




BMJ Open Evaluating artificial intelligence-driven stress echocardiography analysis system (EASE study): A mixed method study

Mahdi Mahdavi ,¹ Nicki Thomas ,² Chris Flood,³ Adele Stewart-Lord ,⁴ Lesley Baillie,⁵ Enrico Grisan,⁶ Patrick Callaghan,⁷ Rosica Panayotova,⁸ Sandeep S Hothi,^{9,10} Virgil Griffith,¹ Sharanya Jayadev,¹ Daniel Frings⁶

To cite: Mahdavi M, Thomas N, Flood C, *et al.* Evaluating artificial intelligence-driven stress echocardiography analysis system (EASE study): A mixed method study. *BMJ Open* 2024;**14**:e079617. doi:10.1136/bmjopen-2023-079617

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2023-079617>).

Received 06 September 2023
Accepted 21 August 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Daniel Frings;
Dan.frings@lsbu.ac.uk

ABSTRACT

Introduction The use and value of artificial intelligence (AI)-driven tools and techniques are under investigation in detecting coronary artery disease (CAD). EchoGo Pro is a patented AI-driven stress echocardiography analysis system produced by Ultromics Ltd. (henceforth Ultromics) to support clinicians in detecting cardiac ischaemia and potential CAD. This manuscript presents the research protocol for a field study to independently evaluate the accuracy, acceptability, implementation barriers, users' experience and willingness to pay, cost-effectiveness and value of EchoGo Pro.

Methods and analysis The 'Evaluating AI-driven stress echocardiography analysis system' (EASE) study is a mixed-method evaluation, which will be conducted in five work packages (WPs). In WP1, we will examine the diagnostic accuracy by comparing test reports generated by EchoGo Pro and three manual raters. In WP2, we will focus on interviewing clinicians, innovation/transformation staff, and patients within the National Health Service (NHS), and staff within Ultromics, to assess the acceptability of this technology. In this WP, we will determine convergence and divergence between EchoGo Pro recommendations and cardiologists' interpretations and will assess what profile of cases is linked with convergence and divergence between EchoGo Pro recommendations and cardiologists' interpretations and how these link to outcomes. In WP4, we will conduct a quantitative cross-sectional survey of trust in AI tools applied to cardiac care settings among clinicians, healthcare commissioners and the general public. Lastly, in WP5, we will estimate the cost of deploying the EchoGo Pro technology, cost-effectiveness and willingness to pay cardiologists, healthcare commissioners and the general public. The results of this evaluation will support evidence-informed decision-making around the widespread adoption of EchoGo Pro and similar technologies in the NHS and other health systems.

Ethics approval and dissemination This research has been approved by the NHS Health Research Authority (IRAS No: 315284) and the London South Bank University Ethics Panel (ETH2223-0164). Alongside journal publications, we will disseminate study methods and findings in conferences, seminars and social media. We will produce additional outputs in appropriate forms, for example, research summaries and policy briefs, for diverse audiences in NHS.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ The main strength of this study lies in its mixed-method approach to evaluation, which offers comprehensive assessment, data triangulation and increased validity of, and confidence in, findings.
- ⇒ The mixed-method approach will yield diverse forms of empirical evidence, financial evidence and stakeholder evidence.
- ⇒ By measuring willingness to pay from different users, this study will inform decisions around investment in healthcare artificial intelligence (AI) applications.
- ⇒ One limitation is that the impact of this AI application at the health system level, in terms of improved outcomes or quality of care, will not be measured in this study.

INTRODUCTION

Coronary artery disease (CAD), also referred to as ischaemic heart disease, is one of the leading causes of premature death globally. CAD was ranked as the second most common cause of disability-adjusted life years in 2019.¹ In older adults (50–75 years), it was the top leading cause of disability and premature deaths in 2019.¹ The potential sequelae of CAD are also remarkably costly to treat. An international systematic showed that on average, acute myocardial infarction, coronary artery bypass graft and percutaneous coronary intervention event costs \$11 686, \$37 611 and \$13 501 respectively for the health systems around the world.²

CAD is the most common cause of premature death in the UK. After adjusting for age effects, 3337 CAD events occur per 100 000 of the population in the UK,³ in turn leading to around 66 000 annual deaths.⁴ In 2015, total direct healthcare costs of CAD reached £2.2 billion; in the same year, total non-medical costs, including informal care and productivity losses due to morbidity and premature death reached £6.9 billion.⁴ Therefore, the

burden of CAD in terms of excess mortality, reduced quality of life and treatment costs makes CAD a major health system challenge. To address this challenge, health systems around the world have set standards of care and surveillance systems to assure accurate diagnosis of CAD among suspected patients (eg, the UK standard of care and monitoring system⁵).

