Evidence generation plan for HTG10877 Technologies for diagnosing endometriosis in primary care

June 2026

1 Purpose of this document

NICE's assessment of EndoSure and Endotest recommends that more evidence is generated while they are being used in the NHS. The other technology that was assessed can only be used in research and is not covered in this plan.

This plan outlines the evidence gaps and what data needs to be collected for a NICE review of the technologies again in the future. It is not a study protocol but suggests an approach to generating the information needed to address the evidence gaps. Evidence generated through other study approaches will also be considered. For assessing comparative treatment effects, well-conducted randomised controlled trials are the preferred source of evidence.

The companies are responsible for ensuring that data collection and analysis takes place. Financial support for evidence generation may be available through the NIHR Invention for Innovation product development award. Academic support may be available through groups such as the NIHR HealthTech Research Centres.

NICE will withdraw the guidance if the companies do not meet the conditions in section 4 on monitoring.

After the end of the evidence generation period (3 years), the companies should submit the evidence to NICE in a format that can be used for decision making. NICE will review all the evidence and assess whether the technologies can be routinely adopted in the NHS.

2 Evidence gaps

This section describes the evidence gaps, why they need to be addressed and their relative importance for future committee decision making.

The committee will not be able to make a positive recommendation without the essential evidence gaps (see section 2.1) being addressed. The company can strengthen the evidence base by also addressing as many other evidence gaps (see section 2.2) as possible. This will help the committee to make a recommendation by ensuring it has a better understanding of the patient or healthcare system benefits of the technology.

2.1 Essential evidence for future committee decision making

Diagnostic accuracy in an NHS primary care setting

There is no data on the diagnostic accuracy of the technologies in NHS primary care. It is important to understand how the technologies perform compared with current NHS practice in primary care. The committee agreed that the different populations and technicalities of operating the tests in primary care compared with secondary care could affect the technologies' diagnostic performance. The committee also highlighted that clearer reporting of failure rates is needed. This could impact diagnostic performance and the need for further testing.

Impact on diagnostic pathway and resource use

More information is needed on how using the technologies would affect resource use during and after implementation to help the committee understand the technologies' long-term resource-use impacts. For example, the additional time needed to carry out a test may result in fewer people being seen in primary care than in standard care. Collecting evidence on this will help the committee understand how using the technologies will affect care in the NHS. Key areas that will help to address this evidence gap are:

- the time and associated resources used to perform the tests by reporting practitioners, including additional reviews by specialists or resources, such as an additional room for doing the test
- the number of people referred for further testing (including imaging or laparoscopy)
- the number of people having medication for endometriosis
- time to treatment

- time to presumptive diagnosis
- long-term resource-use costs, such as number and extent of treatments and number of hospital appointments or emergency department visits
- the costs of technology acquisition, implementing (including training), integrating, maintaining and using the technologies

- cost savings associated with avoided downstream tests.

Patient outcomes and experience

More evidence on patient-related outcome and experience measures, for example intervention uptake, satisfaction, quality of life, acceptability and preferences, will help the committee assess the technologies' real-world uptake and support future cost-effectiveness modelling.

2.2 Evidence that further supports committee decision making

Clinical effectiveness in different subgroups

The impact of the technologies on clinical effectiveness and resource and service use in different subgroups is unknown. The committee discussed that subgroup analyses in the following subgroups would be useful:

- young people
- women, trans men and non-binary people who:
  - are approaching menopause or who have experienced menopause
  - are from ethnic minority backgrounds
  - find transvaginal ultrasound unacceptable, or for whom transvaginal ultrasound is not suitable.

3 Approach to evidence generation

3.1 Evidence gaps and ongoing studies

There are no ongoing or planned studies that evaluate the diagnostic accuracy of these technologies in primary care. The ENDOBEST study (NCT06794424), ADOmiRNA study and a planned NHS pilot study may address some of the gaps for

Endotest around resource use, subgroup analyses in young people and diagnostic accuracy by the end of 2027. But these studies are all in secondary care. Some diagnostic accuracy and laparoscopy rate data may be provided for EndoSure by the ADDEND study (ISRCTN83220665) and a planned NHS pilot.

