

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

HealthTech draft guidance

Artificial intelligence (AI)-assisted echocardiography analysis and reporting to support the diagnosis and monitoring of heart failure

February 2026

Guidance development process

NICE HealthTech guidance evaluates digital technologies, diagnostics and medical devices (including artificial intelligence). It provides evidence-based recommendations about how safe and effective these technologies are, and their cost effectiveness. The guidance supports healthcare professionals and commissioners to ensure that patients get the best possible treatments. NICE aims to promote innovations that meet the needs of patients and the healthcare system.

This guidance has been developed as early-use HealthTech guidance, for HealthTech products that could address an unmet need in the NHS and need more evidence to support routine use.

Find out more on the [NICE webpage on HealthTech guidance](#).

NICE is producing this guidance on AI-assisted echocardiography analysis and reporting to support the diagnosis and monitoring of heart failure in the NHS in England. The diagnostics advisory committee has considered the evidence and the views of clinical and patient experts.

This document has been prepared for consultation with the stakeholders. It summarises the evidence and views that have been

considered, and sets out the recommendations made by the committee. NICE invites comments from the stakeholders for this evaluation and the public. This document should be read along with the [evidence](#).

The committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex or sexual orientation?

After consultation:

- Based on the consultation comments received, the committee may meet again.
- If committee meets again it will consider the evidence, this evaluation consultation document and comments from stakeholders.
- The committee will then prepare the final draft guidance, which will go through a resolution process before the final guidance is agreed.

Note that this document is not NICE's final guidance on AI-assisted echocardiography analysis and reporting to support the diagnosis and monitoring of heart failure. The recommendations in section 1 may change after consultation.

More details are available in [NICE's HealthTech programme manual](#).

Key dates:

Closing date for comments: 24 February 2026

1 Recommendations

More research is needed

1.1 More research is needed on the following artificial intelligence (AI)-assisted echocardiography analysis and reporting technologies to support the diagnosis and monitoring of heart failure, before they can be funded by the NHS:

- EchoConfidence
- EchoGo Heart Failure
- Ligence Heart
- US2.ai

What this means in practice

There is not enough evidence to support funding AI-assisted echocardiography analysis and reporting technologies to support the diagnosis and monitoring of heart failure in the NHS.

Access to the technologies should be through company, research or non-core NHS funding, and clinical or financial risks should be managed appropriately.

What research is needed

More research is needed on:

- the impact on service capacity, including:
 - time saved during echocardiography procedures when the AI technologies are used in NHS secondary, community or primary care settings
 - the impact on number of appointments per day and waiting times
 - time taken for human review of AI findings

- reliability and failure rate (for example, performance with echocardiograms of varying quality)
- diagnostic accuracy and performance when the AI technologies are used by operators with varying levels of experience outside of secondary care
- patient selection and cohorts used in the AI training and validation datasets, to assess whether these reflect the diverse population seen in NHS clinical practice
- healthcare professional-reported ease of use and acceptability
- any harm caused by using the technologies
- for technologies with the additional functionality, accuracy in detecting or distinguishing between heart failure and other less common cardiac conditions (which may co-present) such as amyloidosis and heart valve disease.

Why the committee made these recommendations

Limited evidence suggests that using EchoConfidence or Us2.ai could reduce the time taken for echocardiography analysis and reporting. But there is no evidence to suggest that the time saving results in more people being seen or reduced waiting times.

Most of the clinical-effectiveness evidence for the AI technologies relates to diagnostic accuracy. For all 4 technologies the evidence is uncertain and may not reflect clinical practice in the NHS. This is because many of the studies:

- are retrospective and observational
- were done in single centres with single operators
- were done outside of the UK
- excluded complex cases and poor-quality echocardiography scan images.

Because of the uncertainties in the evidence, it is not certain whether the AI technologies could lead to delays and longer waiting times if healthcare professionals need to check and review the AI findings. It is also uncertain whether EchoGo Heart Failure, Ligence Heart and Us2.ai were developed

using population data that reflects the NHS. So, using these AI technologies in the NHS could pose clinical risks.

