PROTEUS Study: A Prospective Randomized Controlled Trial Evaluating the Use of Artificial Intelligence in Stress Echocardiography

Gary Woodward, PhD [1], Mamta Bajre, MBBS [2], Sanjeev Bhattacharyya, TSC [3], Maria Breen, PhD [1], Virginia Chiocchia, DSMC [4], Helen Dawes, PhD [5], Hakim-Moulay Dehbi, PhD [6], Tine Descamps, PhD [1], Elena Frangou, TSC [7], Carol-Ann Fazakarley, BSc [1], Victoria Harris, DSMC [8], Will Hawkes, PhD [1], Oliver Hewer, MA [1], Casey L Johnson, MSc [9], Samuel Krasner, BA [9], Lynn Laidlaw, TSC , PPI chair [10], Jonathan Lau [1], Tom Marwick, TSC chair [11], Steffen E Petersen, DSMC [3,12], Hania Piotrowska, BA [1], Ged Ridgeway, TSC [13], David P Ripley, TSC [14], Emily Sanderson, MSc [1], Natalie Savage, PhD [9], Rizwan Sarwar, MRCP [1], Louise Tetlow, PhD [1], Ben Thompson, PhD [1], Samantha Thulborn, BSc [1], Victoria Williamson, PhD [1], William Woodward, BSc [9], Ross Upton, PhD [1], and Paul Leeson, PhD [9] *Oxford, UK; London, UK; Bern, Switzerland*

Background Stress echocardiography (SE) is one of the most commonly used diagnostic imaging tests for coronary artery disease (CAD) but requires clinicians to visually assess scans to identify patients who may benefit from invasive investigation and treatment. EchoGo Pro provides an automated interpretation of SE based on artificial intelligence (AI) image analysis. In reader studies, use of EchoGo Pro when making clinical decisions improves diagnostic accuracy and confidence. Prospective evaluation in real world practice is now important to understand the impact of EchoGo Pro on the patient pathway and outcome.

Methods PROTEUS is a randomized, multicenter, 2-armed, noninferiority study aiming to recruit 2,500 participants from National Health Service (NHS) hospitals in the UK referred to SE clinics for investigation of suspected CAD. All participants will undergo a stress echocardiogram protocol as per local hospital policy. Participants will be randomized 1:1 to a control group, representing current practice, or an intervention group, in which clinicians will receive an AI image analysis report (EchoGo Pro, Ultromics Ltd, Oxford, UK) to use during image interpretation, indicating the likelihood of severe CAD. The primary outcome will be appropriateness of clinician decision to refer for coronary angiography. Secondary outcomes will assess other health impacts including appropriate use of other clinical management approaches, impact on variability in decision making, patient and clinician qualitative experience and a health economic analysis.

Discussion This will be the first study to assess the impact of introducing an AI medical diagnostic aid into the standard care pathway of patients with suspected CAD being investigated with SE.

Trial Registration Clinicaltrials.gov registration number NCT05028179, registered on 31 August 2021; ISRCTN: ISRCTN15113915; IRAS ref: 293515; REC ref: 21/NW/0199. (Am Heart J 2023;263:123–132.)

From the [1]Ultromics Ltd., Oxford, UK, [2]Oxford Academic Health Science Network, Magdalen Centre—Whitehead Building, Oxford, UK, [3]Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, West Smithfield, London, UK, [4]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland, [5]Oxford Clinical Allied Technology and Trial Services Unit, Oxford Brookes University, Oxford, UK, [6]UCL Comprehensive Clinical Trials Unit, University College London, London, UK, [7]MRC Clinical Trials Unit at UCL, London, UK, [8]Department of Primary Health Care Sciences, University of Oxford, Oxford, UK, [9]Cardiovascular Clinical Research Facility, RDM Division of Cardiovascular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK, [10]British Society of Echocardiography. (BSE) Wavelength PPI group member, NIHR and BMJ lay reviewer, [11]Baker Heart and Diabetes Institute, Melbourne Victoria, Australia, [12]William Harvey Research Institute, NIHR Barts Biomedical Research Centre, Queen Mary University London, Charterhouse Square, London, UK, [13]Oxford Brain Diagnostics Ltd, Oxford Centre for Innovation, Oxford, UK, [14]School of Medicine, Faculty of Health Sciences and Wellbeing, University of Sunderland, London, UK

Abbreviations: AAC, Accelerated Access Collaborative; AI, Artificial Intelligence; AUROC, Area Under Receiver Operator Characteristic; CAD, Coronary Artery Disease; CMR, Cardiac Magnetic Resonance; CTCA, Computed Tomography Coronary Angiography; DSMC, Data Safety Monitoring Committee; eCRF, Electronic Case Report Form; EDC, Electronic Data Capture; EQ-5D-5L, EuroQol 5 dimensions health questionnaire (5 question length); GLMM, Generalised Logistic Mixed Methodology; HRA, Health Research Authority; ICF, Informed Consent Form; IEP, Image Exchange Portal; MAR, Missing at Random; MNAR, Missing Not at Random; MPS, Myocardial Perfusion Scintigraphy; NHS, National Health Service; NIHR, National Institute for Health and Care Research; PACS, Picture Archiving and Communication System; PIL, Patient Information Leaflet; REC, Research Ethics Committee; ROC, Receiver Operator Characteristic; SAQ-7, Seattle Angina Questionnaire (short); SE, Stress Echocardiography; TSC, Trial Steering Committee; VPN, Virtual Private Network.

