster-randomised implementation trial to evaluate a real-time, AI-dependent point-of-care diagnostic in primary care. The study is designed to meet the specifications of the NICE Evidence Standards Framework for Digital Health Technologies[49] and therefore underpin recommendations for subsequent NHS-wide commissioning. This protocol – and lessons from execution of the trial – will inform a replicable blueprint for evaluation of other similar digital health technologies.

The study is designed to address the unacceptable reality that cardiovascular disease, heart failure particularly, is most frequently detected at a late stage, after disease progression precipitates a hospital admission[17,18]. The NHS Long Term plan prioritises the need to reverse this trend, emphasising the need for increased rates of diagnosis through community (primary care initiated) pathways. [1]. We have recently highlighted the £2,500 saving unlocked by every patient diagnosed with heart failure through such community pathways[18]. Similar models inform a compelling health economic case for community-based detection of AF and VHD.[19,20]

By spanning primary care practices across urban and rural geographies and serving populations with varied sociodemographic and ethnic backgrounds, this study will encompass a uniquely representative patient population. This is intended to consider the well-established concerns around AI bias[50], the digital maturity for technology adoption[51], and to demonstrate the AI-stethoscope's suitability for deployment across diverse patient and clinician populations. The research team will collate qualitative data via clinician surveys, ensuring the AI-stethoscope's efficient integration into primary care, preservation of the traditional clinician-patient interaction and high-level usability of the device.

This study is both decentralised and principally uses "real-world" data; there is no lead/central study site or logistics/travel outside each GP practice. Each will contribute data equally, but without the need to complete any study-specific (cumbersome) data collection instruments. Instead, our use of Discover-NOW enables outcomes measurement with routinely recorded (real world) data. The platform's offer of comprehensive linked (primary and secondary care) clinical and cost data affords patient-level sensitivity analyses. This will allow robust measurement of associations between diagnostic outcomes and use of the AI-stethoscope. Separate to such novel elements of this study, TRICORDER has set several other important precedents for translational AI research, including approval of the first regional NHS clinical guideline for use of a primary care focused AI technology; and sector-wide data governance approval covering use of an AI technology across hundreds of disparate primary care sites.

This research protocol is best interpreted in the context of its limitations. The pragmatic design, which aims to have a minimal impact on GP workflows, may in some cases limit sustained use of the technology by not setting any expectations or requirement for use – though this may somewhat mitigate any Hawthorne effect[52]. This study is the first to accompany a novel AI technology with an NHS sector-approved clinical guideline for use, but variable adherence may limit impact attributable to the technology. This will be addressed systematically, taking an implementation science approach to maximise uptake of the intervention. Finally, the examination of real-world data is universally limited by the inconsistency and variable fidelity of medical coding in capturing specific variables of interest. For example, for HF, this is rarely coded with a granularity that describes preserved, moderately reduced, or reduced ejection fraction. However, the otherwise comprehensive dataflows associated with this study will allow holistic scrutiny of a broad selection of outcomes that need to be understood to underpin any recommendations for system-wide uptake. Ultimately, this will serve to provide a realistic estimate of the potential clinical and health economic impact of the AI-stethoscope technology in NHS primary care.

AUTHORS CONTRIBUTIONS

MAK, PB and NSP were involved in conception and trial design. CFP, CC and EF provided statistical expertise. RAL provided expertise on patient and public involvement. MAK and PB were involved in drafting this manuscript. JM, DBK, MTA, AP, AA, CB, AM, CMP and NSP provided critical revision of the manuscript and intellectual content. All authors were involved in the final approval of the article. MAK and PB contributed equally and are co-first authors of the work. NSP is the guarantor.

ACKNOWLEDGEMENTS

The authors are grateful for the infrastructural support from the National Institute of Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. AM is supported by the NIHR Applied Research Collaboration NW London. The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. We are indebted to the co-development with and feedback from patients from the Pumping Marvellous Foundation heart failure charity.

FUNDING

This work was supported by the National Institute for Health and Care Research Invention for Innovation (i4i) Challenge Award, Accelerated Access Collaborative and NHSX via the AI Award in Health and Social Care; Imperial Health Charity; British Heart Foundation.