The results of the economic model suggest that EchoConfidence and Us2.ai could be cost effective. But this is uncertain because of the uncertainty in the clinical evidence. It is not possible to assess the cost effectiveness of EchoGo Heart Failure and Ligence Heart because of limited evidence. So, more research is needed on all 4 technologies.

2 Information about the technologies

2.1 This assessment included 4 technologies (EchoConfidence, EchoGo Heart Failure, Ligence Heart and Us2.ai) that use artificial intelligence (AI) software to aid the interpretation and quantification of echocardiography images and automate report generation. All the technologies included in this assessment are designed to aid the operator, not replace them.

Table 1 Features of each technology

Technology (company)	Regulatory and DTAC status	Intended use	Target population
EchoConfidence (MyCardium)	Class IIb DTAC in place	<ul style="list-style-type: none">detecting and diagnosing heart failure through screening or clinical echocardiogramsstratifying heart failuremonitoring disease progression and treatment response	Adults with or without underlying cardiac disease
EchoGo Heart Failure (Ultromics)	Class IIa (expected June 2026) DTAC not in place	Providing adjunctive information on a cardiovascular condition to detect HFpEF	Adults over 25 years of age having routine functional cardiovascular assessment or with suspected heart failure
Ligence Heart (Ligence UAB)	Class IIa DTAC not in place	Detecting, measuring and calculating various specifications of structure and function of the heart	Adults 18 years and over who are not in a life-threatening state of health, time is not critical for medical decisions, and

Guidance – Artificial intelligence (AI)-assisted echocardiography analysis and reporting to support the diagnosis and monitoring of heart failure

Page 5 of 18

Issue date: February 2026

© NICE 2026. All rights reserved. Subject to [Notice of rights](#).

			no major therapeutic interventions are needed
US2.ai (EKO Pte Ltd)	Class IIb DTAC in place	<ul style="list-style-type: none"> processing, analysing and measuring TTE images, to provide automated estimation of several cardiac structural and functional parameters accelerating and standardising detection of most forms of heart failure, independent of ejection fraction 	Adults as decision support for detecting heart failure, pulmonary hypertension, cardiac amyloidosis, hypertrophic cardiomyopathy and valve disease (for example, aortic stenosis and mitral regurgitation)

Abbreviations: DTAC, digital technology assessment criteria; HFpEF, heart failure with preserved ejection fraction; TTE, transthoracic echocardiography.

Sustainability

2.2 For information, Carbon Reduction Plans for UK carbon emissions, or a commitment to reducing greenhouse gas emissions and achieving net zero, are published here for each technology:

- EKO Pte Ltd: [UK Carbon Reduction Plan Guidance](#)
- [MyCardium AI carbon reduction policy](#)
- Ligence did not disclose its Carbon Reduction Plan for UK carbon emissions or its net zero commitment.
- Ultromics did not disclose its Carbon Reduction Plan for UK carbon emissions or its net zero commitment.

3 Committee discussion

The diagnostics advisory committee considered evidence on artificial intelligence assisted echocardiography analysis and reporting to support the diagnosis and monitoring of heart failure from several sources. This included evidence submitted by MyCardium, Ultromics, Ligence UAB and EKO Pte Ltd, a review of clinical and cost evidence by the external assessment group (EAG), and responses from stakeholders. Full details are available in the [project documents for this guidance](#).

Guidance – Artificial intelligence (AI)-assisted echocardiography analysis and reporting to support the diagnosis and monitoring of heart failure

Page 6 of 18

Issue date: February 2026

© NICE 2026. All rights reserved. Subject to [Notice of rights](#).

