Technoleg Iechyd Cymru
Health Technology Wales

Evidence Appraisal Report [1]

Capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer

Appraisal summary

Why did Health Technology Wales (HTW) appraise this topic?

Gastro-oesophageal reflux disease (GORD), when acid from the stomach leaks into the oesophagus, is a chronic condition and can cause changes in the cells lining the oesophagus. This may develop into a condition called Barrett's oesophagus, which can be a precursor to oesophageal cancer. The prognosis of oesophageal cancer is poor as the early stages may be asymptomatic and patients often do not present until it is advanced. Patients with GORD may be offered upper gastrointestinal tract endoscopy to check for Barrett's oesophagus or signs of cancer and those with confirmed Barrett's oesophagus are recommended to receive regular endoscopic surveillance.

The prevalence of Barrett's oesophagus is low and rates of progression to oesophageal cancer are also very low, meaning most patients sent for endoscopy will not have these conditions. Capsule sponge devices are a non-endoscopic way of collecting cells from the oesophageal lining that can then be tested for biomarkers of Barrett's oesophagus or oesophageal cancer. The potential advantages of capsule sponge testing include improved comfort for patients compared with endoscopy and the ability to be performed in primary care settings. The intended placement of capsule sponge testing in the patient pathway is as a triage test to determine whether endoscopy is required, how urgently, and how frequently for surveillance. The use of capsule sponge testing could therefore lead to reduced pressure on endoscopy services by only referring, and then prioritising, those that show biomarkers of Barrett's oesophagus or malignancy.

This topic was suggested to HTW by the clinical lead for the National Cancer Recovery Programme.

What evidence did HTW find?

We searched for evidence that could be used to answer the review question: what is the clinical effectiveness and cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer?

[1] Cyfieithu dogfennau HTW wedi'u cyhoeddi o'r Saesneg i'r Gymraeg
Translation of published technical HTW documents from English into Welsh

We identified four clinical guidelines and 17 studies: one health technology assessment, one randomised controlled trial, 12 observational studies, and four economic studies (one of which was one of the 12 observational studies).

The evidence suggests the diagnostic accuracy of capsule sponge devices with trefoil factor 3 (TFF3) testing for proactive screening of Barrett's oesophagus in those with chronic reflux is good, with high sensitivity and specificity. Where reported, positive predictive value (PPV) is low whilst negative predictive value (NPV) is high. For case finding of Barrett's oesophagus using capsule sponge testing with TFF3, p53, and cellular atypia, detection rates suggest potentially high rates of false positives but, importantly, very low rates of false negatives as well. The diagnostic accuracy for case finding also appears to be good, with sensitivity above 90%, and PPV and NPV findings supporting the findings from detection rates. Capsule sponge testing with p53 and cellular atypia for Barrett's oesophagus under surveillance also shows good accuracy for detecting dysplasia or cancer, however, the two biomarkers in isolation may not be sufficiently accurate. The evidence also suggests that using capsule sponge testing, in combination with assessing clinical risk factors, is effective in risk stratifying Barrett's oesophagus patients.

Time to diagnosis and time to treatment were reported in one evaluation of real-world data, with no comparisons to standard care. No data on health-related quality of life were identified. The safety of capsule sponge devices appears to be high, and the incidence of adverse events is low.

Most of the evidence was related to the device Cytosponge, however, evidence is generalisable across Cytosponge and EndoSign devices but not to other non-endoscopic cell collection devices. The majority of studies involved people who were involved in the development of the examined devices, or were employees or founders of the companies that manufacture them. The lack of endoscopic biopsy results after negative capsule sponge tests means the number of true/false negative results is not known. More research is needed on the effect capsule sponge testing has on cancer outcomes and evidence comparing outcomes and patient experiences of capsule sponge testing in primary and secondary care settings would also be beneficial.

Four studies conducting a cost-utility analysis were included in the economic review. Three focused on capsule sponge devices used for initial diagnostic screening, and one focused on using the capsule sponge device for surveillance. Only one study took the perspective of the UK NHS, and results of their base case analysis estimated cost savings of £422 with a reduction of 0.0041 quality-adjusted life years (QALYs) per patient triaged using Cytosponge compared with endoscopy alone. This corresponded to a net monetary benefit of £339, at a cost-effectiveness threshold of £20,000 per QALY, and the study concluded that endoscopy-only screening was not cost effective compared to Cytosponge. However, potentially serious limitations of this study were identified including possible biases in the data used to inform the diagnostic pathway and comparator arm, as well as uncertainties in how representative the clinical data is to the modelled population.