Funders were not involved in the study design, writing of the report or decision to submit the article for publication. All authors had full access to the protocol and take responsibility for the integrity of the manuscript.

COMPETING INTERESTS

M.A.K. has nothing to declare; P.B. has nothing to declare; J.M. has nothing to declare; C.F.P. has nothing to declare; M.T.A. has nothing to declare; A.A. has nothing to declare; D.B.K. has nothing to declare; A.P.. has nothing to declare; C.C. has nothing to declare; E.F. has nothing to declare; R.A.L. has nothing to declare; A.M. has nothing to declare; C.M.P. has nothing to declare; N.S.P. is an advisor to Eko Health Inc.

REFERENCES

1. NHS. NHS Long Term Plan. 2019 (https://www.longtermplan.nhs.uk/).
2. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 2020;396(10258):1204-1222. DOI: 10.1016/S0140-6736(20)30925-9.
3. Organisation WH. Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. World Health Organisation, 2020.
4. Kannel WB, Abbott RD, Savage DD, McNamara PM. Epidemiologic Features of Chronic Atrial Fibrillation. New England Journal of Medicine 1982;306(17):1018-1022. DOI: 10.1056/nejm198204293061703.
5. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA. Independent Risk Factors for Atrial Fibrillation in a Population-Based Cohort: The Framingham Heart Study. JAMA 1994;271(11):840-844. DOI: 10.1001/jama.1994.03510350050036.
6. Krum H, Gilbert RE. Demographics and concomitant disorders in heart failure. The Lancet 2003;362(9378):147-158. DOI: https://doi.org/10.1016/S0140-6736(03)13869-X.
7. EHRLICH JR, NATTEL S, HOHNLOSER SH. Atrial Fibrillation and Congestive Heart Failure: Specific Considerations at the Intersection of Two Common and Important Cardiac Disease Sets. Journal of Cardiovascular Electrophysiology 2002;13(4):399-405. DOI: https://doi.org/10.1046/j.1540-8167.2002.00399.x.
8. Meyer S, Brouwers FP, Voors AA, et al. Sex differences in new-onset heart failure. Clin Res Cardiol 2015;104(4):342-50. (In eng). DOI: 10.1007/s00392-014-0788-x.
9. Brouwers FP, de Boer RA, van der Harst P, et al. Incidence and epidemiology of new onset heart failure with preserved vs. reduced ejection fraction in a community-based cohort: 11-year follow-up of PREVEND. Eur Heart J 2013;34(19):1424-31. (In eng). DOI: 10.1093/eurheartj/eht066.
10. Conrad N, Judge A, Tran J, et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. Lancet (London, England) 2018;391(10120):572-580. (In eng). DOI: 10.1016/s0140-6736(17)32520-5.
11. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet (London, England) 2018;392(10159):1789-1858. (In eng). DOI: 10.1016/s0140-6736(18)32279-7.
12. Roger VL. Epidemiology of heart failure. Circ Res 2013;113(6):646-59. (In eng). DOI: 10.1161/circresaha.113.300268.
13. Mosterd A, Hoes AW. Clinical epidemiology of heart failure. Heart 2007;93(9):1137-46. (In eng). DOI: 10.1136/hrt.2003.025270.