Submitted November 9, 2022; accepted May 4, 2023

Reprint requests: *Paul Leeson, Cardiovascular Clinical Research Facility, RDM Division of Cardiovascular Medicine, University of Oxford, Level 1, John Radcliffe Hospital, Oxford OX3 9DU, UK.*

E-mail address: paul.leeson@cardiov.ox.ac.uk.

0002-8703

Background and rationale

Coronary artery disease (CAD) is responsible for over 9.1 million deaths worldwide, and 63,000 deaths annually in the UK.[1] As the leading cause of premature death, it is thought to cost the UK economy in excess of £19 billion annually.[2] Patients with CAD typically present with chest pain or breathlessness on exertion. However, a large proportion of patients with these symptoms do not have CAD, and to identify patients who require medication or invasive treatment further investigations are required. Recent guidelines have consolidated noninvasive imaging as the standard first line investigation for CAD[3]; including stress echocardiography (SE), myocardial perfusion scintigraphy (MPS), stress cardiac magnetic resonance (CMR), or computed tomography coronary angiography (CTCA), dependent upon local resource availability and expertise. SE has been widely available for the last 20 years, and is performed in 3 quarters of National Health Service (NHS) hospitals across the UK,[4] making stress echocardiography the most widely available imaging modality for diagnosis of CAD.[5] However, whilst SE has high clinical accuracy in expert hands, the reliability of the diagnosis relies on subjective clinical assessment and is therefore highly dependent on operator performance and experience. Estimates of the diagnostic performance of stress echocardiograms vary significantly, with an accuracy ranging from 60% to 94%.[5-9] As a result, the development of reproducible, accurate, and easy to use quantitative analysis techniques, which can be routinely applied in stress echocardiography has been a major focus of investigation.

In recent years, artificial intelligence (AI) has been applied to cardiovascular imaging in order to support diagnosis either through automation of measurements or provision of disease classification.[10] By providing quicker and less variable outputs,[11] AI may reduce cost and improve patient outcomes. We have recently reported fully automated AI interpretation of SE is feasible and, in a reader study, improved clinician diagnostic accuracy and confidence.[12] We now report the protocol of a prospective, randomized, controlled multicenter trial to evaluate the impact that introduction of this AI classifier into the CAD pathway has on the patient pathway and outcomes in the UK. We will collect real-world data from multiple hospitals across the UK on the clinical management decisions made for patients, results from further investigations, cardiac events that occur during follow-up, variation in stress echocardiogram reporting, clinician diagnostic confidence and, from this data, also evaluate health economic impact. To our knowledge, there has been no similar scale, randomized trial, evaluating how AI influences medical decision making and patient outcome. We expect the trial design to provide a framework for researchers who aim to validate future AI diagnostic tools.

Objectives and outcomes

The primary aim of this study is to investigate if the intervention (EchoGo Pro) plus standard care is noninferior to standard care alone in aiding for referral to coronary angiogram following stress echocardiogram. The primary outcome is based on the clinical decision to refer for angiography which will be evaluated based on review of participant records at 6-months. An adjudication committee comprising of at least one accredited cardiologist independent of the trial will review all clinical data. A correct decision to refer for coronary angiography will be confirmed based on the results of an elective coronary angiogram. The adjudication committee will decide there was a correct decision if there is evidence of disease on angiography which would warrant intervention either by percutaneous intervention or coronary artery bypass grafting. This includes presence of $\geq$70% stenosis in a proximal segment of a main coronary artery (left main stem, left anterior descending artery, right coronary artery or left circumflex artery) and/or a clinical decision to undertake intervention either by percutaneous coronary intervention or surgery. To address secondary outcomes, the adjudication committee will also consider clinical events during the follow-up period that suggest there may have been an inappropriate decision to manage the participant medically. The committee will review emergency admissions to hospital to determine whether these were related to cardiovascular events confirmed by electrocardiogram, troponin elevation, and emergency or urgent coronary angiography. If there are any deaths, death certification will be used as evidence of relatedness to cardiovascular disease. Secondary aims include investigating if the intervention plus standard care is superior to standard care alone in aiding for referral to coronary angiogram following stress echocardiogram, and if appropriate clinical management decisions are made when using the intervention plus standard care compared to standard care alone. Further secondary aims include a health economic impact assessment, and establishing the effect of the intervention on the number of unanticipated serious cardiac events and subsequent investigations, clinical management decisions, clinician diagnostic confidence and performance, diagnostic performance variation, and patient-reported cardiac symptoms (Table I). All outcomes will be measured at the conclusion of the study follow-up period.

Hypothesis

The primary null hypothesis is that the intervention is inferior to the comparator (standard care), with the