We conducted a new cost-utility analysis from the NHS Wales perspective to estimate the cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer in people with chronic reflux, compared to endoscopic biopsy. Over a lifetime horizon, the results estimated that use of Cytosponge in primary care, followed by endoscopic biopsy in those with a positive result, is expected to reduce costs by [redacted] per patient with a loss of 0.02 QALYs, corresponding to an incremental cost-effectiveness ratio (ICER) of [redacted], representing the cost savings per QALY lost. This is above the £20,000 cost-effectiveness threshold, indicating that the use of Cytosponge is cost effective in the context where the intervention is less costly and less effective than the comparator. Probabilistic sensitivity analysis suggested a 65.8% probability of cost effectiveness at this threshold. Capsule sponge sensitivity, age and Barrett's oesophagus prevalence were identified as influential drivers of cost effectiveness. Scenarios exploring capsule sponge delivery in secondary and community-based

care settings, as well as the use of the Endosign device, had minimal impact on health economic outcomes, with no change in cost effectiveness conclusions. However, conclusions did change in scenarios exploring younger populations and where age-related utility decline is not considered.

Real-world evidence and feedback from subject experts indicated introducing capsule sponge testing could significantly reduce demand on endoscopy services, which are currently under pressure. This testing could also ensure those most in need have quicker access to endoscopic investigation. However, safety netting and clear patient pathways with defined eligibility criteria would also be needed to ensure patients do not receive unnecessary investigations or inappropriate discharges. Introduction of capsule sponge testing could also address equity of access issues both within Wales and across the UK.

What was the outcome of HTW's appraisal?

HTW is a national body working to improve quality of care in Wales. We collaborate with partners across health, social care, and industry to issue independent guidance that informs commissioning within Wales health and social care. We are supported by an Assessment Group, who ensure our work adheres to high standards of methodological and scientific rigour, and an Appraisal Panel, who consider evidence within the Welsh context and produce HTW guidance. More details on our appraisal process, the assessment group, and the appraisal panel can be found on the HTW website.

In this case, the HTW Assessment Group considered the evidence presented in this Evidence Appraisal Report (EAR069) and concluded there was sufficient evidence for the development of guidance. Please refer to the HTW website for full guidance details.

Evidence Appraisal Report 069 follows below and provides full details for this topic. More comprehensive details of the HTW Guidance and HTW Appraisal Panel considerations can be found on the HTW website.

1. Purpose of the Evidence Appraisal Report

This report aims to identify and summarise evidence that addresses the following question: what is the clinical effectiveness and cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer?

Evidence Appraisal Reports are based on rapid systematic literature searches, with the aim of identifying the best published evidence on the effectiveness and cost-effectiveness of health and social care technologies and models of care and support. Researchers critically evaluate this evidence. The draft Evidence Appraisal Report is reviewed by experts and by Health Technology Wales multidisciplinary advisory groups before publication.

2. Context

Gastro-oesophageal reflux disease (GORD) is when acid from the stomach leaks into the oesophagus, causing symptoms such as heartburn (SHTG 2023). This is a chronic condition and chronic acid reflux can cause changes in the cells lining the oesophagus. This may develop into a condition called Barrett's oesophagus, when the squamous cells of the oesophagus are replaced with columnar cells, which can be a precursor to oesophageal cancer (NICE 2023a). Patients with GORD may be offered upper gastrointestinal tract endoscopy to check for Barrett's oesophagus or signs of cancer.

The prevalence of chronic reflux is uncertain due to variations in definitions, however there are estimated prevalences of between 10% to 30% of adults in developed countries, 8.8% to 25.9% in Europe, and a UK incidence of approximately 5 per 1,000 person-years (NICE 2023b). Barrett's oesophagus has an estimated prevalence of 1.5% to 2.5% of the adult population in the UK and 10% to 15% of people with GORD will develop Barrett's oesophagus (NICE 2023b, SHTG 2023). A recent systematic review and meta-analysis estimated the prevalence of Barrett's oesophagus in Europe was 8.6% amongst those with GORD and 2.1% in those without GORD (Saha et al. 2024). Risk factors for Barrett's oesophagus include male sex assigned at birth, increasing age, being overweight, white ethnicity, and family history of Barrett's oesophagus. The rate of progression from Barrett's oesophagus to cancer is low in the UK at approximately 1% per year and between 3% to 13% over their lifetime (SHTG 2023).

The prognosis of oesophageal cancer is poor as the early stages may be asymptomatic and patients often do not present until it is advanced. Therefore, the standard of care for people with confirmed Barrett's oesophagus is being offered endoscopic surveillance of the upper gastrointestinal tract, as per NICE guideline NG231 (NICE 2023a). Recommendations from NICE state that high-resolution, white light endoscopic surveillance with biopsy should be offered every two to three years for people with long-segment (3 cm or longer) Barrett's oesophagus, or every three to five years for people with short-segment (less than 3 cm) Barrett's oesophagus with intestinal metaplasia. Evidence assessed by NICE showed that endoscopic surveillance of Barrett's oesophagus led to a 30% reduction in mortality compared to those who did not receive surveillance. NICE recommends tailoring the frequency of surveillance, within the intervals shown above, based on the individual's risk factors for oesophageal cancer. It is also recommended that endoscopic surveillance is not offered to people with short-segment Barrett's oesophagus without intestinal metaplasia, as long as the diagnosis has been confirmed at two endoscopies. However, recent studies have shown that regular surveillance of Barrett's oesophagus does not lead to improvements in survival compared to at-need endoscopy (Old et al. 2025).