14. Mamas MA, Sperrin M, Watson MC, et al. Do patients have worse outcomes in heart failure than in cancer? A primary care-based cohort study with 10-year follow-up in Scotland. European Journal of Heart Failure 2017;19(9):1095-1104. (https://doi.org/10.1002/ejhf.822). DOI: https://doi.org/10.1002/ejhf.822.
15. Braunschweig F, Cowie MR, Auricchio A. What are the costs of heart failure? Europace 2011;13 Suppl 2:ii13-7. (In eng). DOI: 10.1093/europace/eur081.
16. Excellence NIfHaC. Resource impact report: Chronic heart failure in adults: diagnosis and management (NG106). https://www.nice.org.uk/guidance/ng106/resources/resource-impact-report-pdf-6537494413#:~:text=2.3%20It%20is%20estimated%20that,into%20Living%20with%20Heart%20failure: National Institute for Health and Care Excellence, 2018.
17. Bottle A, Kim D, Aylin P, Cowie MR, Majeed A, Hayhoe B. Routes to diagnosis of heart failure: observational study using linked data in England. Heart 2018;104(7):600-605. (In eng). DOI: 10.1136/heartjnl-2017-312183.
18. Bachtiger P, Kelshiker MA, Petri CF, et al. Survival and health economic outcomes in heart failure diagnosed at hospital admission versus community settings: a propensity-matched analysis. BMJ Health & Care Informatics 2023;30(1):e100718. DOI: 10.1136/bmjhci-2022-100718.
19. Burdett P, Lip GYH. Atrial fibrillation in the UK: predicting costs of an emerging epidemic recognizing and forecasting the cost drivers of atrial fibrillation-related costs. European Heart Journal - Quality of Care and Clinical Outcomes 2020;8(2):187-194. DOI: 10.1093/ehjqcco/qcaa093.
20. Moore M, Chen J, Mallow PJ, Rizzo JA. The direct health-care burden of valvular heart disease: evidence from US national survey data. Clinicoecon Outcomes Res 2016;8:613-627. (In eng). DOI: 10.2147/ceor.S112691.
21. McDonagh TA, Metra M, Adamo M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2021;42(36):3599-3726. (In eng). DOI: 10.1093/eurheartj/ehab368.
22. Hindricks G, Potpara T, Dagres N, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. European Heart Journal 2021;42(5):373-498. DOI: 10.1093/eurheartj/ehaa612.
23. Vahanian A, Beyersdorf F, Praz F, et al. 2021 ESC/EACTS Guidelines for the management of valvular heart disease: Developed by the Task Force for the management of valvular heart disease of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS). European Heart Journal 2021;43(7):561-632. DOI: 10.1093/eurheartj/ehab395.
24. Fonseca C, Morais H, Mota T, et al. The diagnosis of heart failure in primary care: value of symptoms and signs. Eur J Heart Fail 2004;6(6):795-800, 821-2. (In eng). DOI: 10.1016/j.ejheart.2004.08.002.

25. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2013;62(16):e147-239. (In eng). DOI: 10.1016/j.jacc.2013.05.019.
26. Bachtiger P, Petri CF, Scott FE, et al. Point-of-care screening for heart failure with reduced ejection fraction using artificial intelligence during ECG-enabled stethoscope examination in London, UK: a prospective, observational, multicentre study. Lancet Digit Health 2022 (In eng). DOI: 10.1016/s2589-7500(21)00256-9.
27. Davidsen AH, Andersen S, Halvorsen PA, Schirmer H, Reierth E, Melbye H. Diagnostic accuracy of heart auscultation for detecting valve disease: a systematic review. BMJ Open 2023;13(3):e068121. (In eng). DOI: 10.1136/bmjopen-2022-068121.
28. Chorba JS, Shapiro AM, Le L, et al. Deep Learning Algorithm for Automated Cardiac Murmur Detection via a Digital Stethoscope Platform. Journal of the American Heart Association 2021;10(9):e019905. DOI: 10.1161/JAHA.120.019905.
29. Bauer MS, Kirchner J. Implementation science: What is it and why should I care? Psychiatry Res 2020;283:112376. (In eng). DOI: 10.1016/j.psychres.2019.04.025.
30. Elwy AR, Wasan AD, Gillman AG, et al. Using formative evaluation methods to improve clinical implementation efforts: Description and an example. Psychiatry Res 2020;283:112532. (In eng). DOI: 10.1016/j.psychres.2019.112532.
31. Margaret AH, Anuradha G, Adithya C. Strategies for implementing implementation science: a methodological overview. Emergency Medicine Journal 2016;33(9):660. DOI: 10.1136/emermed-2015-205461.
32. Parikh RB, Teeple S, Navathe AS. Addressing Bias in Artificial Intelligence in Health Care. Jama 2019;322(24):2377-2378. (In eng). DOI: 10.1001/jama.2019.18058.
33. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019;28(3):231-237. (In eng). DOI: 10.1136/bmjqs-2018-008370.
34. Digital N. Patients Registered at a GP Practice. March 2023 2023.
35. Excellence NIfHaC. Chronic heart failure in adults: diagnosis and management. National Institute for Health and Care Excellence, 2018. (NG106) (https://www.nice.org.uk/guidance/ng106/resources/chronic-heart-failure-in-adults-diagnosis-and-management-pdf-66141541311685).
36. Cowie MR. The heart failure epidemic: a UK perspective. Echo Res Pract 2017;4(1):R15-r20. (In eng). DOI: 10.1530/erp-16-0043.
37. Bottle A, Cohen C, Lucas A, et al. How an electronic health record became a real-world research resource: comparison between London's Whole Systems Integrated Care database and the Clinical Practice Research Datalink. BMC Med Inform Decis Mak 2020;20(1):71. (In eng). DOI: 10.1186/s12911-020-1082-7.
38. Yao X, Rushlow DR, Inselman JW, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. Nat Med 2021;27(5):815-819. DOI: 10.1038/s41591-021-01335-4.