CAD diagnosis requires a resource-intensive setting where stress echocardiography plays a pivotal role as a widely accepted test by patients and clinicians.⁶ The diagnosis test and process are usually led by a medical consultant who compares images of left ventricular wall motion acquired before and during exercise or pharmacological stress. A clinical judgement is made thereafter. There is a subjective nature around the tests' interpretation; as a result, clinical judgements may vary between raters, which likely increases inter-rater variability of diagnosis. Manual diagnostic interpretation may be inaccurate due to human errors.⁷ Conversely, the automated and consistent capability of artificial intelligence (AI) may assist clinicians in generating more timely and accurate CAD diagnoses.^{8,9}

AI-driven tools and techniques have been under investigation for the detection and classification of cardiac diseases,⁷ among which EchoGo Pro, an AI-driven stress echocardiography analysis system produced by Ultromics, is designed to support clinicians in detecting cardiac ischaemia and potential CAD.¹⁰ EchoGo Pro provides clinicians with a binary report indicating the presence or absence of CAD. The use of EchoGo Pro in simulated clinical decision-making processes in bench tests has improved the sensitivity of CAD detection by 10% compared with detection based on clinicians' judgement alone.¹¹ Another improvement made by EchoGo Pro concerns improved inter-rater agreement and confidence of clinical decision-makers.¹¹

While lab-based evaluation suggests AI-assisted ratings may be more accurate,¹² its acceptability, potential implementation barriers and clinicians' trust in AI in real-world clinical settings are unknown. Evaluating these aspects of AI-driven technology in front-line real healthcare environments is crucial before introducing it for routine use in patient care pathways. Moreover, there is a need to understand what obstacles, bottlenecks and contextual factors healthcare systems face when implementing new technology. The evaluation of EchoGo Pro provides an in situ test of an AI decision-making support tool's effectiveness in its diagnosis of CAD, clinician and patient acceptability, stakeholder confidence and also downstream service impacts such as improved health outcomes, cost-savings and value for money.

The evaluating AI-driven stress echocardiography analysis system (EASE) study and 'A Prospective Randomised Controlled Trial Evaluating the Use of AI in Stress Echocardiography' (PROTEUS trial)¹³ have been separately funded to evaluate Ultromics' AI technology and to inform the adoption of this technology by the National Health Service (NHS) in the UK. Findings will influence

whether the technology should be recommended for adoption across the NHS. The PROTEUS trial is a phase III trial, which is aimed at examining the efficacy of AI-assisted echocardiograms in a randomised controlled clinical trial. The details of the PROTEUS trial are given elsewhere.¹⁴ In contrast to the PROTEUS trial, the EASE study is fully independent of the technology's developer (Ultromics). Furthermore, the EASE study is a field study, which aims to evaluate accuracy, acceptability, implementation barriers, users' experience and willingness to pay, cost-effectiveness and value of the system.

This research will also examine if reduced inter-rater variability and improved diagnostic accuracy provided using EchoGo Pro can generate impacts downstream at a service level (ie, fewer acute cardiac events among screened patients and fewer false positives leading to inappropriate interventions and treatment). Furthermore, EchoGo Pro benefits to wider public health will be explored in terms of value and quality-adjusted life years (QALYs).

METHODS AND ANALYSIS

Research questions

This study will answer the research questions (RQs) below:

- ▶ RQ1: What processes have been used to deploy EchoGo Pro at sites and what barriers and enablers are highlighted for successful deployment of similar technologies?
- ▶ RQ2: Is the principal intended use of EchoGo Pro acceptable to healthcare professionals and patients? What factors predict intention to use and how do clinicians respond to discrepant information from AI decision support recommendations?
- ▶ RQ3: To what extent do EchoGo Pro recommendations align to cardiologists' judgements and what profile of cases is linked with divergence between EchoGo Pro recommendations and cardiologists' judgements?
- ▶ RQ4: How do cardiologists, healthcare commissioners and the public feel about the use of artificial intelligence for the analysis of stress echocardiograms?
- ▶ RQ5: Would EchoGo Pro be expected to deliver value?