3. Health technology

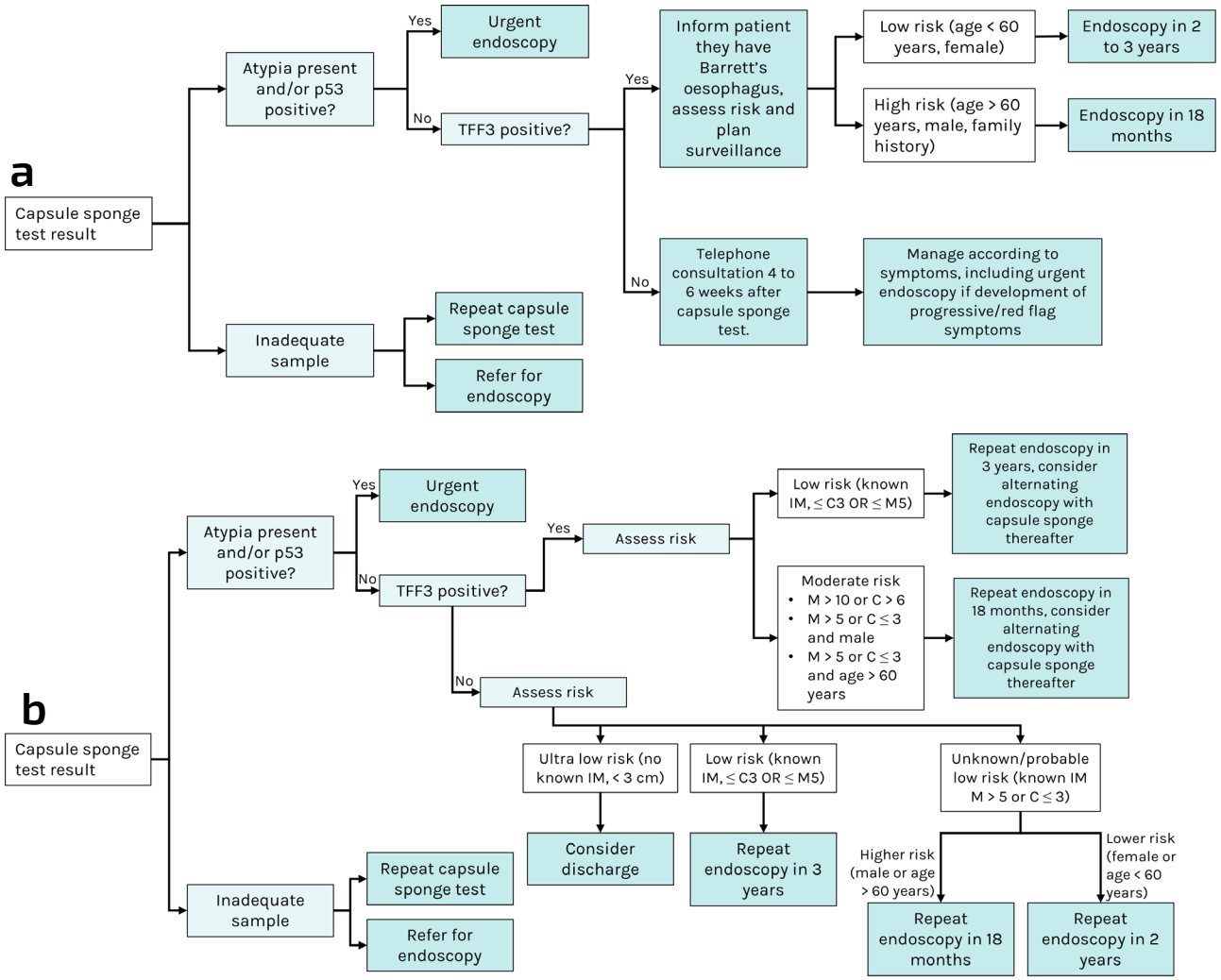
Barrett's oesophagus surveillance and investigation of people with GORD adds to the demand on endoscopy services and, as indicated by the prevalences above, the majority of patients will be found not to have the conditions looked for. Capsule sponge devices are a non-endoscopic way of collecting cells from the oesophageal lining and, with appropriate cytological testing, biomarkers of Barrett's oesophagus or oesophageal cancer can then be looked for in the collected samples. The potential advantages of capsule sponge testing include improved comfort for patients compared with endoscopy, the ability to be performed in primary care settings, and the potential to reduce demand on endoscopy services by acting as a triage test.