39. Kahn M, Grayson AD, Chaggar PS, et al. Primary care heart failure service identifies a missed cohort of heart failure patients with reduced ejection fraction. Eur Heart J 2022;43(5):405-412. (In eng). DOI: 10.1093/eurheartj/ehab629.
40. Ltd SE. Simple randomisation service. https://www.sealedenvelope.com/simple-randomiser/v1/. Sealed Envelope Ltd; 2022.
41. Council GM. Good Medical Practice. General Medical Council; 2013.
42. Brooke J. SUS: A quick and dirty usability scale. Usability Eval Ind 1995;189.
43. Bottle A, Cohen C, Lucas A, et al. How an electronic health record became a real-world research resource: comparison between London's Whole Systems Integrated Care database and the Clinical Practice Research Datalink. BMC Medical Informatics and Decision Making 2020;20(1):71. DOI: 10.1186/s12911-020-1082-7.
44. Greene SJ, O'Brien EC, Mentz RJ, et al. Home-Time After Discharge Among Patients Hospitalized With Heart Failure. J Am Coll Cardiol 2018;71(23):2643-2652. (In eng). DOI: 10.1016/j.jacc.2018.03.517.
45. Lam MB, Riley KE, Zheng J, Orav EJ, Jha AK, Burke LG. Healthy days at home: A population-based quality measure for cancer patients at the end of life. Cancer 2021;127(22):4249-4257. (In eng). DOI: 10.1002/cncr.33817.
46. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implementation Science 2011;6(1):42. DOI: 10.1186/1748-5908-6-42.
47. Rivera SC, Liu X, Chan A-W, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. BMJ 2020;370:m3210. DOI: 10.1136/bmj.m3210.
48. Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nature Medicine 2020;26(9):1364-1374. DOI: 10.1038/s41591-020-1034-x.
49. Unsworth H, Dillon B, Collinson L, et al. The NICE Evidence Standards Framework for digital health and care technologies – Developing and maintaining an innovative evidence framework with global impact. DIGITAL HEALTH 2021;7:205520762110186. DOI: 10.1177/20552076211018617.
50. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366(6464):447-453. DOI: doi:10.1126/science.aax2342.
51. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017;19(11):e367. (In eng). DOI: 10.2196/jmir.8775.
52. Chiesa M, Hobbs S. Making sense of social research: how useful is the Hawthorne Effect? European Journal of Social Psychology 2008;38(1):67-74. DOI: https://doi.org/10.1002/ejsp.401.