The condition

- 3.1 Heart failure is common. It affects over 1 million people in the UK, with 200,000 new diagnoses annually and 800,000 people with the condition on GP registers ([British Heart Foundation Statistics Factsheet - UK, 2026](#)). Heart failure is when the heart cannot pump blood effectively because of structural or functional abnormalities. This may develop gradually (chronic, often linked to hypertension or diabetes) or suddenly (acute, for example after myocardial infarction, arrhythmia, infection, or uncontrolled hypertension). Heart failure significantly impacts quality of life and can lead to disability and early death. Around 80% of heart failure diagnoses in England happen in hospital, despite 40% of patients having symptoms that could have prompted earlier assessment (British Heart Foundation Statistics Factsheet - UK, 2026). Heart failure is classified by left ventricular ejection fraction (LVEF) measured using echocardiography. Heart failure with:
- preserved ejection fraction (HFpEF) is defined as an LVEF of 50% or more
 - reduced ejection fraction (HFrEF) is defined as an LVEF of 40% or less
 - mildly reduced ejection fraction (HFmrEF) is an intermediate category with an LVEF of 41% to 49%.

Current practice

- 3.2 For both acute and chronic onset of heart failure, initial clinical assessment includes a blood test for N-terminal pro-B-type natriuretic peptide (NT-proBNP). If NT-proBNP thresholds are exceeded, then this is followed by confirmatory diagnosis with transthoracic echocardiography (TTE). TTE is the primary diagnostic tool for heart failure and is used in around 87% of diagnoses ([National Institute for Cardiovascular Outcomes Research \[NICOR\] National Heart Failure Audit, 2025](#)). In the NHS

it is usually done in secondary care by a specialist cardiac physiologist. TTE detects abnormalities and defects in the heart's chambers and valves and provides measurements of blood flow and the heart's pumping ability. When an echocardiogram detects abnormal ejection fraction, abnormalities in the motion of the heart wall, or hypertrophy, this can indicate heart failure. Diagnosis with TTE determines whether heart failure is left or right sided, or biventricular. The TTE process typically takes between 45 and 60 minutes. After TTE, people have an appointment with a specialist for clinical assessment and review of the TTE findings, and confirmation of diagnosis.

Unmet need

- 3.3 There is a significant backlog for echocardiography in England, with waiting lists rising to 235,476 people in June 2025 ([NHS England, 2025](#)). NICE quality standards require 90% of suspected heart failure referrals to be investigated using echocardiography (typically TTE). But only half of hospitals meet this target ([NICOR National Heart Failure Audit, 2025](#)). Although suspected heart failure cases should be seen within 6 weeks, only about two thirds of referrals meet this standard ([NHS England, 2024; PDF only](#)).
- 3.4 This can lead to delays in diagnosis and treatment, and in some cases people are unable to get a TTE appointment. Delays to heart failure diagnoses may lead to poorer health outcomes and increased use of healthcare resources.
- 3.5 Artificial intelligence (AI) technologies that aid the interpretation and quantification of echocardiography images, and automate report generation, have the potential to reduce TTE procedure times. This could possibly lead to more appointments being available and a reduction in echocardiography waiting lists, which could lead to a reduction in time to diagnosis.

Clinical effectiveness

Patient considerations

- 3.6 Patient experts explained that symptoms of heart failure can develop gradually and become debilitating. Symptoms may include severe breathlessness and fatigue, and water retention leading to swollen ankles. These symptoms lead to a sedentary lifestyle because they prevent people from performing daily activities and taking part in social activities such as sports. Symptoms may also disrupt education, training and work activities and opportunities. People with heart failure may develop other health complications over time including chest infections, heart palpitations, loss of appetite and weight loss. Patient experts explained that long waiting times for echocardiography and subsequent treatment leads to symptoms worsening and potentially the need for emergency surgery. They highlighted the need for people to move faster through the diagnostic pathway and so have the most appropriate treatment as soon as possible. Reduced waiting times would enable earlier diagnosis and treatment, which would improve health outcomes.