Two capsule sponge devices that are CE marked and in use in the UK were identified: Cytosponge (Medtronic) and EndoSign (Cyted Ltd). Both consist of a polyester bristle-like sponge, attached to a string, that is bunched inside a vegetarian gelatine capsule that is similar in size to a pill. Patients are instructed to swallow the capsule, whilst the end of the string is retained outside of the body. Cytosponge requires a clinician to bundle the thread and place this in the patient's mouth, whereas EndoSign has the thread bundled alongside the capsule within an applicator that is used to place them both together on the back of the patient's tongue. Once in the stomach, the gelatine capsule dissolves and the bunched sponge expands. The sponge is left in place for seven (EndoSign) or 7.5 (Cytosponge) minutes, after which the sponge is pulled up from the stomach by the string and the bristles of the sponge collect cells from the oesophagus lining on the way up. The sponge is then placed in a preservation fluid and the sample is sent to a laboratory for biomarker testing. This process does not require sedation, but an anaesthetic throat spray may be administered after the capsule has been swallowed (NICE 2020, SHTG 2023).

Cyted Ltd, the manufacturer of EndoSign, also carries out the cytological analysis of samples collected with EndoSign and Cytosponge devices (Cyted Health 2024). The biomarkers examined with capsule sponge samples are Trefoil Factor 3 (TFF3), tumour protein 53 (p53), and haematoxylin and eosin staining to look for changes in cell morphology and cellular atypia (NICE 2020, SHTG 2023). Cells that are positive for TFF3 are pre-cancerous and indicate intestinal metaplasia and, therefore, likely Barrett's oesophagus. The presence of p53 indicates likely malignant changes of Barrett's oesophagus cells. Other biomarkers for use with cell collection devices, such as methylated DNA markers, have been investigated; however, as these are not used by Cyted Ltd when analysing samples, they have not been included in this appraisal. If glandular cells are not present in the sample, then it is deemed a low-confidence result as the sponge may not have reached the stomach.

The intended use of capsule sponge devices is as a triage and risk stratification test for people with GORD and heartburn symptoms, and those under regular surveillance for confirmed Barrett's oesophagus, to help determine whether upper gastrointestinal endoscopy is needed, and how urgently. Potential patient pathways, involving capsule sponge triage, are shown in Figure 1. Capsule sponge testing can be performed in both primary and secondary care settings. Based on the biomarkers identified, the appropriate next course of action for the patient can then be decided, including continuing with routine surveillance, referring for a routine endoscopy, or referring for an urgent endoscopy.

Capsule sponge testing is currently used in Betsi Cadwaladr University Health Board (BCUHB) for Barrett's oesophagus surveillance after a local implementation trial. Evaluations of capsule sponge use have also taken place in Cardiff & Vale University Health Board (CVUHB), Cwm Taf Morgannwg University Health Board (CTMUHB) and Powys Teaching Health Board (PTHB).



Based on pathways developed by the SBRI-funded Celtic Capsule project. The pathways presented here have been simplified and are demonstrative only; they do not represent intended or recommended pathways for implementation as this is outside of HTW's remit.

Abbreviations: C, circumferential length of Barrett's oesophagus (cm); IM, intestinal metaplasia; M, maximal length of Barrett's oesophagus (cm); TFF3, trefoil factor 3

Figure 1 – Potential patient pathways for patients being investigated for chronic reflux symptoms (a) and patients under surveillance for Barrett's oesophagus (b)

4. Guidelines

4.1 NICE

NICE have produced guidance on the monitoring and management of Barrett's oesophagus and stage 1 oesophageal adenocarcinoma (OAC), and recommendations on the current standard of care are discussed in Section 2 (NICE 2023a). As part of NG231, a systematic review of comparative evidence on non-endoscopic surveillance techniques was included. Three studies were included, two comparing Cytosponge to endoscopy with biopsy and one comparing balloon cytology to histology. The latter study is not relevant to this appraisal. The NICE evidence review found no evidence on clinical outcomes, and the sensitivity of the investigated devices did not meet the clinical decision threshold of 0.9 set by the committee. The specificity threshold of 0.8 was met. The quality of the evidence on diagnostic accuracy was rated as low. NICE concluded that no recommendations could be made based on this evidence.

4.2 European Society of Gastrointestinal Endoscopy

An evidence-based guideline from the European Society for Gastrointestinal Endoscopy (ESGE) included a strong recommendation on the use of swallowable, non-endoscopic cell collection devices for case finding of Barrett's oesophagus (ESGE 2023). The recommendation states that devices such as Cytosponge, combined with a cytopathological assessment and biomarker TTF3 can be used as an alternative to endoscopy for case finding of Barrett's oesophagus. They also state that the use of other non-endoscopic devices cannot be recommended yet based on low quality evidence for these. One cross-sectional study and one RCT were included in the evidence for capsule sponges and the quality of this evidence was ranked as high.