Care pathway and EchoGo Pro

If CAD is suspected, stress echocardiography (SE) is one of a range of possible, recommended non-invasive tests, particularly in European Society of Cardiology and American College of Cardiology guidance documents.^{15–17} An SE is performed while deliberately increasing the heart rate with exercise or with medication.¹⁸ EchoGo Pro intersects with this pathway at the point that the SE images are analysed.⁸ Specifically, it is sent from the test machine to a cloud-based server, which analyses the study, returns values for key cardiac indexes and provides decision-making support. After SE images are uploaded into a secure cloud, the system returns basic and advanced

cardiac performance index (eg, strain, ejection fraction and left ventricular global longitudinal strain), as a part of the technology's core service package, EchoGo Core. EchoGo Pro (the module under investigation) also additionally provides clinicians with decision support by generating diagnostic recommendations for the identification of CAD (a binary, higher risk flag is returned by the system). It is the latter binary judgements which are the focus of the current evaluation.

Work packages

To address the research questions, we propose a mixed-method evaluation including accuracy bench testing, person-centred interviews and focus groups, a retrospective outcome data analysis, a cross-sectional acceptability and barriers quasi-experimental survey, and a health economic analysis. Research activities are organised into work packages (WPs) which will be conducted between March 2022 and March 2024.

Work package 1: independent accuracy bench testing

Bench testing examines the technical performance of the technology in isolation from the system it will sit within. We will examine the diagnostic accuracy by comparing test reports generated through EchoGo Pro and three manual raters of whom each will review 100 stress tests. Key outcomes to be evaluated are agreement rates between EchoGo Pro recommendations and expert

raters including index estimates and diagnostic alignment (table 1).

To detect inter-rater reliability between EchoGo Pro and manual raters, power analysis (using R's *n.cohen.kappa* function, assuming 5% of cases are judged as high risk, a power of 80% and a two-sided alpha of 0.05) suggests a sample of 108 is required to detect a difference between a hypothetical kappa of 90% and a true kappa of 55%. Thus, our target sample size (100 matched AI/clinical data points per rater, and three raters) approaches this threshold on an individual rater basis, and our total clinician and total AI sample (300) exceeds it.

A questionnaire will capture clinicians' ratings of heart segments rating wall abnormality (using the American Heart Association's 17-segment model¹⁹) and the degree of abnormality in each segment showing ischaemia. Further information to be captured includes whether the SE image is indicative of the presence of CAD or not, and the time taken to review each image. Inter-rater reliability analysis will be used to test the agreement between EchoGo Pro and human raters.²⁰ The details of the analysis plan are given in online supplemental file 1.

Work package 2: person-centred qualitative study

We will take a person-centred approach to fully understand the implications of EchoGo Pro for clinicians, Ultromics staff and patients, thus answering RQ2. WP2

WP	Outcome measures	Instrument/operationalisation	Instrument
1	Diagnostic alignment	Rates of agreement between EchoGo Pro and human expert raters.	Bench testing
2	Acceptability of AI system	Experienced cognitive and emotional responses of users of EchoGo Pro.	Interviews/focus groups
3	Alignment between clinician and EchoGo Pro report	Convergence or divergence between two diagnosis methods as below (i) convergence (with a diagnosis of CAD), (ii) convergence (no diagnosis of CAD), (iii) divergence (clinician identifying CAD) and (iv) divergence (AI diagnosing CAD).	Questionnaires
3	False positive AI prediction	Where AI predicts CAD, not identified by a reference criterion	Questionnaires
3	False positive clinician diagnosis	Where the clinician predicts CAD, not identified by a reference criterion.	Questionnaires
3	False negative AI prediction	Where AI predicts no CAD, CAD is confirmed by a reference criterion.	Questionnaires
3	False negative clinician diagnosis	Where the clinician predicts no CAD, CAD confirmed by a reference criterion	Questionnaires
4	Trust in AI	Perceived reliability, competence and intention to trust AI.	Questionnaires
5	Costs associated with health services/healthcare resources used	Costs of diagnosis models (with and without using EchoGo Pro).	Expert opinion/modelling
5	Value attributed to EchoGo Pro	The monetary value of EchoGo Pro services for CAD diagnosis.	Questionnaires/modelling

Note: A reference criterion refers to CAD identified by angiogram or CAD-related events that occur during a 6-month follow-up after a stress echocardiography test including elective event, hospital admission or death.
AI, artificial intelligence; CAD, coronary artery disease; WP, work package.

uses a qualitative design to provide rich data on the acceptability of use and implementation barriers of the technology. Semistructured interviews with clinicians and innovation/transformation staff (~8 participants per hospital), and relevant staff in Ultromics (~8 participants) will explore attitudes and perceptions of the use of the technology and barriers/enablers of successful deployment of the technology in healthcare settings.²¹ We will also undertake interviews and focus groups with patients in 2 NHS hospitals (8–10 patients per hospital) to explore patient attitudes and perceived value of technology.