Evidence base

- 3.7 There were 19 studies included in the clinical-effectiveness review. Most studies assessed Us2.ai (11 studies). There were 3 studies each on EchoConfidence and EchoGo Heart Failure. Two studies assessed Ligence Heart. Most of the evidence was on aspects of diagnostic performance, with limited evidence on procedure time and clinical outcomes. There was no diagnostic accuracy evidence for Ligence Heart.
- 3.8 The committee noted that there was a lack of UK-based and real-world data, with most studies being based outside the UK or in controlled settings. Some studies were based in single centres and

had single operators. Most of the evidence was from retrospective studies and some was from unpublished reports. The committee considered other limitations in the evidence base such as the exclusion of complex cases and poor-quality echocardiogram images. The committee concluded that the available evidence had limited generalisability to UK clinical practice in the NHS.

Diagnostic accuracy

- 3.9 Diagnostic accuracy was reported for 3 of the technologies (EchoConfidence, EchoGo Heart Failure and Us2.ai) across 5 studies. Two of these studies were UK based (FEATHER study interim analysis for EchoConfidence, and Campbell et al. [2025] for Us2.ai). The committee noted that the AI technologies generally show good performance for detecting abnormalities indicative of heart failure and related parameters when compared to human measurements or multiparametric clinical scoring tools. The committee also noted that standard TTE and clinical assessment without AI was both the comparator and reference standard. So, the AI technologies could not demonstrate superiority to standard TTE, only equivalence.
- 3.10 A clinical expert explained that in practice, TTE is only one component of the overall diagnostic process, which also takes into account blood tests such as NT-proBNP, clinical assessment and multiparametric clinical scoring tools. Also, heart failure symptoms may be caused by other conditions such as amyloidosis, valve disease, pulmonary hypertension and pericardial constriction, which some AI technologies may miss. The committee concluded that overall, the diagnostic performance of the AI technologies was generally good, but real-world diagnostic performance when used as intended in UK clinical practice is uncertain.

Procedure time, waiting times and system impact

- 3.11 The committee considered that a reduction in procedure time, leading to more appointments being available and a reduction in waiting lists was a key unmet need in echocardiography clinics. A clinical expert explained that around a quarter of people with suspected heart failure are still waiting to have their echocardiogram at the time of the initially scheduled specialist clinical review appointment. So, these people either reschedule or do not attend the specialist clinical review appointment. Evidence on the impact of AI technologies on TTE procedure time was available for 2 of the technologies, EchoConfidence and Us2.ai. For EchoConfidence, the interim analysis from the FEATHER study reported that in a UK community care setting, the AI technology reduced the mean time for analysis of echocardiographic parameters from 553 seconds and 587 seconds for 2 human readers to 3.2 seconds. For Us2.ai, 2 studies based in Japan reported data on procedure time ([Hirata et al. \[2024\]](#) and [Sakamoto et al. \[2025\]](#)). The committee noted that Hirata et al. reported an overall time saving of 524 seconds, similar to that reported for EchoConfidence in the FEATHER study.
- 3.12 The EAG explained that there was insufficient detail in the studies to be certain about what the time saving related to. Clinical experts considered that although these time savings were promising, it was unclear whether they would translate into routine clinical practice in the NHS. The committee said that the available evidence on procedure time was limited and not generalisable to NHS clinical practice in secondary care. It concluded that it is uncertain whether reported time savings would translate into additional appointments and reduced waiting times.
- 3.13 Clinical experts said that in some instances using the AI technologies may lead to increases in procedure times because of healthcare professionals needing to check and review the AI

findings and potentially intervene. Introducing any further delays could cause harm because of delays in diagnosis and treatment.