https://doi.org/10.1016/j.ahj.2023.05.003

Table I. PROTEUS trial objectives and outcome measures

Objective	Outcome measure
Primary	
• Investigate if the intervention plus standard care is non-inferior to standard care alone for aiding appropriate referral to coronary angiogram following stress echocardiogram.	-AUROC for the ability to make an appropriate referral to coronary angiogram.
Secondary	
• Investigate if the intervention plus standard care is superior to standard care alone for aiding referral to coronary angiogram following stress echocardiogram.	-AUROC for the ability to make an appropriate referral to coronary angiogram.
• Investigate the appropriate clinical management decisions made when using the intervention plus standard care compared standard care alone.	-AUROC for appropriate clinical decision for coronary angiography or medical management.
• Establish if deploying the intervention into NHS sites affects the number of unanticipated serious cardiac events.	-Number of acute cardiac events not related to elective cardiac procedures.
• Establish if clinical management decision is affected by review of the intervention report.	-Change in clinical management decision.
• Establish if using the intervention affects clinician's diagnostic confidence.	-Clinician diagnostic confidence in their interpretation of stress echocardiogram report, self-reported.
• Establish if clinician diagnostic performance variance is reduced with use of the intervention.	- Diagnostic performance variance using both inter-clinician and in-site stress echocardiogram interpretation accuracy.
• Establish if using the intervention affects the number of subsequent investigations for cardiovascular disease.	-Number of investigations following stress echo.
• Establish if using the intervention affects participant reported CAD symptoms.	-Participant reported symptoms.
• Assess the health economic impact of implementation and use of the intervention in NHS units.	-Participant reported health related quality of life.

difference in area under the receiver operating characteristic (AUROC) between comparator and intervention greater or equal to the noninferiority margin of 0.05 (Comparator-Intervention $\geq$ 0.05).

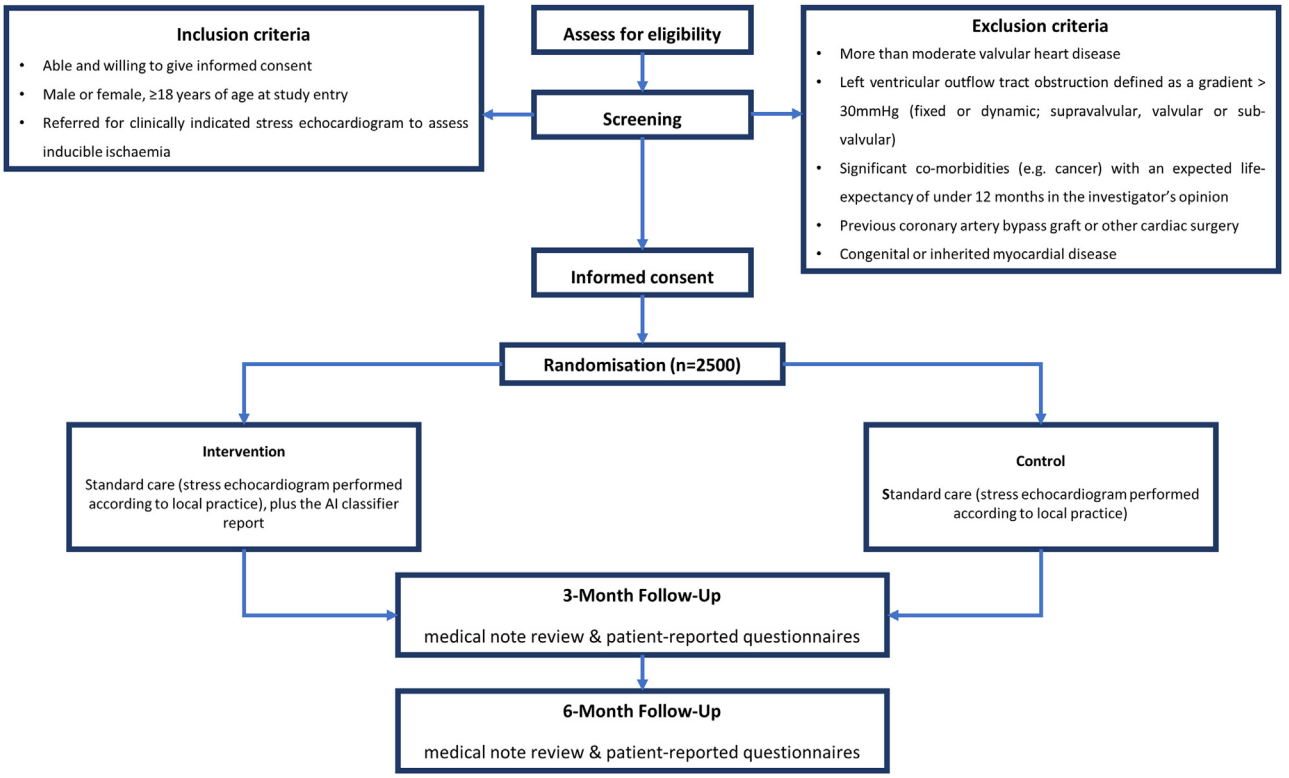
Methods

PROTEUS is a randomized, multicenter, 2-armed, noninferiority study aiming to recruit 2,500 participants from up to 20 NHS hospitals in the UK referred to SE clinics for investigation of suspected CAD. All participants will undergo a stress echocardiogram protocol as per local hospital policy. Participants will be randomised 1:1 to a control group, representing current practice, or an intervention group, in which clinicians will receive an AI image analysis report (EchoGo Pro, Ultromics Ltd, Oxford, UK) to use during image interpretation, indicating the likelihood of severe CAD. There are two remote follow-up visits at 3 and 6-months postrandomisation consisting of both a medical note review (conducted by the study team) and 2 patient-reported questionnaires (Figure 1). This protocol is prepared in accordance with the SPIRIT statement including the SPIRIT-AI extension.

Participants who are enrolled in the trial will normally complete 4 study visits over the course of the trial:

I. Screening assessment (Visit 0)
II. Baseline Study Visit (Visit 1)
III. 3-Month Follow-up (Visit 2) at 3-months postrandomization
IV. 6-Month Follow-up (Visit 3) at 6-months postrandomization

Figure 1



PROTEUS Study overview and visit schedule. Provides an overview of the study, describing the inclusion/exclusion criteria, randomization arms, and study visits.