4.3 American Gastroenterological Association

The American Gastroenterological Association (AGA) produced a clinical practice update on new technologies for the surveillance and screening of Barrett's oesophagus based on an expert review (Muthusamy et al. 2022). Included in this was best practice advice that non-endoscopic cell collection devices may be considered as an option to screen for Barrett's oesophagus. This was based on one RCT, one cohort study, one case-control study and one prospective cohort study on capsule sponges with TTF3/atypia/p53 testing.

4.4 American College of Gastroenterology

The American College of Gastroenterology (ACG) carried out a 'selective literature review', in which the 'strongest evidence pertaining to each question' was selected and used to create an evidence-based guideline on the diagnosis and management of Barrett's oesophagus (Shaheen et al. 2022). This included a conditional recommendation that swallowable, non-endoscopic capsule sponge devices combined with a biomarker are acceptable alternatives to endoscopy for screening for Barrett's oesophagus in those with chronic reflux symptoms and other risk factors. This was based on one retrospective analysis of five prospective cohort analyses, one RCT, one economic analysis, four case-control studies and two prospective cohort studies. The quality of the evidence was ranked as very low.

5. Effectiveness

We searched for evidence that could be used to answer the review question: what is the clinical effectiveness and cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer?

For details on the methodology used to identify evidence for this report, refer to Appendix 1.

5.1 Overview

Evidence on the effectiveness of capsule sponge devices was extracted from one health technology assessment (HTA) (SHTG 2023) and 13 primary studies. The primary studies included one RCT (Fitzgerald et al. 2020), six prospective cohort studies (Angel et al. 2025, Chien et al. 2024a, Chien & Glen 2025, Eluri et al. 2022, Gourgiotis et al. 2025, Kadri et al. 2010, Tan et al. 2025), two case-control studies (Ross-Innes et al. 2015, Ross-Innes et al. 2017), one retrospective cohort analysis (Chien et al. 2024b), one cross-sectional study (Norton et al. 2025), and one cross-sectional study followed by a real-world prospective pilot (Pilonis et al. 2022). The report by SHTG discussed the results of a systematic review, as well as three of the primary studies mentioned above. However, after reviewing the report, only four of the studies in the systematic review were relevant to this appraisal and, therefore, we decided to extract data from individual studies directly.

Details of the included studies are shown in Tables A1, A2 and A3.

5.2 Diagnostic accuracy

5.2.1 Proactive screening of people with GORD taking medication

The Barrett's oEsophagus Screening Trial 1 (BEST1) was a prospective cohort study of people with chronic reflux being managed with acid suppressants, who underwent a Cytosponge-TFF3 test in primary care (Kadri et al. 2010). The sensitivity and specificity of the Cytosponge to detect Barrett's oesophagus with a cut-off segment length of 1 cm or more were 73.3% (95% confidence interval [CI] 44.9 to 92.2%) and 93.8% (95% CI 91.3 to 95.8%), respectively, compared to a reference standard of endoscopic biopsy. The positive predictive value (PPV) was 26.8% (95% CI 14.2 to 42.9%) and the negative predictive value (NPV) was 99.1% (95% CI 97.8 to 99.8%). However, when a cut-off segment length of 2 cm was used, the sensitivity of the Cytosponge test improved to 90.0% (95% CI 55.5 to 99.7%), whilst the specificity remained similar at 93.5% (95% CI 90.9 to 95.5%).

BEST3 was an RCT examining the use of Cytosponge testing for chronic reflux in primary care settings in England (Fitzgerald et al. 2020). More than 13,000 patients were in the trial, with 6,983 assigned to the intervention group (Cytosponge followed by endoscopy if the result was positive). However, only 1,654 patients successfully swallowed the capsule; it was also optional whether patients took the Cytosponge test and this may have introduced selection bias. The PPV for detecting Barrett's oesophagus, dysplasia or oesophago-gastric cancer in 221 participants with TFF3-positive Cytosponge results, who underwent subsequent endoscopy, was 59%.

BEST2 was a case-control study conducted in secondary settings across England, in which cases were individuals with a previous diagnosis of Barrett's oesophagus attending for their monitoring endoscopy, and controls were individuals referred for endoscopy because of dyspepsia and/or reflux symptoms (Ross-Innes et al. 2015). The sensitivity of Cytosponge-TFF3 testing to detect Barrett's oesophagus was 79.9% (95% CI 76.4 to 83.0%) in this study, whilst the

specificity was 92.4% (95% CI 89.5 to 94.7%). When patients who received a second Cytosponge test during the study period were included, the sensitivity of Cytosponge increased to 89.7% (95% CI 82.3 to 94.8%).

Results for all studies are shown in Table 1.