Community and primary care settings

- 3.14 Clinical experts explained that there is an increasing trend for TTE procedures to be done in community settings, such as community diagnostic centres. (See the [NHS England webpage on community diagnostic centres](#).) Clinical experts highlighted that in community or primary care, use by healthcare professionals such as GPs, community nurses and pharmacists, would affect diagnostic performance. So, they said that the available evidence would not be generalisable to these potential future use cases.
- 3.15 The EAG highlighted that the external assessment report presented evidence from studies that may be relevant to use of the AI technologies in community or primary care. This included studies that looked at novice operator performance compared with standard expert-led echocardiography ([Huang et al. 2024a](#)). The studies demonstrated some potential for the technologies to be used in community care settings. The committee concluded that although there was some evidence to support use of the AI technologies in community or primary care settings, these use cases were outside the scope of the assessment. It further concluded that more research was needed on diagnostic performance and procedure times, when used by different operators of varying levels of experience in different care settings (see [section 1](#)).

Cost effectiveness

Conceptual model structure and assumptions

- 3.16 The EAG constructed a conceptual Markov model with a 1-year time horizon to capture the impact of reduced waiting times from shorter TTE durations when using AI technologies. The EAG

explained that because of the limited evidence base, it needed to make multiple assumptions to develop the conceptual model.

- 3.17 AI technologies would be used as an adjunct to standard TTE, and a specialist clinical assessment is needed to diagnose heart failure. So, diagnostic accuracy was assumed to be unaltered when using the AI technologies. Clinical experts agreed that this was a reasonable and safe assumption. Most people (79%) were assumed to enter the model in an acute episode, while the remainder entered from the waiting list. A clinical expert suggested that in practice most people would likely enter the care pathway from primary care. The EAG explained that the sensitivity analysis demonstrated that this had no substantial impact on cost effectiveness. But it noted that if the AI technologies did reduce waiting times, then this could potentially reduce hospitalisation.
- 3.18 The committee noted that downstream benefits of earlier diagnosis were not modelled because of the lack of evidence in this area, and uncertainty around current waiting times. It said that the conceptual model and assumptions were appropriate for assessing the potential cost effectiveness of the AI technologies. But it concluded that the limited evidence base and number of assumptions meant that the economic results were highly uncertain.

TTE appointment time

- 3.19 The time for a standard TTE appointment was 45 minutes, based on clinical opinion, and the EAG assumed 10 TTE appointments per day. In the base case for EchoConfidence and Us2.ai, the EAG assumed a reduced appointment time of 36 minutes (based on a reduced TTE procedure time). For EchoConfidence this was based on data from the interim FEATHER study and for Us2.ai it was based on data from Hirata et al. (2024). This resulted in an increase in the number of TTE appointments to 12 per day. The

model assumed that in standard cardiology clinics (that is, not one-stop diagnostic centres) the waiting time for specialist clinical review remained unchanged. The committee recalled the uncertainty in the evidence on procedure time saving and its limited generalisability (see [sections 3.11 and 3.12](#)). It concluded that further research was needed on real-world time savings when used in NHS clinical practice to better inform any future economic models.

Technology costs

3.20 The committee noted that the cost per scan for each technology was made up of a number of separate components, including:

- software cost per scan
- system set-up and training costs
- IT support costs
- staff costs.

The conceptual model base-case analysis included a total cost per scan of £4.26 for EchoConfidence and £7.70 for Us2.ai.

Plausibility of cost effectiveness

3.21 In the EAG's base-case analysis, EchoConfidence was cost saving compared with standard care (cost difference of -£3.14 and quality-adjusted life year [QALY] difference of 0.0005). This was mainly because of the reduced staff time per TTE offsetting its cost per use. The base-case analysis for Us2.ai showed that it was more effective but more costly than standard care (cost difference of £0.92 and QALY difference of 0.0005), with an incremental cost-effectiveness ratio (ICER) of £1,674 per QALY gained. This was because in the conceptual model, cost savings from the shorter procedure time when using Us2.ai were not sufficient to fully offset the earlier treatment costs incurred when more people get an

earlier diagnosis. The committee understood that these results were uncertain because of the limitations in the evidence base for procedure time savings (see [sections 3.11 and 3.12](#)).