Figure Legends

Figure 1. NICE diagnostic pathway for HF. *There are substantial delays to confirmatory echocardiographic testing and specialist review for those are referred for suspected HF through this pathway, such that only 10% of patients complete the pathway to time and target.*

Figure 2: *Map of London, with the boroughs of North West London stratified by Index of Multiple Deprivation deciles*

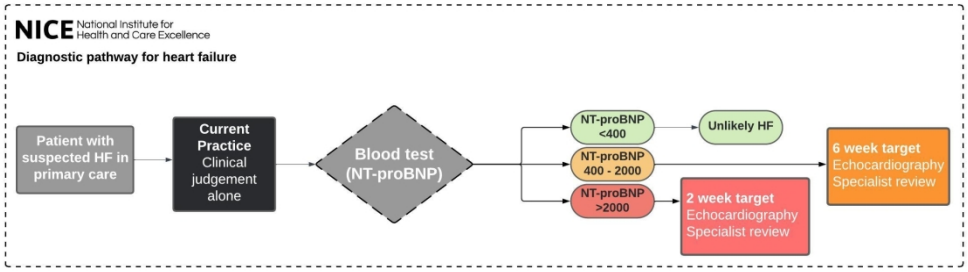


Figure 1. NICE diagnostic pathway for HF. There are substantial delays to confirmatory echocardiographic testing and specialist review for those are referred for suspected HF through this pathway, such that only 10% of patients complete the pathway to time and target.

381x114mm (160 x 160 DPI)

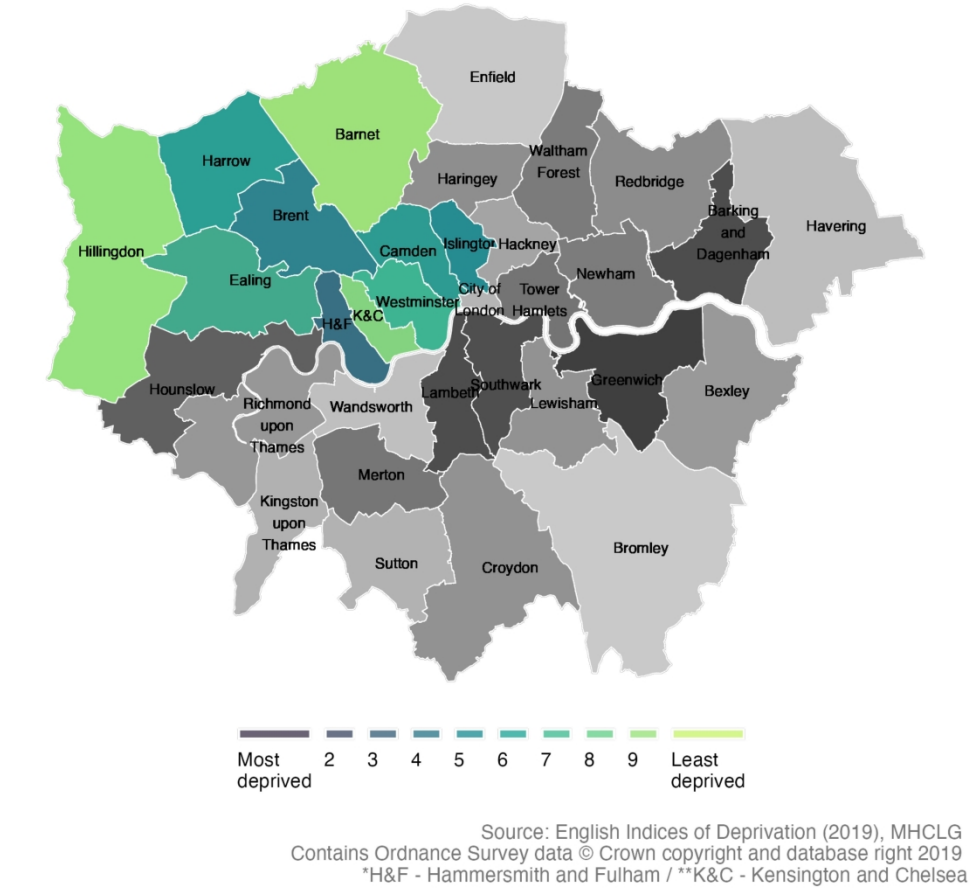
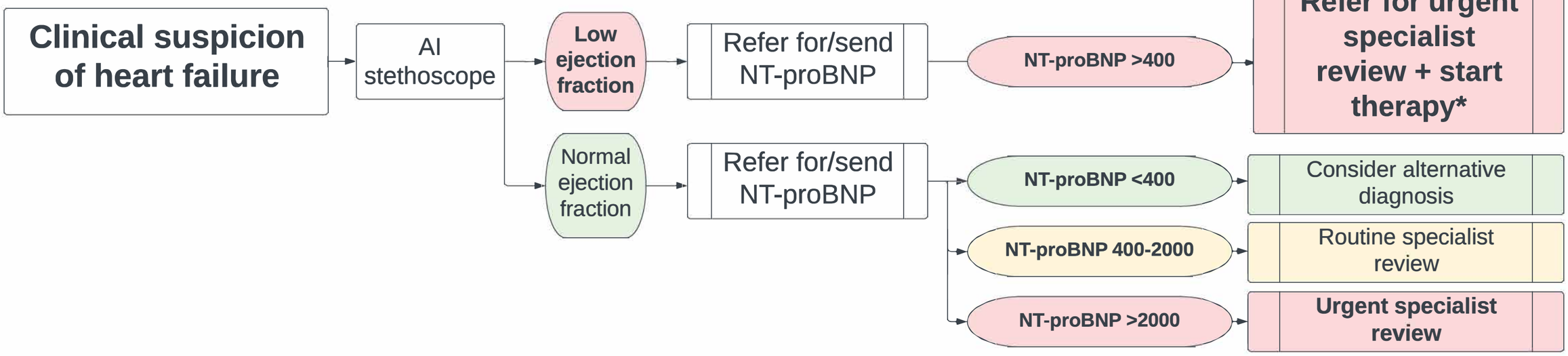