Qualitative data will be analysed using Braun and Clarke's thematic analysis approach²² in six steps: (1) familiarising oneself with data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes and (6) producing a report. We will take an inductive analysis approach, with

the themes identified strongly linked to the data themselves. The details of the analysis plan are given in online supplemental file 1.

Work package 3: comparative evaluation

This WP will assess expert consultant interpretation of stress echocardiograms vs. Ultromics AI assessment of stress echocardiograms, specifically for the diagnosis of obstructive epicardial coronary disease (figure 1). We will create a reference criterion to compare both expert interpretation and AI assessment against it for the diagnosis. This criterion will provide a binary response whether the case is severe CAD or normal.²³ This criterion will use the binary information of angiogram (ischaemia identified or no ischaemia), further hospital admissions, cardiac events or death due to CAD during a 6-month follow-up post period after the time when SE was performed.²³ WP3

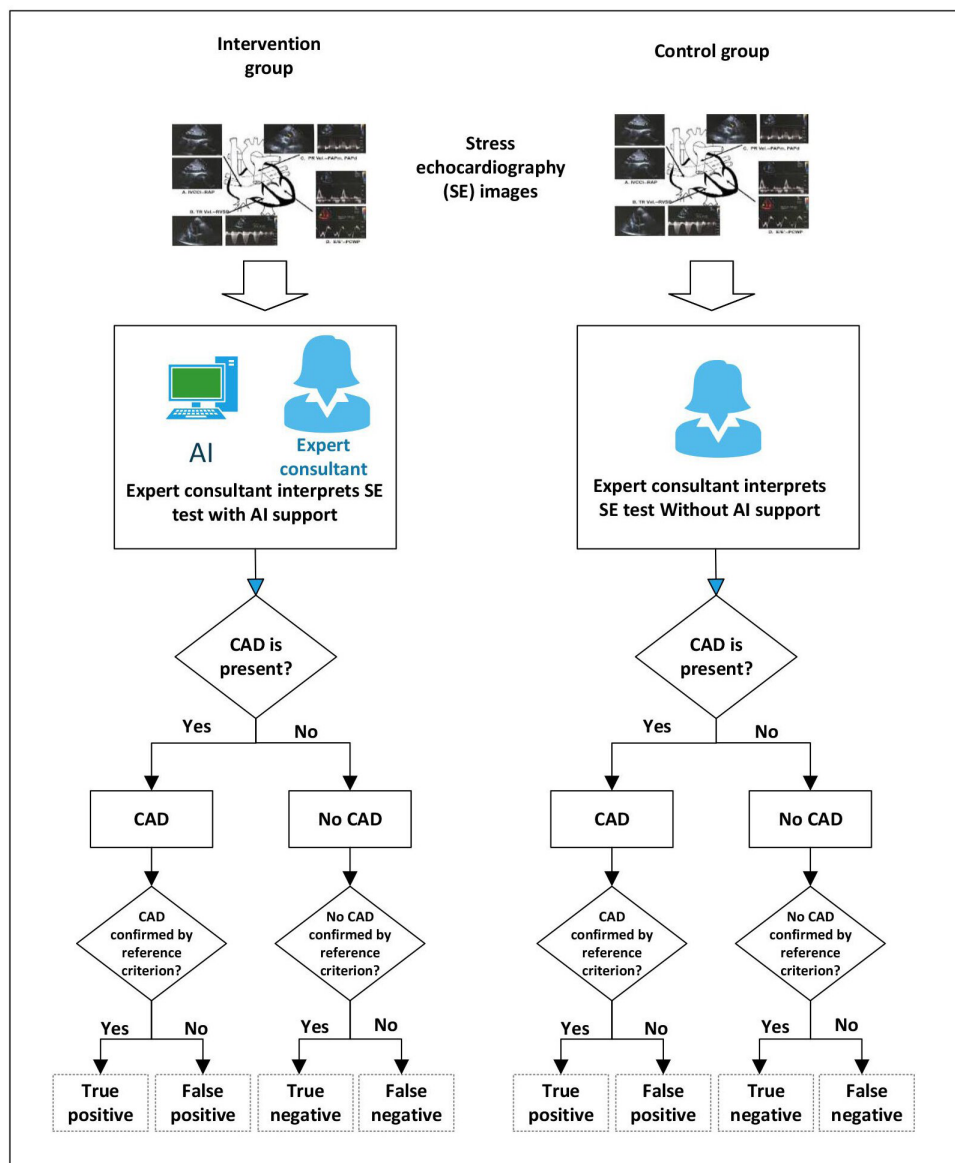


Figure 1 Steps to evaluate and compare the accuracy of diagnoses between intervention and control groups in work package 3. SE, stress echocardiography test; CAD, coronary artery disease; AI, artificial intelligence-driven echocardiography analysis system, EchoGo Pro.