- 3.22 EchoConfidence remained cost saving in all sensitivity analyses, while Us2.ai had ICERs ranging from £93 to £2,684 per QALY gained. The results for both technologies were most sensitive to the impact of waiting-time reduction and the proportion of people diagnosed in a one-stop diagnostic clinic. While the results for Us2.ai were also sensitive to the proportion of inpatients having TTE, and the grade of staff delivering TTE. The committee noted that a key driver of cost effectiveness in the modelling was reduced procedure time using the AI technologies (which could potentially increase throughput, thereby reduce waiting lists and time to diagnosis). It concluded that while procedure time was key to the results of the model, the available evidence for reduced procedure time lacked robustness and generalisability, so the cost-effectiveness estimates were uncertain.

Equality considerations

- 3.23 The committee noted that heart failure is more prevalent in people from some ethnic backgrounds including South Asian and Black British populations. The committee agreed that it is important to understand how and on what datasets the AI technologies have been trained. A lack of external validation in UK or similar populations may limit suitability of the technologies for UK practice and pose clinical risks. The EAG explained that there is a lack of consistency in how external validation cohorts are reported in the studies. It also noted that there was some UK cohort validation data for EchoConfidence. The committee concluded that more research is needed to understand how the AI technologies are trained and more transparency is needed around the populations used for validation (see [section 1](#)).

Adverse events

- 3.24 The studies in the clinical review did not report any adverse events. The committee noted that this could reflect the retrospective nature of the evidence base and the lack of real-world evidence. The committee considered that although there is no direct evidence to indicate that there is any risk of harm from using the AI technologies, the clinical risk is uncertain given the current limited evidence base.

Evidence gaps

- 3.25 The committee considered the evidence gaps highlighted in the external assessment report. These included:
- Procedure time, including:
 - time taken for automation of echocardiographic measurements
 - time taken for automation of TTE report
 - overall procedure time
 - Clinical outcomes, including:
 - time to heart failure diagnosis
 - time to treatment initiation
 - patient-reported health-related quality of life
 - Validation in cohorts representative of the UK population, including:
 - diagnostic accuracy, interchangeability, agreement and correlation with human measurements
 - diagnostic performance in detecting and classifying heart failure.
 - Acceptability of the AI technologies, including:
 - ease of use
 - confidence in accuracy of automation (need for human review)

- feasibility of implementation in different settings (primary and secondary care) with staff of varying skill levels.
- Adverse events, including:
 - inaccurate measurements or incorrect diagnoses
 - AI failure rate
- Care pathway uncertainties needed for future model development, including downstream treatment costs and utilities associated with treated and untreated heart failure.

The committee concluded that further research was needed on these areas (see [section 1](#)).

Ongoing studies

3.26 The committee noted that there are a number of ongoing studies on the AI technologies. These include 2 RCTs (TARTAN-HF and SYMPHONY-HF) on Us2.ai investigating the use of AI-assisted echocardiography as part of screening strategies. For EchoConfidence, the FEATHER study is a double-blind evaluation of AI for heart failure diagnosis and stratification on unselected consecutive patients referred for evaluation to community cardiology services. There are also 2 ongoing studies on Ligence Heart. The committee concluded that although there are several ongoing studies that may provide further evidence on AI-assisted echocardiography analysis and reporting to support the diagnosis and monitoring of heart failure in the NHS, they will not address all the identified evidence gaps (see [section 3.25](#)).

4 Committee members and NICE project team

This topic was considered by [specialist committee members appointed for this topic](#) and [NICE's diagnostics advisory committee](#), which is a standing advisory committee of NICE.

Committee members are asked to declare any interests in the technology to be evaluated. If it is considered there is a conflict of interest, the member is excluded from participating further in that evaluation.

The [minutes of each committee meeting](#), which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Chair

Thomas Clutton-Brock

Chair, diagnostics advisory committee

NICE project team

Each evaluation is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the evaluation), a technical adviser, a project manager and an associate director.

Iain Willets and Simon Webster

Technical leads

Amy Crossley

Technical adviser

Donna Barnes

Project manager

Lizzy Latimer

Associate director

ISBN: [\[to be added at publication\]](#)