5.2.2 Case finding of Barrett's oesophagus

A cross-sectional study, conducted as part of a charity campaign in England, offered EndoSign tests to members of the public who had self-identified as having chronic heartburn and were deemed to be high-risk for oesophageal disease (Norton et al. 2025). The biomarkers TFF3, p53, and cellular atypia were tested for. Out of 60 tests performed, 54 provided conclusive results and 12 of these were positive (breakdown by which biomarkers were positive not reported). Eleven of the 12 participants with positive EndoSign results then accepted the offer for endoscopy in a private clinic. The PPV for the detection of Barrett's oesophagus was 72.7% (95% CI 43.5 to 91.7%).

A prospective cohort analysis from NHS England found that the sensitivity and specificity of capsule sponge testing with a result showing abnormal biomarkers to detect endoscopically-confirmed Barrett's oesophagus were 90.9% and 74.4%, respectively (Angel et al. 2025). The PPV and NPV were 34.1% and 98.2%, respectively. The biomarkers TFF3, p53 and cellular atypia were used in this study. When looking specifically at capsule sponge testing with a positive TFF3 result to detect endoscopically-confirmed Barrett's oesophagus, the sensitivity was 87.5%, specificity was 75.5%, PPV was 34.6% and NPV was 97.6%. The accuracy of capsule sponge testing with abnormal biomarker results to detect Barrett's oesophagus, oesophageal cancer, or atrophic cancer was similar to the results above, with a sensitivity of 90.2%, specificity of 76.8%, PPV of 43.5% and NPV of 97.6%. These results suggest good accuracy, with low levels of false negatives, but also high levels of false positives.

The results for both studies are shown in Table 1.

5.2.3 Barrett's oesophagus under surveillance

In a case-control study, the median sensitivity of Cytosponge with p53 testing to detect high-grade dysplasia or intramucosal cancer was 58% (Ross-Innes et al. 2017) compared with endoscopic biopsy. The specificity of p53 testing was 96%. The sensitivity and specificity of Cytosponge testing for glandular atypia to detect high-grade dysplasia or intramucosal cancer were 64% and 94%, respectively. The authors concluded that neither marker was sensitive or specific enough to be used individually.

Following this study, a cross-sectional study including a prospective cohort analysis examined the use of Cytosponge with cellular atypia and p53 testing for Barrett's oesophagus surveillance (Pilonis et al. 2022). A training cohort (n = 557) and a validation cohort (n = 334) were used to assess the accuracy of three diagnostic models compared to the reference standard of endoscopic biopsy in the cross-sectional study. The model using positive biomarkers from Cytosponge testing alone had a sensitivity of 74% (95% CI 65 to 83%) and 89% (95% CI 77 to 97%) for detecting high-grade dysplasia or intramucosal cancer in the training and validation cohorts, respectively. The specificities were 86% (95% CI 83 to 89%) and 84% (95% CI 80 to 88%), respectively, in the two cohorts. The overall accuracy of this model, as demonstrated by the area under the receiver operating characteristic curve (AUC), was 80% in the training cohort and 86% in the validation cohort. The AUC of positive Cytosponge markers to detect any grade of dysplasia were 77% and 80% in the cohorts, respectively. The addition of clinical risk factors to the model did not notably improve sensitivity (77% in the training cohort and 80% in the validation cohort).

In the real-world prospective cohort (n = 223), Cytosponge results indicated 39 participants had cellular atypia, p53 overexpression, or both. Within these 39 participants, the PPV of Cytosponge to detect high-grade dysplasia or intramucosal cancer was 31% and the PPV to detect any grade of dysplasia was 44%. These participants had had their Barrett's surveillance delayed by the Covid-19 pandemic; there may, therefore, have been a higher prevalence of dysplasia in this cohort due to delayed investigation.

In a prospective cohort study by Tan et al. (2025), patients were risk stratified based on clinical risk factors and the results of capsule sponge testing with the biomarkers p53 and atypia. The sensitivity of the test (comparing moderate and high-risk stratification to low risk) to detect any level of dysplasia or cancer was 87.2% (95% CI 77.9 to 93.1%), and to detect high-grade dysplasia or cancer was 94.4% (95% CI 80.0 to 99.0%). The NPVs for the low- and moderate-risk groups were consistently high for both ruling out any level of dysplasia or cancer, and high-grade dysplasia or cancer, with all values above 90%. The PPVs in the moderate- and high-risk groups were considerably lower for both any level of dysplasia or cancer and high-grade dysplasia or cancer. These were lower than 10% in the moderate-risk group and were 37.7% and 19.6% in the high-risk group for any level of dysplasia or cancer and high-grade dysplasia or cancer, respectively.

Results for all studies are shown in Table 1.