will determine the extent to which EchoGo Pro recommendations align with cardiologists' interpretations of SE images, what profile of cases is linked with convergence/divergence between EchoGo Pro recommendations and cardiologists' interpretations (see Table for definitions) and how these are linked to outcomes. We will determine the number of divergent cases and test for identifying characteristics of such cases. We will compare health outcomes between convergent and divergent cases. We will describe how often clinical interventions were offered allowing insight into what may have happened if the AI had been provided and followed.

Power analyses using R's WebPower package (wp.logistic function) suggest that detecting a CAD prevalence of 4% in the divergent cases and 6% in the convergent cases (ie, 1% below and above the estimated 5% prevalence) would require an $n=1068$ (assuming power=0.80, alpha=0.05). In terms of detecting false positive angiogram test results, Gambre and colleagues report a positive predictive value of 86.2% for the detection of stenosis via catheter coronary angiography.²⁴ Power analysis with the same assumptions suggests detecting false negative rates which differ between 13% and 17.3% would require a sample of 619 tests (with power=0.80, alpha=0.05). We anticipate 893 patients in the intervention arm of the PROTEUS trial will be available for analysis.

To test for potential health inequality and inequity issues, we will also adopt a series of sensitivity analyses by testing the moderating effect of age and socioeconomic status for example, ethnic group and education. These analyses will be undertaken for the primary outcome measure, false negative SE results and true negative angiogram results.

Work package 4: survey of clinician trust in AI

WP4 will be a quantitative cross-sectional survey of clinicians with varying years of experience. We will measure confidence in AI and preparedness to use AI in two clinical scenarios: (a) at the point of care to obtain a diagnostic recommendation and (b) after consultation to provide a second opinion.²⁵ The clinician survey will be focused on the intention to use AI. This translates into an overall willingness to trust in AI, positivity or negativity regarding the system, potential conflict of interest between AI and clinicians, lack of transparency in the system and perception about AI. We will explore concerns around the use of AI that influence intention to use. Concerns potentially include changed levels of professional autonomy, independent decision-making and the fear of ceding control by adopting AI tools that may disrupt patient-physician relationships in the future.²⁶

Our sample for the clinician survey will consist of 60 cardiologists who currently read and interpret stress echo tests. They will be recruited through a third-party research recruitment panel (Qualtrics). We will develop a questionnaire and validate it by cardiologists already supporting this research, focussing on demographic and socioeconomic status, cardiac experience measures,

attitude to AI, degree of openness to and intention to use AI tools and techniques, reliability, competence and integrity of cardiac AI tools. We will conduct descriptive and least square regression analyses to identify determinant factors of trust in AI (see analysis details in online supplemental file 1).

Work package 5: health economics studies

We will calculate disaggregated costs and benefits of the new proposed service whereby the total costs of providing EchoGo Pro (including cost savings) and the total gains associated with the new AI technology overall can be reported separately. Additionally, a cost per unit of effect (ie, clinical improvement) will be estimated. Cost data at the hospital level would be incorporated alongside life expectancy data and any effectiveness data as measured through EuroQol EQ-5D to create QALYs and an average overall cost per change in QALY. The cost-effectiveness analyses will be achieved by comparing the changes in costs between the intervention and comparator groups, alongside changes in quality of life (using EQ-5D-5L).²⁷

We take a health services (NHS) provider perspective, in contrast with a societal perspective, for cost estimation. This will include all set-up costs and potential cost savings for NHS budgets associated with more efficient use of clinic time. Some direct costs incurred by patients and providers will be measured, including direct medical costs and other out-of-pocket expenses to capture a wider societal perspective and the impact of costs from the viewpoint of patients and carers. Further cost-benefit analyses will be conducted in the form of WTP methods to determine what the public, decision-makers and providers would be willing to pay for such AI technology. Hypothetical future potential costs and consequences will be set out for respondents in a scenario provided to value the technology. This scenario will be co-produced with the patient and public involvement (PPI) group to ensure it is easily understandable and reflects the potential real-world model of care. We will offer this scenario where the AI technology could be shown to have improved diagnostic accuracy and reduced waiting times.