AI integration

In order to integrate the AI into the study, a secure connection will need to be established between each recruiting site and Ultromics. This will allow for the transfer of stress echocardiogram images to Ultromics for processing, and the return of the AI analysis report to the site for review. Sites will be connected either via a secure virtual private network (VPN), or image exchange portal (IEP). This will require communication between site IT/Network teams and connectivity specialists at Ultromics. No software will be installed on site. Additionally, the picture archiving and communication system (PACS) at each site will be configured in order to allow data transfer between each site and Ultromics. For sites whose PACS is unable to be configured, configurations will be established on the ultrasound machines used for stress echocardiogram image acquisition. Stress echocardiogram images will then be exported directly from the ultrasound machine, and AI analysis reports will then be returned as a PDF imported into a secured, access-restricted folder located on a server at the site. Limited participant details will be included in the file name of the PDF for reidentification at the site.

Sample size

The study is powered to demonstrate there is no decrease in performance of stress echocardiography introduced by use of an AI classifier in clinical decision making. Parity between clinician and AI/clinician over 6 months (AUROC = 0.95) is therefore assumed based on previously obtained results utilising the intervention.[12] Using a kappa (ratio non-diseased/diseased) of 12, and assuming 80% power with alpha 0.05, a total sample size of 1,534 participants (767 per arm) is required under a noninferiority margin of 0.05. If the margin is reduced to 0.04 then a sample of 2,405 (1,203 participants per arm) is required. Sensitivity and specificity is assumed to be similar to that from the EVAREST study[13] for clinicians over 12 months (AUROC = 0.88) for the comparator. If the intervention maintains an AUROC of 0.95 over 12 months, as assumed under the 6 months noninferiority objective, then a sample size of 2,132 participants would be powered to show superiority. This increase in 10% improvement would correspond with the data previously observed.[12] Therefore, we aim to recruit 2,500 participants, allowing for a predicted recruitment drop-out rate of 5% to 10%. Under this sample size the trial will be

powered for both primary and secondary aims, including a reduced noninferiority margin of 0.04. Interim analysis will be used to update any assumptions where necessary.

Eligibility criteria

Patients are eligible for the trial if:

(1) Age $\geq$ 18 years
(2) Undergoing clinically indicated stress echocardiogram to assess inducible ischemia
(3) Able and willing to give informed consent

Patients are not eligible for the study if any of the following are met:

(1) Patient has more than moderate valvular heart disease
(2) Patient has a left ventricular outflow tract obstruction defined as a gradient > 30mmHg (fixed or dynamic; supravalvular, valvular or subvalvular)
(3) Patient has significant co-morbidities (eg, cancer) with an expected life-expectancy of under 12 months in the investigator's opinion
(4) Patient has had a previous coronary artery bypass graft or other cardiac surgery
(5) Patient has congenital or inherited myocardial disease
(6) Decision to use pacemaker or vasodilator stress for the test

Recruitment

The clinical care team at participating NHS sites will be asked to review their medical records and identify potential participants. Potential participants will be attending their appointment as part of their routine clinical care, and thus no additional strategies to recruit participants will be necessary. Potential participants will be invited to join the study by means of an invitation letter and participant information leaflet (PIL) from their standard care team. The PIL will detail no less than: the exact nature of the study; what it will involve for the participant; the implications and constraints of the protocol; the known side effects and any risks involved in taking part. It will be clearly stated that the participant is free to withdraw from the study at any time for any reason without prejudice to future care, and with no obligation to give the reason for withdrawal. If potential participants are interested in joining the study, they will be offered the choice to contact the study team directly or give permission for the study team to contact them by their preferred method (phone, email etc). Potential participants will have the opportunity to discuss the study with a member of the study team prior to taking part and will be given sufficient time to consider joining the study.

Informed consent

Informed consent will be obtained from each participant, by suitably qualified study personnel who have been approved by the local Principal Investigator prior to any study procedures taking place. Due to the COVID-19 pandemic, additional methods of providing informed consent may be employed to avoid unnecessary contact. In addition to acquiring written consent face-to-face, participants may be offered the opportunity to provide consent via email, post, or verbally over the telephone. The informed consent process will be completed on the day of the screening study visit or in advance of the baseline visit if consenting via phone, email, or post. For written consent provided face-to-face, via email, or phone, the participant must personally sign and date the latest approved version of the informed consent form (ICF) before any study specific procedures are performed. For verbal consent, the participant will not sign the ICF; however, the study team will record on a physical ICF that this participant consented via telephone. The method of consent will also be recorded in the electronic data capture (EDC) system used for the study. The original ICF will be retained at the study site, and a copy will be provided to the participant and filed in the medical notes of the participant. Participants shall be asked to provide consent for the collection of data at the baseline medical assessment, as well as medical note reviews at 3 and 6-months following trial entry. Consent shall include explicit permission for the data collected to be used by Ultromics for commercial benefit. Biological specimens shall not be taken for any reason other than standard care as per local healthcare provider policy. No specific post-trial provision is required for enrolled participants. The sponsor has appropriate trial insurance in place.