Table 1 summarises the evidence gaps and ongoing studies that might address them. Information about evidence status is derived from the external assessment group's report; evidence not meeting the scope and inclusion criteria is not included. The table shows the evidence available to the committee when the guidance was published.

Table 1 Evidence gaps and ongoing studies

Evidence gap	EndoSure	Endotest
Diagnostic accuracy in NHS primary care setting	No evidence	No evidence
Impact on diagnostic pathway and resource use	No evidence	No evidence Ongoing study
Patient outcomes and experience	Limited evidence	Limited evidence Ongoing study
Clinical effectiveness in different subgroups	No evidence Ongoing study	No evidence Ongoing study

3.2 Data sources

No existing data registries or audits were identified as potential sources of real-world data that may address the evidence gaps. Some existing data sources, such as the NHS England Secure Data Environment (SDE) service and Hospital Episode Statistics (HES), could support evidence generation by providing long-term outcomes if they are linked to primary data collection. NICE's real-world evidence framework provides detailed guidance on assessing the suitability of a real-world data source to answer a specific research question.

3.3 Evidence collection plan

The suggested approach to addressing the evidence gaps for the technologies is a cross-sectional diagnostic accuracy study and a real-world comparative cohort study.

Cross-sectional diagnostic accuracy study

A cross-sectional diagnostic accuracy study would assess the accuracy of the technology against a reference standard (MRI). Reported accuracy should include sensitivity, specificity, and negative and positive predictive values.

The committee noted that the gold standard for these tests is laparoscopy. But it would be unethical to use this when there is a negative test result to confirm true negativity. Given the lengthy waiting list for laparoscopy for diagnosis and treatment, the committee discussed that using laparoscopy for evidence-generation purposes would not be an appropriate or ethical use of NHS resources. The committee agreed that using MRI would be an appropriate reference standard instead. It also heard that access to MRI may be limited. Latent class analyses or other statistical analyses may be appropriate when a gold standard is unsuitable, if multiple test results (and clinical history data) are readily available.

Real-world comparative cohort study

This approach would follow an intervention arm and a control arm and compare their outcomes. This design would allow the clinical impact of the technologies and the resource use associated with their implementation in primary care to be assessed. Qualitative data could be generated through appropriate methods like surveys, focus groups or interviews, as highlighted in NICE's real-world evidence framework. This could include reported outcomes (acceptability, usability and preferences) from people using the service.

Despite consistent eligibility criteria, non-random assignment to interventions can lead to confounding bias, complicating interpretation of the treatment effect. So, approaches should be used that balance confounding factors across comparison groups, for example using propensity score methods. To achieve this robustly, data collection will need to include prognostic factors related to both the intervention delivered and patient outcomes. These should be defined with input from clinical specialists. Incomplete records and demographically imbalanced groups can lead to bias if they are unaccounted for.

Data collection should follow a predefined protocol. Quality assurance processes should be put in place to ensure the integrity and consistency of data collection. NICE's real-world evidence framework provides guidance on planning, doing and reporting real-world evidence studies. It also provides best-practice principles for a robust design of real-world evidence when assessing comparative treatment effects using a prospective cohort study design.

3.4 Data to be collected

Diagnostic accuracy study

- Patient characteristics, including age, ethnicity, menopausal status and conditions that may complicate imaging
- Diagnostic accuracy, including sensitivity, specificity, and positive and negative predictive values
- Prevalence of disease
- Multiple pathology rate
- Test failure rate
- Adverse events

Real-world comparative cohort study

- Patient characteristics as above for diagnostic study
- Details of the technology
- Health-related quality-of-life data (baseline and after the intervention), including symptom reduction, ideally collected using the EQ-5D-3L questionnaire
- Detail of the technology (software name, version and configuration settings)
- Image details (including anatomical location, projection when considering X-rays and manufacturer of CT or X-ray machine)
- Time from symptom onset to referral, presumptive diagnosis and definitive diagnosis
- Time from presentation in primary care to referral, presumptive diagnosis and definitive diagnosis
- Diagnostic laparoscopy rate