We will aim to recruit samples from three population groups: cardiologists, healthcare commissioners and the public. We aim to recruit $n=60$ cardiologists (as an opportunity sample, associated with WP4). To test if the public samples' willingness to pay differs from the others and to detect a small effect difference ($d=0.2$, $\alpha=0.05$, power=0.95), a sample size of 327 completed cases is needed. This allows tests of differences between cardiologists' WTP and the commissioners' WTP at effect size $d=0.47$ (with the same assumptions) and differences between WTP in the public and cardiologists at $d=0.51$ (same assumptions). The commissioners' sample will likely be smaller and will either be used for descriptive statistics or if $n>30$, for inferential statistics with associated achieved power analysis reported.²⁸ The details of the analysis plan are given in online supplemental file 1.

ETHICS AND DISSEMINATION

This research has received a favourable opinion from the NHS Health Research Authority (IRAS No: 315284) and approval from the London South Bank University Ethics Panel (ETH2223-0164). Study sites will have oversight on the implementation of the research protocol in their sites and comply with national regulations on the use of data.

Access to study participants, who comprise healthcare professionals and patients, will be facilitated by gatekeepers in the relevant NHS sites and other organisations. Gatekeepers will inform potential participants through the invitations to participate in the study and will send them participant information sheets. Potential participants who are interested in joining the study will contact the research team to arrange for an interview/survey. All practical arrangements for the interviews will be undertaken via work email. The work emails of the interviewees will be stored safely and deleted at the end of the study.

We will seek consent from cardiologists reviewing SE images (WP1), the participants of staff interviews and patient interviews/focus groups (WP2), clinician surveys (WP4) and WTP study (WP5), before collecting any data. In WP1, we will ask cardiologists who review SE images to consent before completing reviews. For bench testing in WP1, we will use anonymised SE images.

The dissemination plan of this study endeavours to communicate the findings of this research to diverse stakeholders in a comprehensive and impactful manner. Targeting audiences such as cardiologists, clinicians and healthcare commissioners, we will submit detailed articles to academic journals, emphasising the study's insights into the accuracy, acceptability, implementation barriers, user experience, cost-effectiveness and the broader value of the EchoGo Pro analysis tool.

We have planned to present study findings in four conferences of relevant medical and AI domains which will further amplify our dissemination efforts. Leveraging social media platforms, particularly LinkedIn will facilitate the efficient distribution of key findings and engage a wider audience. We will present findings to cardiologists involved in the study and will collaborate with Ultromics. Within the NHS, research summaries and policy briefs will be circulated through our already-established communication channels. Public engagement will be prioritised through talks, webinars and accessible educational materials for a broader understanding of AI applications in cardiac care.

Patient and public involvement

Our research team members are well-experienced in working with PPI representatives. Our PPI group consists of 6 people with lived experience of stress echocardiogram, mostly recruited via the British Heart Foundation (Heart Voices). The members are diverse in terms of age, gender and ethnic group. The PPI group has been and will continue to be consulted regarding the importance and relevance of the research questions and the acceptability of the research design for patients. PPI representative(s)

will also inform the development of study instruments, participant information sheets, and consent forms and dissemination of findings. This is achieved via meetings every 6–8 weeks.

DISCUSSION

A field evaluation of the impact of AI-driven SE analysis as a decision-making aid is crucial before full-fledged implementation in wider health systems. The current research will use a mixed-method design to answer research questions on the accuracy, reliability, implementation barriers and enablers, and broader value of EchoGo Pro in cardiac care settings.

The main strength of this study lies in its approach to answering research questions. Specifically, we take a mixed-method, multilevel approach to research questions to shed light on the technology implementation barriers at the patient level, the care provider team, the organisational level, the funding level and so forth. This approach offers numerous benefits such as comprehensive multifaceted evaluation, data triangulation and increased validity of, and confidence in, findings.

By leveraging the strengths of a mixed-method approach, this study aims to generate a range of evidence that is required to support decision-making and ensure the successful adoption of AI technology. We expect to generate different types of relevant evidence thus: (1) empirical evidence for example, through service description and surveys to identify enablers and barriers of technology implementation; (2) financial evidence for example, benefit-costs analysis, cost-effectiveness analysis and willingness to pay studies to help assess the feasibility and sustainability of AI tools in real life; (3) stakeholder evidence from interviews and focus groups to capture experience and perception of people affected by or involved in the implementation process such as clinicians and patients; and (4) effectiveness and efficiency evidence that provides comprehensive understanding of the intervention's likely effects on quality of care and patient health outcomes. These types of evidence complement each other and increase the likelihood of effective implementation of the technology and provide lessons to be learnt by other healthcare AI applications.²⁹

