HTE10067 Artificial intelligence assisted echocardiography to support the diagnosis and monitoring of heart failure

Final Protocol

Produced by: CEDAR (Centre for Healthcare Evaluation, Device Assessment, and Research)

Authors: Ayesha Rahim (Principal Researcher), Dr Huey Yi Chong (Health Economist), Dr Meg Kiseleva (Systematic Reviewer), Agatha Hills (Evaluation Scientist), Dr Jasmine Rollings (Senior Evaluation Scientist), Megan Dale (Principal Health Economist), Dr Rhys Morris (Director)

Correspondence to: CEDAR (Centre for Healthcare Evaluation, Device Assessment, and Research), Cardiff and Vale University Health Board, Cardiff Medicentre, CF14 4UJ

Date completed: 22/09/2025

Contents

1. Decision problem

The topic of this Early Value Assessment (EVA) is artificial intelligence (AI) assisted echocardiography to support the diagnosis and monitoring of heart failure.

Table 1 summarises the decision problem to be addressed in this assessment. Further detail on each element can be found in the published scope for the assessment.

Table 1. Summary table of the decision problem

Item	Description
Population(s)	Adults with suspected heart failure or who require ongoing monitoring for diagnosed heart failure.
Subgroups	If the evidence allows, the following subgroups may be considered: • People with acute onset of symptoms presenting as an emergency. • People with chronic symptoms initially referred from primary care.
Intervention(s)	AI technologies (adjunctive) for use in echocardiography to aid the diagnosis or monitoring of heart failure including: • EchoConfidence (MyCardium) • EchoGo Heart Failure (Ultromics) • Ligence (Ligence UAB) • Us2.v2 (Us2.ai)
Comparator	Echocardiography used in the diagnosis or monitoring of heart failure without adjunctive AI technologies
Setting	• Secondary care, echocardiography suite • Secondary care, emergency or ward setting
Outcomes eligible for inclusion	Diagnostic and intermediate outcomes: • Diagnostic accuracy outcomes including sensitivity, specificity, likelihood ratios, positive predictive value (PPV) and negative predictive value (NPV), receiver operator curves (ROC) • Diagnostic yield outcomes including false positive (FP) and false negative (FN) rates • Classification and stratification of heart failure by type and severity • Time to diagnosis • Time to initiation of treatment • Biomarker responses (e.g. changes to NT-proBNP) Clinical outcomes:

	• All-cause mortality: • Cardiovascular mortality • Heart-failure hospitalisations • Composite of cardiovascular death or hospitalisation due to heart failure • NYHA class shifts • 6-minute walk test Patient-reported outcomes: • Kansas City Cardiomyopathy Questionnaire (KCCQ) score improvements • Minnesota Living with Heart Failure Questionnaire (MLHFQ) score improvements • Generic health-related quality of life (e.g. EQ‑5D‑3L, SF36) • Service user acceptability, views, experience and satisfaction • Carer acceptability, views, experience and satisfaction Costs and resource use: • Cost of technology • Cost of treatment and management • Cost of training • Cost of down-stream diagnostic tests (e.g. cardiac MRI) • Staff time at different specialisms and levels of pay • Staff cost at different specialisms and levels of pay • Health service resource use in different settings • Cost of health service resource use in different settings
Economic analysis	Costs will be considered from an NHS and Personal Social Services perspective. Costs and resource use outcomes for consideration should include: • Costs of the technologies including implementation costs, license fees, hardware, software and any operational costs (IT support, cybersecurity) • Training costs The time horizon for estimating the clinical and economic value should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.

1.1 Objectives

The purpose of this Early Value Assessment (EVA) is to summarise and critically appraise existing evidence on the clinical-effectiveness and cost-effectiveness of artificial intelligence (AI) assisted echocardiography to support the diagnosis and monitoring of heart failure.