5.2.4 Barrett's oesophagus under surveillance after treatment

A prospective cohort study examined the accuracy of Cytosponge with TFF3 testing to detect residual Barrett's oesophagus in patients who had received ablative treatment for the condition (Eluri et al. 2022). When Barrett's oesophagus was defined as columnar epithelium of greater than or equal to 1 cm in the tubular oesophagus, the sensitivity of Cytosponge was 74% (95% CI 49 to 91%), the specificity was 85% (95% CI 78 to 91%), and the overall accuracy was 84% (95% CI 77 to 89%) compared to endoscopic biopsy (n = 142). The AUC was 0.74. The adjusted odds ratio (OR) of a positive Cytosponge result in Barrett's oesophagus cases compared to controls was 17.1 (95% CI 5.2 to 55.9). When using a definition of Barrett's oesophagus which included patients with endoscopic columnar epithelium of any length with concurrent biopsies showing intestinal metaplasia, the sensitivity, specificity and overall accuracy were 63%, 87%, and 82%, respectively. The AUC in this scenario was 0.75. A sensitivity analysis of all adequate Cytosponge samples (n = 175, including 33 that had only endoscopic results without biopsy) was performed and this resulted in sensitivity, specificity and overall accuracy of Cytosponge of 69%, 84%, and 81%, respectively. The AUC was 0.75. However, it should be noted that ablative treatment received within the past two months is a contraindication to the use of Cytosponge (NICE 2020).

Results for this study are shown in Table 1.

Table 1 – Capsule sponge compared to endoscopy with biopsy: diagnostic accuracy

Evidence source	Number of participants	Population	Diagnostic accuracy	Comments
Proactive screening of people with GORD				
Kadri et al. (2010) Prospective cohort study	n = 501 32 participants did not attend for gastroscopy and were classed as not having BO.	Median (range) age 62 (56 to 66) years 45.7% male 95.8% White, 4.2% other ethnicity GORD impact scores: 7.0% very well controlled, 19.8% fairly well controlled, uncontrolled 27.1%, poorly controlled 38.9%, very poorly controlled 7.2% Current use of acid suppressants: 13.4% antacids, 7.6% H_2 antagonists, 57.0% proton pump inhibitors, 1.8% H_2 and proton pump inhibitors, 20.2% none	<u>To detect BO with a cut-off segment length of 1 cm or more</u> Sensitivity 73.3% (95% CI 44.9 to 92.2%) Specificity 93.8% (95% CI 91.3 to 95.8%) PPV 26.8% (95% CI 14.2 to 42.9%) NPV 99.1% (95% CI 97.8 to 99.8%) <u>To detect BO with a cut-off segment length of 2 cm or more</u> Sensitivity 90.0% (95% CI 55.5 to 99.7%) Specificity 93.5% (95% CI 90.9 to 95.5%)	• Participants who successfully swallowed the Cytosponge device were invited to attend for a gastroscopy within three weeks of the Cytosponge procedure. BO was defined as endoscopically visible columnar lined epithelium arising at least 1 cm circumferentially above the gastro-oesophageal junction with IM. If BO was present, four biopsies every 2 cm were collected according to surveillance guidelines. • Endoscopists and histopathologists were blinded to the result of the Cytosponge test.
Ross-Innes et al. (2015) Case-control study	n = 1,110 (647 cases, 463 controls)	<u>Cases</u> Median (IQR) age 66 (58 to 73) years Male:female ratio 4:1 96.8% white, 1.8% other ethnicity Median (IQR) BMI 28.1 (25.6 to 31.2) Maximum length of BO (median [IQR]) 5 (3 to 8) cm <u>Controls</u> Median (IQR) age 56 (44 to 66) years Male:female ratio 1.0:1.3 92.5% white, 7.3% other ethnicity Median (IQR) BMI 26.8 (24.0 to 30.2) Maximum length of BO NA	<u>To detect BO</u> Sensitivity 79.9% (95% CI 76.4 to 83.0%) Specificity 92.4% (95% CI 89.5 to 94.7%) <u>To detect BO, including participants who had a second surveillance test</u> Sensitivity 89.7% (95% CI 82.3 to 94.8%)	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Cases were individuals with a previous diagnosis of BO attending for their monitoring endoscopy. Controls were individuals referred to endoscopy because of dyspepsia and/or reflux symptoms. • Endoscopy was performed within one hour of Cytosponge testing. • BO was defined as endoscopically visible columnar-lined oesophagus that measured at least 1 cm circumferentially or at least 3 cm in non-circumferential tongues with documented histopathological evidence of IM on at least one biopsy in the course of their endoscopic history. • Participants under surveillance for BO who happened to undergo a second surveillance endoscopy for clinical purposes during the