Randomization

At study entry a unique study record will be created for each participant on a secure web-based EDC system provided by Castor EDC. A unique study identification number will be generated for each participant. Participants will be randomized using a secure web-based platform integrated into the EDC system. All members of the immediate study team at each participating site will have access to the EDC and randomization tool within. The immediate study team will randomize each participant after valid consent has been acquired and at the baseline visit (Visit 1). Randomization does not impact the stress echocardiogram procedure, and can occur immediately prior to, during, or after the stress echocardiogram procedure. The allocation will not be blinded to the immediate study team. Participants will not have access to the EDC system and thus will not know which randomization arm they have been assigned to.

Randomization will be stratified by gender, age, recruiting centre, and pre-existing CAD (defined as previous Acute Coronary Syndrome, or evidence of CAD confirmed on imaging sufficient to change their management course from primary prevention to secondary prevention). The randomization schedule will be generated using Castor EDC by the trial statistician. Adherence to the randomization schedule will be monitored by the statistician throughout the trial.

To minimize potential bias in outcome assessments, clinicians will be asked not to share the details of the AI classifier's report received for participants within the intervention arm with the participant or with colleagues who may undertake future investigations e.g. angiography. If a clinician subsequently considers there is a clinical need to know the details of the AI classifier's report, this can be made available at the discretion of the local primary investigator.

Intervention

Participants randomized to the intervention group will undergo a stress echocardiogram at baseline (Visit 1) in accordance with the local healthcare standard protocols. No concomitant care is prohibited during the trial. In addition, participants will have their stress echocardiogram images sent for additional assessment using the EchoGo Pro AI classifier. Once uploaded, the stress echocardiogram images undergo a technical quality control check prior to auto contouring with the AI to verify consistency and completeness (confirm images are present and that no errors have occurred). Stress echocardiograms are eligible for processing through the AI device if they contain the required views for peak and rest images (apical 2 chamber and 4 chamber views as well as parasternal short axis mid ventricle level view) and pass quality assessment. If stress echocardiogram images do not meet the requirements for processing, a rejection report will be generated. Feedback (including additional guidance for image acquisition best practices) will be provided to the site as needed during the course of the study. Automated AI generated contours of the endocardial border will then be used to extract geometric and kinematic features of endocardial motion which are used by the AI classifier to predict disease. The accuracy of automated contours is assessed by competent and qualified operators (all required to hold accreditation from their respective recognized professional bodies ie, British Society of Echocardiography), employed by Ultromics Ltd. The referring clinician will be provided with a copy of the AI analysis report within 15-20 minutes to use as a support to their clinical decision-making process (Supplementary File 1). This report is binary in nature, providing a result in which the regions of interest within the echocardiogram are suggestive of a high or low probability of severe CAD (where severe CAD is defined based on the presence of $\geq$70% stenosis in the proximal to mid left anterior descending artery (LAD), proximal left circumflex (LCx) or proximal to mid right coronary artery (RCA) and/or CAD related event within 6- months of SE, consistent with previous definitions of severe CAD). Data will be collected on the effect of the AI analysis report on the diagnostic confidence of each clinician. The version of the AI classifier will be the same as whichever is regulatory cleared and currently in use in standard service to show a fair representation of device performance within the healthcare setting.

Comparator group

Participants in the control group will also undergo a stress echocardiogram at baseline (Visit 1) as for the intervention group. Participant stress echocardiogram images will be sent for assessment as for the intervention group, but these AI analysis reports will not be returned to the treating clinician. Participants in the control group will continue down their respective care pathway based on the treating clinician's decision only, without the AI analysis diagnostic aid. Participants in the comparator group will be adjudicated after study completion as for the intervention participants in order to assess secondary outcomes.

Withdrawals

Participants will have the right to withdraw from the study at any time without giving a reason and without impacting any future care they receive. Reasons for withdrawal, if given, will be recorded on the electronic case report form (eCRF), along with the date of withdrawal. In addition, should a participant be randomized for study entry but meet any exclusion criteria during the baseline stress echocardiogram, the participant will be withdrawn from the study as a screen fail. Participants who decide to withdraw will be asked for permission to use data already collected, and if their notes can be reviewed for the purposes of follow-up/data collection.

Participant confidentiality

Role based logins will allow restricted access to the online EDC to defined study team members for specified purposes. The participant's date of birth and trial identification number will be used for identification on the eCRFs. The participant's email address will be recorded to enable follow-up using the EQ-5D-5L and entered into the secure database, but only viewable to research staff at sites. All study documents will be stored securely and only accessible to the study team and authorized personnel. All team members who have access to participant data will be subject to the Data Protection Act 2018, General Data Protection Regulations (GDPR) and all local organization requirements. Archiving will take place

following the completion of the study and publication of study results, in accordance with NHS guidelines, for 25 years.

Data collection

All collected participant data will be entered into the EDC throughout the course of the study. Both the EQ-5D-5L and short Seattle Angina Questionnaire (SAQ-7) will be used as patient-reported health questionnaires throughout the course of the study. Baseline participant demographic information will be collected by the immediate study team via participant medical note review. The patient-reported questionnaires will be provided to the participants at baseline (Visit 1) by the study team and acquired before the participant leaves the clinic. In addition, details of the stress echocardiogram procedure will also be entered into the EDC by the study team. The immediate study team will also be responsible for collecting participant outcome data via medical note review, and providing patient-reported health questionnaires at both 3 and 6-months (Visits 2 and 3) postrandomization. The clinical care teams at each site will send out reminders to the participants asking them to return the questionnaires at 3 and 6-months postrandomization. Reminders will be sent 2 weeks later if a completed questionnaire has not been received. Participants will be made aware of the questionnaires at the point of enrolment, with estimated time expenditure stated to ensure transparency and improved adherence to the protocol. The majority of the clinical outcomes data shall be retrieved from the 6-month medical note review at site. Other secondary outcomes data on self-reported participant symptoms will be collected from the EQ-5D-5L and SAQ-7 at 3 and 6-months post baseline examination, and data from the EQ-5D-5L questionnaire will be used to assess the intervention cost/benefit impact. Follow-up data will also be used to collect all information on further tests and procedures to allow a health economic analysis. Further information shall be collected through self-reporting to assess clinician diagnostic confidence in the AI classifier at baseline, 3 and 6-months.