The EASE study has a number of limitations that warrant consideration. First, in terms of evaluation scope, the assessment of expert consultant interpretation of SE vs. Ultromics AI assessment of SE only relates to significant epicardial coronary stenosis detection by stress echo. As such, microvascular disease falls outside the scope of the assessment. Microvascular disease and the broader question of anatomical-functional discrepancies may be the subject of future work. Second, the rescoping of the project in December 2022 has introduced a temporal delay, potentially compromising the study's original timeline. This could impact the accuracy and relevance of findings, especially given the dynamic nature of AI technologies. In WP1, we may overestimate the reliability

of images' ratings by cardiologists as each cardiologist will review images only once. Due to time constraints, we would not be able to deliver beyond the WP1 objective for example to test reproducibility on re-acquired images from the same study, or between repeat studies performed on a different day, or test-retest accuracy of the same image. Furthermore, in WP1, we do not have access to downstream outcomes or adjudication to use as an objective reference criterion for comparison/accuracy evaluation. In WP2, we may not have sufficient time to recruit sufficient interviewees or iterate a process to reach thematic saturation. Our evaluation in WP3 does not include a 'shadow mode assessment' (defined as evaluating the use of an AI system on an actual clinical situation but without influencing it) as recommended by the AI assessment guidelines.³⁰ The sample size for WP3 and the clinician survey in WP4 may present challenges in achieving comprehensive insights and a representative participant sample. Lastly, in WP5, since we will take a health services provider perspective to costs and benefits, we will be unable to consider broader societal costs and benefits of the technology. The willingness to pay questionnaire in WP5 highlights the potential improved accuracy of CAD diagnosis by the use of AI technology. However, we recognise the actual performance indicators cannot be included in the questionnaire given the study timelines. The study acknowledges these limitations, and users should exercise caution in generalising findings beyond the outlined constraints.

In conclusion, the findings which emerge from this work will provide evidence on key dimensions of the technology that altogether shape overall trust in AI tools in cardiac care and beyond. Our evaluation of an AI application to cardiac care will enable wide-ranging analyses which together have the potential to provide insight into and inform decisions around the entire ecosystem of AI tools in cardiac care, including non-metric-based factors such as cardiologists' perception of job security, task shifting to technology and patient-doctor relationship.²⁶

Author affiliations

- ¹School of Applied Sciences, London South Bank University, London, UK
- ²Institute of Health and Social Care, London South Bank University, London, UK
- ³School of Health and Social Care, London South Bank University, London, UK
- ⁴School of Allied and Community Health, London South Bank University, London, UK
- ⁵Department of Adult and Midwifery Studies, London South Bank University, London, UK
- ⁶London South Bank University, London, UK
- ⁷LSBU Doctoral College, London South Bank University, London, UK
- ⁸Stepping Hill Hospital, Stockport, UK
- ⁹Department of Cardiology, Royal Wolverhampton NHS Trust, Wolverhampton, UK
- ¹⁰Institute of Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Acknowledgements We thank the PPI representatives for their contribution to the ongoing implementation and dissemination of this project. We are grateful to South London Clinical Research Network, Christopher Ward, and Francesca Temple-Brown for providing us with study support services.

Contributors DF, CF, NT, EG, AS-L and PC conceptualised the study, developed the study design and acquired the funding. MM contributed to the development of study design and drafted the paper. VG and SJ contributed to study implementation

and project management. SH, RP, LB and all other co-author critically reviewed and approved the final version of the manuscript. DF is the guarantor.

Funding This work was supported by NHSx and NHS Accelerated Access Collaborative grant number REI2122-0030 (AI Award in Health and care, Technology Specific Evaluation Team). This report is independent research and the views expressed in this publication are those of the authors and not necessarily those of the NHS, NHSX or the Department of Health and Social Care.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

- Mahdi Mahdavi <http://orcid.org/0000-0002-7383-3455>
 Nicki Thomas <http://orcid.org/0000-0002-8311-7356>
 Adele Stewart-Lord <http://orcid.org/0000-0003-1594-5616>