- rASRM (revised American Society for Reproductive Medicine) rating, where laparoscopy is available
- Time to treatment
- Results of any additional downstream tests
- Number of primary-care consultations
- Number of referrals and referral requests
- Number of hospital appointments and emergency department visits
- Health-related quality-of-life data (ideally EQ-5D-5L and Endometriosis Health Profile-30 questionnaire, and including anxiety)
- Imaging details
- Multiple pathology rate
- Adverse events
- Test failure rates
- Patient-related experience measures, ideally intervention uptake, satisfaction, acceptability and preferences
- Costs associated with the technologies, including acquisitions, implementation (for example, cost of staffing, an additional room, training and transportation of samples) and additional downstream testing.

3.5 Evidence generation period

This will be 3 years to allow for setting up, implementing the test, data collection, analysis and reporting.

3.6 Following best practice in study methodology

Following best practice in conducting studies is paramount to ensuring the reliability and validity of the research findings. Adherence to rigorous guidelines and established standards is crucial for generating credible evidence that can ultimately improve patient care. The NICE real-world evidence framework details some key considerations.

Within the context of an early-use assessment a key factor to consider as part of the informed consent process is to ensure that patients (and their carers, as appropriate) understand that data will be collected to address the evidence gaps identified in

section 2. Where applicable this should take account of NICE guidance about shared decision making.

4 Monitoring

NICE will contact the companies:

- within 6 months of publication of this plan to confirm agreements are in place to generate the evidence
- annually to confirm that the data is being collected and analysed as planned.

The companies should tell NICE as soon as possible of anything that may affect ongoing evidence generation, including:

- any substantial risk that the evidence will not be collected as planned
- new safety concerns
- the technology significantly changing in a way that affects the evidence generation process.

If data collection is expected to end later than planned, the companies should contact NICE to arrange an extension to the evidence generation period. NICE reserves the right to withdraw the guidance if data collection is delayed, or if it is unlikely to resolve the evidence gaps.

5 Minimum evidence standards

During the evidence generation period, new technologies may become available. This section summarises the minimum evidence requirements that a new technology would need to meet to be considered in the NICE evaluation after the evidence generation period.

The committee noted that there is some evidence showing that the technologies have good diagnostic accuracy (sensitivity and specificity above 90%). But this is uncertain primarily because of the lack of data in NHS primary-care settings. It heard that the technologies may improve quality of life by reducing time to diagnosis and treatment, but evidence for this was also limited.

For new technologies, the committee indicated that it may, in future, be able to recommend technologies in this topic area that have UK-based evidence for:

- non-inferiority of the technology compared with standard care in terms of diagnostic accuracy
- user engagement with the technology, including satisfaction, acceptability and preferences
- cost savings resulting from resource use associated with the technologies.

Companies can strengthen the evidence base by having evidence for diagnostic accuracy and uptake rates in different subpopulations.

6 Implementation considerations

The following considerations around implementing the evidence generation process have been identified through working with system partners.

6.1 Implementing the guidance

- The committee heard that diagnosis is often delayed and access to care is unequal for certain population groups, including:
  - people with female reproductive organs who do not identify as women, including trans men and non-binary people
  - young people
  - women, trans men and non-binary people:
    - from ethnic minority backgrounds (frequent misdiagnosis or dismissal of pain, or cultural barriers to discussing menstrual health)
    - with learning disabilities or who have difficulties communicating their symptoms
    - who find transvaginal ultrasound unacceptable
    - who do not have access to healthcare professionals in primary care with high levels of expertise in transvaginal ultrasound.
- The technologies may offer additional benefit to women, trans men or non-binary people with a needle phobia.

- Reporting intervention-related adverse events (for example, worsening pain) is essential to assess any risk associated with the use of the technologies in the NHS.

6.2 Implementing data collection

- The technologies are currently only available in the private healthcare sector.
- It is likely that evidence generation within the NHS will cause additional resource and staffing burden. The process is most likely to succeed with dedicated research staff to reduce the burden on NHS staff.
- Practical requirements, such as fasting, medication cessation, long appointment times and the need for a quiet clinical environment, may result in missed appointments and incomplete datasets.

ISBN: [to be added at publication]