Adverse events

Due to the nature of the disease, it is likely that participants will experience serious adverse events during the time they are enrolled in the trial. In addition, participants in the study population are likely to experience comorbidities. Deterioration of existing medical conditions, hospital admissions due to acute illness, and new medical problems are also expected. Therefore, events meeting the criteria of an "expected" or "unexpected" adverse event or serious adverse event will not be reported unless they are also classed as "related." Post market surveillance of the AI classifier is conducted as part of its regulatory obligations for CE marking to identify and analyze any performance errors.

Data management

Baseline and outcome data will be collected via eCRFs onto a password protected EDC system provided by Castor EDC. Data will be entered at site and subject to validation checks to ensure completeness and validity of data collected. Members of the study team will monitor data validity and completeness. Audits will be undertaken following a risk-based approach.

Clinical data recorded on the eCRFs is stored on Castor EDC servers in the UK and compliant with relevant data protection legislation. No study data will be stored on local machines. Any paper-based materials such as consent forms containing participant identifiers will be held securely at the research sites following local Trust procedures. Stress echocardiogram images sent to Ultromics for the purpose of the study are automatically deidentified during the transfer pathway from each site to Ultromics and ultimately stored on a secure Azure UK server, with access restricted by secure log-in to authorized individuals within Ultromics only. All data management procedures are conducted per the PROTEUS Data Management Plan (DMP) and according to Ultromics Ltd standard operating procedures (SOPs).

Statistical methods

A study adjudication committee, comprised of at least one accredited cardiologist, will review outcome data collected during the 6-month medical records review as described above in the description of primary and secondary outcomes. The adjudication committee will be blinded to the results of the stress echocardiogram and/or the EchoGo Pro report. The committee will review the uploaded discharge summaries and/or cardiac investigation/procedure reports to confirm participant outcome. Receiver operator characteristic (ROC) curve analysis will be used to assess the accuracy of deciding to refer a participant for coronary angiogram in PROTEUS. The DeLong test will be used to obtain the significance of the difference in AUROC between comparator and intervention. Testing will be one-sided with a non-inferiority margin of 0.05 (Comparator-Intervention $\geq$ 0.05).

For secondary analysis, ROC curve analysis will be used to test for superiority of the intervention over the control using the principles applied for the primary analysis and for evidence of change in medical management decisions. For other outcomes based on estimation of odds ratio between control and intervention arms we propose a Generalized Logistic Mixed Model (GLMM) method using the binary composite secondary outcome. Depending on the outcome either logistic or linear mixed models will be used. This modelling method will include both fixed and random effects, allowing to account for the expected

site-specific variation in the random effects. The fixed effects will include a variable defining the echo being processed through the comparator or intervention arm. The outcome will be defined using a dichotomous variable defining whether the participant received the appropriate decision in terms of referral to elective invasive coronary angiogram or not. Residuals of this model should show no evidence of deviating from the linearity assumption of GLMM and transformations will be explored if required. Analysis can be performed either in SAS-STAT 14.3 (glimmix procedure) or R 4.0.3 (lme4 package). The following secondary outcomes are not suitable for analysis using GLMM:

- Change in health-related Quality of Life (measured by EQ-5D-5L at trial entry, 3 and 6-months)
- Change in CAD symptoms and impact on participant health status (measured by patient-reported SAQ-7 trial entry, 3 and 6-months)

A longitudinal model will be fitted to address whether there is a significant change over time.

Subgroup analyses will be based on additional features, including:

- Biological sex (male vs female)
- Age
- Coronary disease (known coronary disease vs no disease)
- LV function (regional wall motion at rest vs normal LV function)
- Stress Echo (contrast vs noncontrast stress echo)
- Stress Echo test (exercise vs pharmacological stress echo test)
- Number of stress echocardiograms performed at site (high vs low volume)
- Ultrasound machine vendor
- Per-recommendation analysis (ie, as per the AI classifier's report recommendation)

When a significant association is reported, subgroup analysis of the outcome stratified over this variable will be performed. Further analysis will look at the learning and decision curves of clinicians based on their expert level.

Missing and inconclusive data will be defined as:

1. Missing at random (MAR) - Unrelated to any of the observed or unobserved data
2. Missing not at random (MNAR) - Related to the unobserved outcome data
3. Invalid inconclusive outcomes - Uninterpretable or missing outcome features not related to unobserved outcome data.
4. Valid inconclusive outcomes - Data where imaging diagnostic features have been obtained, but the result is not clearly positive or negative.

There is no single "optimal" approach to analyzing missing or inconclusive results. In these scenarios within this this study, where possible, diagnostic accuracy will be analyzed in line with how the test in question would be used in clinical practice. This will not be conducted for MAR and invalid inconclusive if confirmation is present that the cause is an intrinsic property of the test (an objective quality) rather than factors effecting false positive and false negative rates.[14,15] If multiple imputation methods based on correlation with related factors is possible, this will be implemented instead of the complete exclusion of the data.