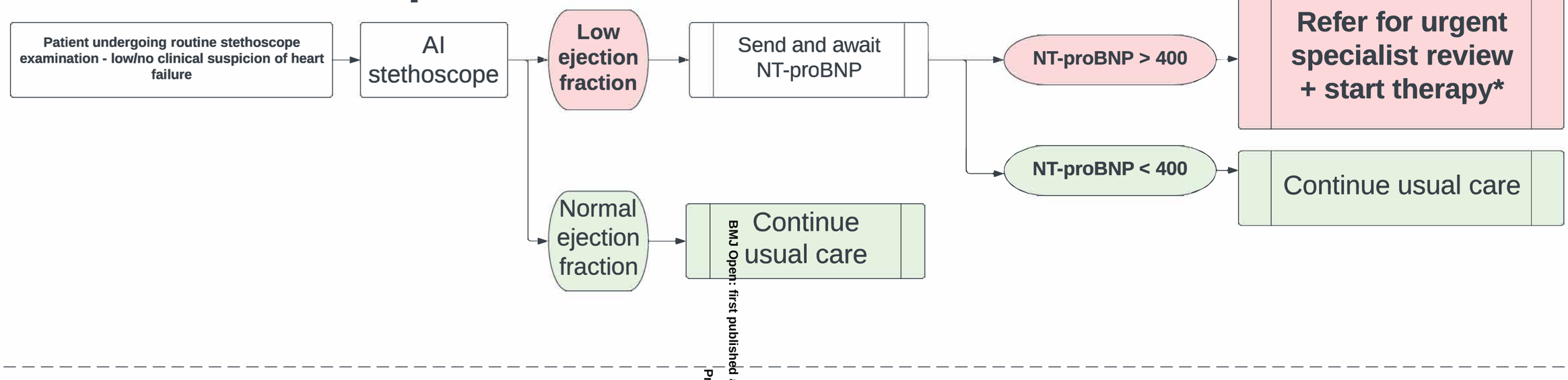
Figure 2: Map of London, with the boroughs of North West London stratified by Index of Multiple Deprivation deciles

352x364mm (300 x 300 DPI)