REFERENCES

- 1 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. *Lancet* 2020;396:1204–22.
- 2 Nicholson G, Gandra S, Halbert R. Global Epidemiology of Ischemic Heart Disease: Results from the Global Burden of Disease Study. *CEOR* 2016;8:495–506.
- 3 Khan MA, Hashim MJ, Mustafa H, *et al*. Global Epidemiology of Ischemic Heart Disease: Results from the Global Burden of Disease Study. *Cureus* 2020;12:e9349.
- 4 Heart & Circulatory Disease Statistics, 2023. Available: <https://www.bhf.org.uk/what-we-do/our-research/heart-statistics/heart-statistics-publications/cardiovascular-disease-statistics-2023>
- 5 Cowie A, Buckley J, Doherty P, *et al*. Standards and core components for cardiovascular disease prevention and rehabilitation. *Heart* 2019;105:510–5.
- 6 Moss AJ, Williams MC, Newby DE, *et al*. The Updated NICE Guidelines: Cardiac CT as the First-Line Test for Coronary Artery Disease. *Curr Cardiovasc Imaging Rep* 2017;10:15.
- 7 de Siqueira VS, Borges MM, Furtado RG, *et al*. Artificial intelligence applied to support medical decisions for the automatic analysis of echocardiogram images: A systematic review. *Artif Intell Med* 2021;120:S0933-3657(21)00158-5.
- 8 Alsharqi M, Woodward WJ, Mumith JA, *et al*. Artificial intelligence and echocardiography. *Echo Res Pract* 2018;5:R115–25.
- 9 Massalha S, Clarkin O, Thornhill R, *et al*. Decision Support Tools, Systems, and Artificial Intelligence in Cardiac Imaging. *Can J Cardiol* 2018;34:827–38.
- 10 GitHub - equinor/EchoCore: Echo core - functions for echo framework app API, Available: <https://github.com/equinor/EchoCore>
- 11 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.
- 12 O'Driscoll JM, Hawkes W, Beqiri A, *et al*. Left ventricular assessment with artificial intelligence increases the diagnostic accuracy of stress echocardiography. *Eur Heart J Open* 2022;2:oeac059.



- 13 Breen M. PROTEUS: A PRO spective randomised controlled T rial E valuating the U se of artificial intelligence in S tress echocardiography. 2022.
- 14 PROTEUS: Evaluating the Use of Artificial Intelligence to Support Stress Echocardiography Testing for Heart Disease - Full Text View - ClinicalTrials.gov, Available: <https://clinicaltrials.gov/ct2/show/NCT05028179>
- 15 Gulati M, Levy PD, Writing Committee Members. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J Am Coll Cardiol* 2021;78:e187–285.
- 16 Erratum: 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes (European Heart Journal (2019) Doi: 10.1093/eurheartj/ehz425). *Eur Heart J* 2019.
- 17 Knuuti J, Wijns W, Saraste A, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J* 2020;41:407–77.
- 18 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.
- 19 Pellikka PA, Nagueh SF, Elhendy AA, et al. American Society of Echocardiography recommendations for performance, interpretation, and application of stress echocardiography. *J Am Soc Echocardiogr* 2007;20:1021–41.
- 20 Fleiss JL, Levin B, Cho Paik M. Statistical Methods for Rates and Proportions, 2004. 10.1002/0471445428 Available: <https://doi.org/10.1002/0471445428>
- 21 Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017;17:88.
- 22 Clarke V, Braun V, Hayfield N. Thematic analysis. Qualitative psychology: A practical guide to research methods. 2015;222:248.
- 23 Woodward G, Bajre M, Bhattacharyya S, et al. PROTEUS Study: A Prospective Randomized Controlled Trial Evaluating the Use of Artificial Intelligence in Stress Echocardiography. *Am Heart J* 2023;263:123–32.
- 24 Gambre AS, Liew C, Hettiarachchi G, et al. Accuracy and clinical outcomes of coronary CT angiography for patients with suspected coronary artery disease: a single-centre study in Singapore. *Singapore Med J* 2018;59:413–8.
- 25 Ford E, Edelman N, Somers L, et al. Barriers and facilitators to the adoption of electronic clinical decision support systems: a qualitative interview study with UK general practitioners. *BMC Med Inform Decis Mak* 2021;21:193.
- 26 Starke G, van den Brule R, Elger BS, et al. Intentional machines: A defence of trust in medical artificial intelligence. *Bioethics* 2022;36:154–61.
- 27 Dyer MTD, Goldsmith KA, Sharples LS, et al. A review of health utilities using the EQ-5D in studies of cardiovascular disease. *Health Qual Life Outcomes* 2010;8:13.
- 28 Heinrich-Heine-Universität Düsseldorf/ImpressumKontakt. G*Power: Statistical Power Analyses, Available: <http://www.psych.uni-duesseldorf.de/aap/projects/>
- 29 Brocklehurst PR, McKenna G, Schimmel M, et al. How do we incorporate patient views into the design of healthcare services for older people: a discussion paper. *BMC Oral Health* 2018;18.
- 30 Vasey B, Novak A, Ather S, et al. DECIDE-AI: a new reporting guideline and its relevance to artificial intelligence studies in radiology. *Clin Radiol* 2023;78:130–6.