For valid inconclusive results and MNAR, the outcome measure will be assigned to one of the binary outcome variables via an outcomes adjudication committee informed by additional relevant metadata (i.e. patient intervention, other imaging, medical records etc.).

Sensitivity analyses will be carried out on the primary and secondary outcomes to assess the impact of missing data; if there is greater than 20% missing data, appropriate imputation methods will be utilized, if there is less than 20% missing data, missing data sensitivity analyses will not be performed. As follows, all cases should be included in the performance analysis as far as practicable and should minimize associated performance biases.

Interim analyses

An interim analysis will be performed by the trial statistician when 30% of the original sample size has been recruited (750 participants) based on 3-month follow-up data. This analysis will be to check the assumption related to referrals for coronary angiography and incidence of severe disease i.e. the estimated kappa (ratio non-diseased/diseased) = 12) as well as the drop-out ratio being between 5% to 10%. Based on this information the Trial Steering Committee (TSC) will decide as to whether to adjust the sample size. There is no planned adjustment of the significance level foreseen, this would be kept at 5% for both interim and final analysis.

Trial oversight

The trial will be overseen by an independent TSC. The TSC will include an independent Chair, a majority of independent members, a PPI member, and the Chief Investigator and Clinical Lead. The Clinical Project Manager or delegate will act as the committee facilitator. The TSC will review overall progress of the trial and report to the funder. It will act on recommendations of the Data Safety Monitoring Committee (DSMC).

The trial sponsor has established a study specific DSMC, the membership of which includes independent subject matter experts. The independent DSMC will review progress and emerging safety data for the trial. It will provide recommendations to the TSC and to the Sponsor on trial conduct as necessary.

Dissemination policy

The results from the trial shall be disseminated through conference attendance and publication in peer reviewed open access journals. The protocol shall be published in an open access journal and permission has been sought from the trial participants to share anonymous data for future research purposes.

Discussion

The PROTEUS study will be the first randomized controlled trial to assess the impact of an AI medical device on the standard care pathway of suspected CAD patients. As the proposed use of AI within healthcare is gaining greater visibility, robust validation is required to maintain the confidence of patients and doctors in its application. This trial primarily aims to ensures there is no negative impact of introducing an AI classifier into the patient pathway while evaluating the broader impacts of the AI classifier on patient outcomes, clinician and patient experience and health economics. There are however, a range of other conditions and protocols relevant to the use of stress echocardiogram that were not included in the study design. Future technical developments will be required to generate results for these patients.

This study uses a predefined electronic randomization, in which the participants are blinded to the allocation; however, the referring clinician will be unblinded to the allocation, allowing us to measure changes in clinician confidence based on the recommendation of the technology and the impact of the medical device in the care pathway. Data will also be collected from the participants to measure their perceptions of AI in healthcare, as well as self-assessed symptom reporting. As the care pathway for those participants within the intervention arm and control arm are identical, the risk of adverse events related to the trial is reduced, and concerns regarding protocol adherence are removed.

Potential implementation problems for the study may lie in the integration of different ultrasound vendors and connectivity methods found in the healthcare settings that enroll to the trial. Robust connectivity assessments and trial echocardiogram studies will be processed prior to the study kick-off meetings in order to reduce unforeseen issues, alongside regular service monitoring to detect and resolve issues should they appear. Additionally, the study design and funding allowed for follow up to six months after stress echocardiogram. Although the majority of events will be captured during this period, it is likely additional adverse acute coronary events will occur after this six-month time window. A subsequent study to evaluate longer term follow up will be important.

As one of the leading causes of premature death,[2] CAD presents healthcare providers with a resource and economic burden which is set to rise.[16] For these reasons, we believe that the results from the PROTEUS study could present an important opportunity to provide cost effective and efficient changes to the standard care pathway, which could result in positive patient outcomes. In addition, we believe use of a large multicenter randomized trial to evaluate a novel AI diagnostic aid establishes an exemplar for future safe and effective adoption of AI into clinical practice.

Trial status

The trial commenced recruitment in September 2021 and is expected to continue until December 2023. Delays due to the COVID-19 pandemic have caused study timelines to be extended from the original estimated study end date of December 2022. At the time of manuscript submission, 1,941 participants have been recruited to the study. Protocol version at time of manuscript submission: Version 6.0 (January 19, 2023).

Sponsor contact information and responsibilities

Dr Jaco Jacob, Ultromics Ltd. Cascade Way, 4630 Kingsgate, Oxford Business Park, Oxford OX4 2SU, PROTEUS@Ultromics.com

The responsibilities of the sponsor include study design, management, data collection, analysis and interpretation, report writing, and publication submission.

Authors' contributions

GW, PL, RS, MB, BT, HD, RU, VW developed the study concept, protocols, funding applications, and initiated the project. OH, WH, WW assisted in the further development of the protocol. TD, WH, ES, HMD developed the Statistical Analysis Plan for the study. MBa developed the Health Economics Analysis Plan. VW and CF developed and ran the Qualitative Sub-study. BT, MB, CJ, NS, HP, ST, CJ, PL, SK were responsible for the day-to-day running of the trial. JL, CJ, and NS oversaw the Data Management. SB, EF, LL, TM, GD, DR made up the Trial Steering committee and VC, VH, SP the Data Safety and Monitoring Committee. CJ, LT, MB, TD drafted the manuscript. All authors contributed to and approved the final manuscript.