Clinical suspicion of heart failure



Low/no clinical suspicion of heart failure

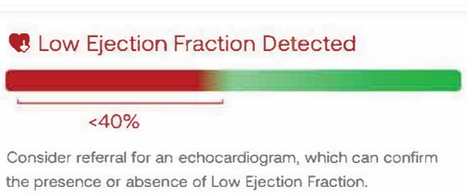


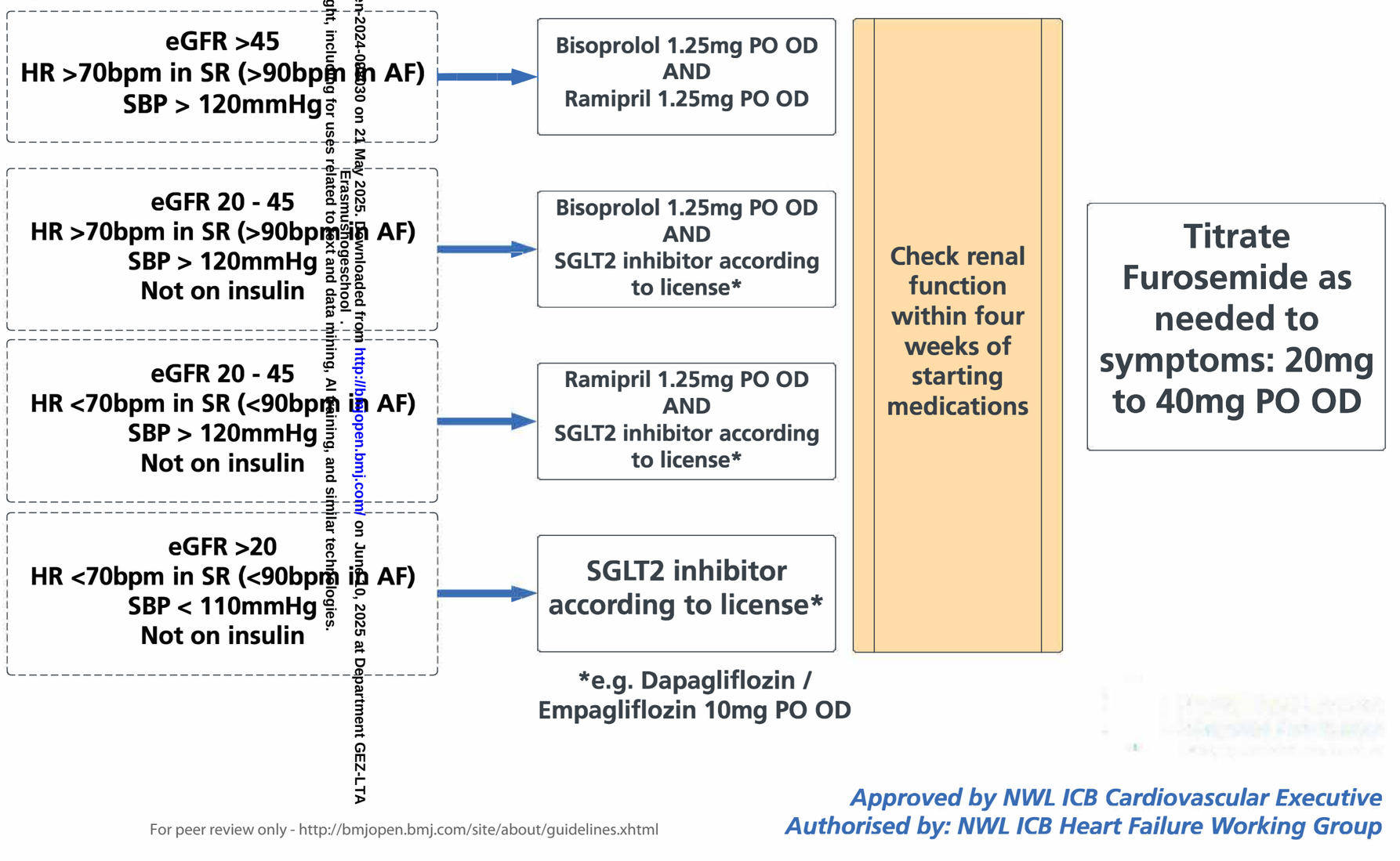
*Initiating Heart Failure prognostic therapy in Primary Care with AI stethoscope