The research questions this assessment will aim to answer are:

- What is the clinical effectiveness and cost-effectiveness of AI assisted echocardiography to support the diagnosis and monitoring of heart failure?
- What are the risks and safety considerations associated with using these technologies?
- What are the key gaps in the evidence for these technologies, and how might these be addressed?
- What are the practical, cost and resource implications of introducing the technologies into the current care pathway?

The following objectives are proposed to address the research questions:

Clinical Effectiveness

- Identify and assess evidence relating to the use and clinical effectiveness of the included technologies as it pertains to the scope
- For evidence not directly related to the scope, outline the potential generalisability and limitations of the evidence
- Report on any potential safety issues
- Report the evidence gaps, highlighting what data may need to be collected to inform these gaps

Cost Effectiveness

- Identify and assess economic evidence relating to the use of the included technologies within the scope
- Report on the technology costs and develop a simple model using available inputs, reporting on the plausibility of cost effectiveness

- Develop a conceptual economic model, or identify a suitable existing model, related to the scope, that can be used to inform future research and data collection

- Report available model inputs and evidence gaps

2. Evidence review methods

An independent search for relevant evidence will be conducted by the EAG. Evidence relevant to the scope will be identified using a combination of databases of published evidence and evidence provided by technology manufacturers through NICE. Rapid review methods will be adopted by the EAG, such as those outlined in the Cochrane Rapid Review Methods Guidance. A full systematic review will not be conducted as this is outside of the scope for an EVA, as outlined in sections 3.7 - 3.16 of the Early value assessment interim statement (NICE, 2022).

2.1 Inclusion criteria

Table 2 outlines the inclusion and exclusion criteria considered for the evidence identified. If a large volume of evidence is identified, certain evidence may be prioritised for inclusion (see Section 2.3).

Table 2. Inclusion and exclusion criteria

	Inclusion Criteria	Exclusion Criteria
Population	Adults suspected of having heart failure or who require ongoing monitoring for diagnosed heart failure.	Adults suspected of other cardiac conditions, or those receiving echocardiograms for other clinical purposes, unless there is no evidence available that is directly relevant to the diagnosis or monitoring of heart failure. Animal studies will be excluded.
Intervention	• EchoConfidence (Mycardium) • EchoGo Heart Failure (Ultromics) • Ligence (Ligence UAB) • Us2.v2 (Us2.ai)	Any AI or digital technologies not named in the scope. AI technologies that are not adjunctive to or compatible with existing echocardiography systems in the NHS.

Comparators	Echocardiography used in the diagnosis or monitoring of heart failure without AI technologies	Comparative evidence where the comparator does not reflect the UK clinical pathway, unless no relevant evidence is identified with comparators relevant to the UK.
Setting	• Secondary care, echocardiography suite • Secondary care, emergency or ward setting Relevant evidence from community care settings may be considered for inclusion, if feasible in the given timescale.[α]	Evidence from settings outside of the UK, unless no relevant evidence is identified from UK-based settings.
Outcomes	Those included in the scope.	Evidence will be excluded if no relevant outcomes are reported. If a subsection of outcomes are relevant to the scope, these alone will be reported.
Study designs	Randomised controlled trials, retrospective and prospective observational studies, diagnostic accuracy studies, cross-sectional studies, case series.	Narrative reviews, case reports, editorials, letters.
Publication type	Full-text publications. Abstracts and conference proceedings will only be included if there is sufficient information reported for the purposes of this assessment.	Abstracts and conference proceedings will be excluded if: - there is an associated full-text paper available for the same study, - there is insufficient information reported for adequate assessment, - there is a large volume of relevant full-text publications available.
Key: α: Applicable to the clinical evidence review only. Evidence from community care settings will not inform the economic modelling, as this is outside of the scope.		

2.2 Search strategy

Searches will be developed in MEDLINE (Ovid) by an experienced Information Specialist. Search terms will include free-text terms and controlled vocabulary (e.g. MeSH). The search strategy will consist of two parts combined with the “OR” operator: one part constructed around the device and manufacturer names and the

other aimed at retrieving relevant literature that does not name the devices in the meta-data indexed by the databases, and this part will pragmatically combine population and intervention terms in a way that ensures that the screening stage is feasible within the agreed timeframe. The search strategy will be peer-reviewed by a second Information Specialist. A draft search strategy is available in Appendix A. The search strategy will be translated to each database.