Availability of data and materials

The sponsor and relevant collaborators alone will have access to the whole dataset for analysis once the trial is complete. The model site agreement (contract between the sponsor and the clinical sites) states that sites can publish/present data collected at their own site once the main multicentre publication has occurred. The AI classifier is CE marked and commercially available in the UK. Use of the AI classifier is restricted to employees of Ultromics Ltd.

Ethics approval and consent to participate

The protocol, informed consent form, participant information leaflet, and proposed advertising materials were reviewed by the PPIE group, approved by study sponsor and the North West - Preston Research Ethics Committee (REC) for the National Health Service Health Research Authority (NHS HRA) (Reference 21/NW/0199).

Amendments to documents including the protocol, PIL, ICF, patient facing eCRFs, and other patient facing advertising and recruitment material, will be submitted to an appropriate REC and the NHS HRA by Ultromics Ltd. as necessary. All approved amendments will then be circulated to all participating NHS centres via email.

Consent for publication

Publication mechanism was detailed in the patient facing documentation and consent process. Ethically approved example documentation can be provided upon reasonable request.

Funding

This research was funded by the Artificial Intelligence in Health and Care Award, a competitive process run by the ACC in partnership with NHSx and the NIHR. The views expressed are those of the authors and not necessarily those of the NHSx, the NIHR or AAC. The funding body had oversight of the design of the protocol, but no involvement in the writing of this manuscript.

Declaration of Competing Interest

GW, RS, WH, TD, ES, MB, OH, LT, JL, RU are all employees of Ultromics Ltd GW, RS, WH, BT, JL hold share options for Ultromics Ltd PL and RU are founders and shareholders of Ultromics Ltd and have patents in the field of AI and imaging. SEP provides consultancy to Circle Cardiovascular Imaging, Inc, Calgary, Alberta, Canada. All other authors declare that they have no competing interests.

Acknowledgments

The authors would like to express gratitude to the generosity of the funders of this study—the Accelerated Access Collaborative (ACC) in partnership with the National Institute for Health and Care Research (NIHR) and NHSx, without whom this impactful and innovative work would not be possible. Additionally, the authors would like to acknowledge all participating NHS sites including all primary investigators and study delivery team members.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ahj.2023.05.003.

References

1. Foundation, BH. Global heart and circulatory diseases factsheet. 2021.
2. Foundtion, BH., UK Factsheet, British Heart Foundation. 2021.
3. Skinner JS, Smeeth L, Kendall JM, et al. NICE guidance. Chest pain of recent onset: assessment and diagnosis of recent onset chest pain or discomfort of suspected cardiac origin. Heart 2010;96:974–8.
4. Bhattacharyya S, Chehab O, Khattar R, et al. Stress echocardiography in clinical practice: a United Kingdom National Health Service Survey on behalf of the British Society of Echocardiography. Eur Heart J Cardiovasc Imaging 2014;15:158–63.
5. Woodward W, Dockerill C, McCourt A, et al. Real-world performance and accuracy of stress echocardiography: the EVAREST observational multi-centre study. Eur Heart J Cardiovasc Imaging 2022;23:689–98.
6. Sicari R, Nihoyannopoulos P, Evangelista A, et al. Stress Echocardiography Expert Consensus Statement–Executive Summary: European Association of Echocardiography (EAE) (a registered branch of the ESC). Eur Heart J 2009;30:278–89.
7. Secretariat MA. Stress Echocardiography for the diagnosis of coronary artery disease: an evidence-based analysis. Ont Health Technology Assess Ser 2010;10:1–61.
8. Neglia D, Rovai D, Caselli C, et al. Detection of significant coronary artery disease by noninvasive anatomical and functional imaging. Circ Cardiovasc Imaging 2015;8:e002179.
9. Pellikka PA, Arruda-Olson A, Chaudhry FA, et al. Guidelines for performance, interpretation, and application of stress echocardiography in ischemic heart disease: From the American Society of Echocardiography. J Am Soc Echocardiogr, 2020;33:1-41 e8.
10. Dey D, Slomka PJ, Leeson P, et al. Artificial Intelligence in Cardiovascular Imaging: JACC State-of-the-Art Review. J Am Coll Cardiol 2019;73:1317–35.
11. Karagodin I, Carvalho Singulane C, Woodward GM, et al. Echocardiographic Correlates of In-Hospital Death in Patients with Acute COVID-19 Infection: The World Alliance Societies of Echocardiography (WASE-COVID) Study. J Am Soc Echocardiogr 2021;34:819–30.
12. Upton R, Mumith A, Beqiri A, et al. Automated Echocardiographic Detection of Severe Coronary Artery Disease Using Artificial Intelligence. JACC Cardiovasc Imaging 2022;15:715–27.
13. Woodward W, McCourt A, Dockerill C, et al. Real-world use and accuracy of stress echocardiography: preliminary insights from the EVAREST study. Eur Heart J 2020;41(suppl 2).
14. Altman DG, Bland JM. Missing data. BMJ 2007;334:424.
15. Shinkins B, Thompson M, Mallett S, et al. Diagnostic accuracy studies: how to report and analyse inconclusive test results. BMJ 2013;346:f2778.
16. Bloom DE, Cafiero ET, Jane-Llopis E, et al. The global economic burden of non-communicable diseases. Geneva: World Economic Forum 2011.