NT-proBNP > 400

AND

AI stethoscope





Approved by NWL ICB Cardiovascular Executive
Authorised by: NWL ICB Heart Failure Working Group

SPIRIT-AI

SPIRIT-AI Checklist: Recommended items to address in a protocol and related documents for clinical trials evaluating AI interventions

Section		SPIRIT 2013 Item[a]	SPIRIT-AI Item		Addressed on Page No [b]
Administrative Information					
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	SPIRIT-AI 1(i) Elaboration	Indicate that the intervention involves artificial intelligence / machine learning and specify the type of model.	1
			SPIRIT-AI 1(ii) Elaboration	Specify the intended use of the AI intervention.	1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry			5
	2b	All items from the World Health Organization Trial Registration Data Set			Throughout protocol
Protocol version	3	Date and version identifier			1
Funding	4	Sources and types of financial, material, and other support			4
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors			2
	5b	Name and contact information for the trial sponsor			4
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities			4
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)			4
Introduction					
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	SPIRIT-AI 6a (i) Extension	Explain the intended use of the AI intervention in the context of the clinical pathway, including its purpose and its intended users (e.g., healthcare professionals, patients, public).	9, 10
			SPIRIT-AI 6a (ii) Extension	Describe any pre-existing evidence for the AI intervention.	10

SPIRIT-AI

	6b	Explanation for choice of comparators			10
Objectives	7	Specific objectives or hypotheses			11
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)			11
Methods: Participants, Interventions and Outcomes					
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	SPIRIT-AI 9 Extension	Describe the onsite and offsite requirements needed to integrate the AI intervention into the trial setting.	11
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	SPIRIT-AI 10 (i) Elaboration	State the inclusion/exclusion criteria at the level of participants.	16
			SPIRIT-AI 10 (ii) Extension	State the inclusion/exclusion criteria at the level of the input data.	17
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	SPIRIT-AI 11a (i) Extension	State which version of the AI algorithm will be used.	20
			SPIRIT-AI 11a (ii) Extension	Specify the procedure for acquiring and selecting the input data for the AI intervention.	19, 20
			SPIRIT-AI 11a (iii) Extension	Specify the procedure for assessing and handling poor quality or unavailable input data.	17, 20
			SPIRIT-AI 11a (iv) Extension	Specify whether there is human-AI interaction in the handling of the input data, and what level of expertise is required for use.	19, 20
			SPIRIT-AI 11a (v) Extension	Specify the output of the AI intervention.	21
			SPIRIT-AI 11a (vi) Extension	Explain the procedure for how the AI intervention's output will contribute to decision-making or other elements of clinical practice.	20
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)			23
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)			23
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial			23
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended			9, 16

SPIRIT-AI

Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)			12
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations			11
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size			11
Methods: Assignment of Interventions (For Controlled Trials)					
Sequence generation	16A	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions			19
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned			19
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions			19
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how			23
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial			24
Methods: Data Collection, Management, And Analysis					
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol			22
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols			222
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol			2

SPIRIT-AI

Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol			22
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)			22
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)			22
Methods: Monitoring					
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed			24
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial			23
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	SPIRIT-AI 22 Extension	Specify any plans to identify and analyse performance errors. If there are no plans for this, explain why not.	23
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor			23
Ethics and Dissemination					
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval			1
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)			4
Consent or ascent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)			19
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable			Not applicable
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial			22, 24
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site			25

SPIRIT-AI

Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	SPIRIT-AI 29 Extension	State whether and how the AI intervention and/or its code can be accessed, including any restrictions to access or re-use.	24
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation			Not applicable
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions			25, 26
	31b	Authorship eligibility guidelines and any intended use of professional writers			25
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code			Not applicable
Appendices					
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates			Appendix
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable			Not applicable

[a] It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items.
[b] Indicates page numbers to be completed by authors during protocol development.