The following bibliographic databases will be searched:

- Medline ALL (Ovid)
- Embase (Ovid)
- Cochrane Database of Systematic Reviews (CDSR)
- Cochrane Central Register of Controlled Trials (CENTRAL)
- International HTA database (INAHTA)

The following clinical trial registries will be searched:

- ClinicalTrials.gov
- International Clinical Trials Registry Platform (ICTRP)

Where possible, the EAG will identify additional studies from the information companies provide to NICE. To identify studies that have not been retrieved by the database searches, company websites will also be searched for relevant publications. The EAG may consider searching for pre-prints should there be a paucity of evidence identified from the sources described above.

The EAG anticipates that the main search strategy described above will identify both clinical and economic evidence relevant to the technologies in scope. Where evidence is identified as relevant to the wider decision problem, but is not specific to the technologies, the EAG will consider a brief summary in the final report. Additionally, a targeted literature search may be conducted to update model inputs for this assessment, if necessary. To increase the relevance of the results, filters will be added to this search, such as the economic evaluations and models (CADTH,

2025), health utilities (CADTH, 2025), and the NICE UK geographic filters (Ayiku, 2017).

2.3 Study selection

Retrieved references will be imported into EndNote and deduplicated, after which they will be imported into the screening tool Rayyan, where deduplication will be completed and records screened. The titles and abstracts of the identified studies will be screened by one reviewer and a minimum of 20% of excluded records will be checked by a second reviewer against the pre-specified inclusion and exclusion criteria (see Table 2). The AI screening tool available within Rayyan will not be used and all decisions will be made by the review team. Where a record appears to meet the eligibility criteria, or where a decision cannot be made based on the information provided in the titles and abstracts alone, it will be progressed to the full-text screening stage. The full texts of the articles progressed to this stage will be obtained and screened by one reviewer, with a random 20% of exclusions checked by a second reviewer. Only records where it is made explicit that the examined technology is one of those in scope will be included in the review. A list of studies excluded at the full-text stage, with reasons for their exclusion, will be presented in an appendix in the report.

Where a large volume of evidence is identified, a pragmatic approach to study selection may be taken, in line with the Early value assessment interim statement (NICE, 2022). Prioritisation of studies to be included may be based on factors such as study design, availability of relevant outcomes and extent of generalisability to a UK population. Clinical experts may be consulted to inform these decisions. Any decisions made and approaches taken by the EAG will be flagged with the NICE team for discussion and presented transparently in the final report.

The evidence review aims to identify the most relevant evidence relating to the decision question defined in the scope. If no evidence directly relevant to the technologies in scope is available, a broader evidence base may be evaluated. This will be decided upon in conjunction with NICE and transparently reported by the EAG.

2.4 Data extraction strategy

Where available, the following data will be extracted from studies: study information (i.e., author, year), study design, study dates, intervention characteristics (i.e., intervention name), comparator, participant characteristics (i.e., demographics, comorbidities) and participant outcomes which are relevant to the scope. Data extraction will be conducted by one reviewer using Microsoft Word and checked by a second.

2.5 Quality assessment strategy

Formal quality assessment of individual studies using validated checklists will not be conducted, in line with the expectations outlined in the Early value assessment interim statement (NICE, 2022). A narrative summary of the key strengths and limitations of the evidence will be presented in the final report. This summary will highlight potential biases in individual studies, discuss how these impact on the certainty of the results and outline how this might impact generalisability to NHS clinical practice.

2.6 Methods of synthesis and analysis

The EAG will consider meta-analysis methods to synthesise relevant clinical evidence. However, due to the nature and purpose of an Early Value Assessment (EVA), it is not anticipated that there will be enough data available to conduct a meta-analysis.

Results for both clinical and economic literature will therefore be presented in a suitable tabular format accompanied by a narrative synthesis of the data, considering available evidence relating to all aspects of the scope (for example, population, setting, comparators). Methodological issues with included studies will be noted along with any identified risks of bias which may impact study results. A discussion outlining the applicability of the evidence to the scope of the EVA will be included, as well as consideration of the generalisability of evidence to clinical practice in the NHS.

3. Economic analysis methods

The scope of the economic analysis will depend on data availability. If data allows, an early economic model will be developed, to compare costs and outcomes of AI

technologies in the scope. An appropriate time horizon will be used to capture costs and outcomes, based on available evidence. Existing models will be identified, prioritising models used in NICE guidance. If unavailable, this may expand to literature. Existing models will be assessed to determine their suitability as the basis for the conceptual model. Clinical expert validation will be sought to assess the face validity of the EAG early economic model and any model modifications on the conceptual model. However, if data are not sufficient to populate the early economic model, a simple cost comparison model may be developed to assess costs, staff time and waiting times between technologies.

3.1 Model development

The economic analysis will be performed in line with the NICE reference case. The perspective of NHS and Personal Social Services will be undertaken. Costs will be expressed in 2024 prices and where applicable, costs will be inflated using NHS Cost Inflation Index (NHSCII). If sufficient data are available, quality-adjusted life years (QALYs) will be the primary outcome in the economic analysis, calculated using utility values for each intervention. All costs and outcomes will be discounted at 3.5% per annum.

A simple decision analytic model will be developed in Microsoft Excel to compare AI-assisted echocardiography and echocardiography alone in the diagnosis of heart failure. Where possible, the impact of AI-assisted echocardiography on staff time, procedure time, and subsequently the waiting time will be considered. If evidence allows, relevant costs and outcomes of events occurring when waiting for an echocardiography appointment may be modelled. Diagnostic performance of each intervention (sensitivity, specificity and other relevant diagnostic accuracy outcomes) will be considered. Where possible, downstream costs and QALYs will be modelled based on diagnostic yield outcomes (true and false positives and negatives) and stratified by HF type. One-off costs and QALYs of each diagnostic outcome may be applied. Where sufficient clinical evidence is available, a longer-term model may be considered to capture the impact of earlier diagnosis and improved care.

If data allow, the following subgroups will be evaluated. In each subgroup, relevant health resource use, training costs, different staff delivering the service and subsequent treatment based on diagnosis will be considered.

- People with acute onset of symptoms presenting as an emergency
- People with chronic symptoms referred from primary care

The model inputs will be informed by evidence identified from the EAG search, published literature, company submissions and consultation with clinical experts and patient representatives. If necessary, additional targeted searches may be considered to update model inputs. Resource use and costs for each intervention will be estimated, including technology costs, training costs, NHS staff costs. Resource use and costs associated with managing heart failure will be obtained from relevant clinical guidelines, published literature and clinical expert inputs. Model assumptions will be clearly described and informed by evidence or advice from clinical experts. Where possible, other relevant outcomes will be reported. This may include different staff delivering the service (grade and time), discriminative ability for HF types and severity, timeliness of diagnosis, time to initiation of treatment, and the proportion of false positive and false negative cases.

Deterministic and probabilistic sensitivity analyses will be undertaken to identify the key cost drivers and to explore the impact of uncertainty, where possible.

3.2 Conceptual modelling

If available, a suitable existing model will be identified as the conceptual model, and any model modifications for future analysis will be outlined.

Where existing models are not available or not suitable, the EAG will conceptualise the model structure that represents the pathways and parameters needed for this population. This could form the basis for future economic analysis to compare long-term costs and benefits between technologies. However, it is not feasible to develop an executable model and fully populate it. Model inputs that are likely to be key drivers of future cost-effectiveness will be identified.

3.3 Cost of reversing a decision

The cost of reversing a decision will be estimated, which may include any up-front costs to purchase the equipment and setting up the service, training costs and any costs that could not be recouped following their implementation in NHS. This will be explored in a sensitivity analysis, if appropriate.

4. Evidence gap analysis

Evidence gaps will be identified and summarised in a suitable tabular format. The EAG will use the available evidence base in combination with input from clinical experts to determine where the evidence gaps lie and how these gaps may be addressed. Suggested methods of evidence generation to fill evidence gaps will be summarised in the report and key outcomes that may be required will be identified.

5. Handling information from the companies and other stakeholders

All data submitted by the companies in evidence and information requests by NICE, or data submitted by other stakeholders will be considered by the EAG if received by 10/10/2025. Information arriving after this date will not be considered. If the data included in the information provided meets the inclusion criteria for the review, they will be extracted and quality assessed following the procedures outlined in this protocol. The EAG may seek clarification or additional information from companies and other stakeholders where necessary. All correspondence between the EAG and companies will happen through NICE.

Any 'commercial in confidence' data provided by a company and specified as such will be highlighted in blue and underlined in the assessment report. Any 'academic in confidence' data provided by company(s), and specified as such, will be highlighted in yellow and underlined in the assessment report. If confidential information is included in the economic model, the EAG will provide a copy of the model with 'dummy variable values' for the confidential values (using non-confidential values).

6. Competing interests of authors

None.

7. References

Ayiku, L., Levay, P., Hudson, T., Craven, J., Barrett, E., Finnegan, A. and Adams, R. (2017), The medline UK filter: development and validation of a geographic search filter to retrieve research about the UK from OVID medline. Health Info Libr J, 34: 200-216. https://doi.org/10.1111/hir.12187

Economic Evaluations & Models - MEDLINE. In: Canada's Drug Agency Search Filters Database. Ottawa: Canada's Drug Agency; 2025: https://searchfilters.cda-amc.ca/link/16. Accessed 2025-09-17.

Economic - Health Utilities / Quality of Life - MEDLINE. In: Canada's Drug Agency Search Filters Database. Ottawa: Canada's Drug Agency; 2025: https://searchfilters.cda-amc.ca/link/19. Accessed 2025-09-17.

National Institute for Health and Care Excellence (NICE). (2025) Artificial intelligence assisted echocardiography to support the diagnosis and monitoring of heart failure [GID-HTE10067] Final scope. Available from: Project documents | Artificial Intelligence assisted echocardiography to support diagnosis of heart failure: Early Value Assessment | Guidance | NICE

Early value assessment interim statement (NICE). (2022) Available from: 1 Introduction | Early value assessment interim statement | Guidance | NICE

Appendix A: Draft search strategy

Ovid MEDLINE(R) ALL <1946 to September 09, 2025>

#	Query	Hits
1	EchoGo*.mp.	10
2	Ultromics.mp.	5
3	ligence*.mp.	6
4	"Us2.ai*".mp.	9
5	"Us2.v2".mp.	0
6	"eko.ai*".mp.	0
7	"A*STAR Biomedical Research Council".mp.	3
8	"A*STAR Exploit Technologies".mp.	1
9	EchoConfidence*.mp.	0
10	MyCardium*.mp.	12
11	or/1-10	40
12	exp Heart Failure/	161678
13	((heart or cardiac) adj2 (failure or insufficiency or decompensation)).tw.	258448
14	(HFrEF or HFmrEF or HFpEF).tw.	10361
15	or/12-14	297381
16	((echocardi* or "echo cardi*" or "transthoracic cardi*" or ((heart or cardi*) adj2 (ultraso* or sonogra*))) adj10 (AI or "artificial intelligence*" or "deep learning" or "machine learning" or "neural network*" or CNN or DNN or "augmented intelligence" or "automated recognition")).tw.	903
17	15 and 16	131
18	11 or 17	169
19	limit 18 to english language	164
20	exp animals/ not humans.sh.	5376333
21	19 not 20	160