Evidence source	Number of participants	Population	Diagnostic accuracy	Comments
				study period were invited to take a Cytosponge test again. • Those scoring Cytosponge samples were blinded to the patient's diagnosis and histocytopathologists reviewing biopsy results were blinded to Cytosponge results.
Fitzgerald et al. (2020) RCT	<u>Intervention group</u> n = 6,834. 1,654 successfully swallowed the capsule sponge device, 221 with positive TFF3 result underwent endoscopy	<u>Intervention group</u> Age distribution 50 to 59 years 20%, 60 to 69 years 34%, 70 to 79 years 37%, 80 to 89 years 8%, 90 to 99 years 1% 48% male Median (IQR) Index of Multiple Deprivation decile NR	PPV for BO, dysplasia or oesophago-gastric cancer: 59%	• Cytosponge test was optional in intervention group, ITT analysis used. • Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Primary care setting. • 24% of participants randomised to intervention group successfully swallowed Cytosponge device. • 150 participants (9%) in intervention arm had low-confidence result after repeat Cytosponge testing.
Case finding of Barrett's oesophagus				
Norton et al. (2025) Cross-sectional study	n = 60 (12 positive Endosign tests, 11 accepted endoscopy offer)	(Of 78 participants invited to undergo EndoSign test): Mean age 57.1 ± 9.4 years 85.9% male Demographics of the 60 participants who successfully swallowed the capsule, and were included in analysis, NR	54 of 60 tests conclusive Among those with any cellular abnormality detected on EndoSign, the PPV for the detection of BO was 72.7% (95% CI 43.5 to 91.7%)	• Included self-referred individuals who had chronic heartburn who were deemed to be high-risk. • The study was part of a charity campaign that was supported by Cyted. • Unclear whether/how many participants who underwent endoscopy also had biopsies taken. • Not part of NHS pathways. EndoSign testing carried out in mobile units, those with positive results were sent to a private clinic for confirmatory gastroscopy. Anyone with clinically actionable findings was referred to their GP for ongoing care.

Evidence source	Number of participants	Population	Diagnostic accuracy	Comments
Angel et al. (2025) Prospective cohort analysis	n = 871 (808 successfully swallowed capsule sponge device, 763 adequate samples) 331 patients underwent endoscopy	Median (IQR) age 54 (41.0 to 65.5) years 40.1% male Patients with adequate samples: Median (IQR) age 54 (41.0 to 64.0) years for males, 56 (42.6 to 65.7) years for females	<u>Abnormal biomarker to detect endoscopically-confirmed BO</u> Sensitivity 90.9% Specificity 74.4% PPV 34.1% NPV 98.2% <u>Positive for TFF3 to detect endoscopically-confirmed BO</u> Sensitivity 87.5% Specificity 75.5% PPV 34.6% NPV 97.6% <u>Abnormal biomarker to detect endoscopically-confirmed BO, oesophageal cancer, or atrophic gastritis</u> Sensitivity 90.2% Specificity 76.8% PPV 43.5% NPV 97.6%	• Study started during the COVID-19 pandemic when usual endoscopy services were disrupted. • All patients were recruited to the DELTA or NHS England evaluations reported elsewhere. • For those who had a negative capsule sponge test and were not offered endoscopy, a review of the Medilogik EMS database was undertaken at 1, 2 and 3 years from the test to see if they had been referred back to endoscopy and to review subsequent endoscopy findings. • From November 2020 to June 2023, the Cytosponge device was used and from July 2023 onwards, the EndoSign device was used. • Only patients with abnormal, inadequate or failed capsule sponge tests or ongoing/concerning symptoms had endoscopy.
Barrett's oesophagus under surveillance				
Ross-Innes et al. (2017) Case-control study	Discovery cohort (n = 468)	Non-dysplastic BO (n = 376): Median (IQR) age 64 (56 to 71) years Male:female ratio 3.8:1 97% white, 2% other ethnicity, less than 1% refused to disclose Median (IQR) BMI 28.1 (25.5 to 30.8) BO with HGD or IMC (n = 92): Median (IQR) age 69 (63 to 74) years Male:female ratio 7.4:1 99% white, 1% other ethnicity Median (IQR) BMI 28.8 (26.1 to 31.1) Inclusion criteria: all	<u>To detect HGD or IMC</u> p53: median (IQR) sensitivity 58% (44 to 70%), median (IQR) specificity 96% (92 to 98%) Glandular atypia: median (IQR) sensitivity 64% (50 to 77%), median (IQR) specificity 94% (90 to 97%)	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Endoscopy was performed within one hour of Cytosponge testing. • Biopsy samples were taken from any visible lesions and from each quadrant, every 2 cm. • Pathologists reviewing biopsy results were blinded to Cytosponge results.

Evidence source	Number of participants	Population	Diagnostic accuracy	Comments
		BO patients with IM and a TFF3-positive Cytosponge test. No minimum BO segment length was required provided participants had a least one TFF3-positive cell on Cytosponge.		
Pilonis et al. (2022) Cross-sectional study and prospective cohort analysis	Cross-sectional study (n = 891) Prospective cohort analysis (n = 223)	Cross-sectional study: Training cohort n = 557 Median (IQR) age 65 (59 to 72) years 81% male 98% white, 2% other ethnicity Median (IQR) BO maximum segment length 5 (3 to 8) cm Median (IQR) BO circumferential length 3 (1 to 6) cm Median (IQR) BMI 28.25 (25.61 to 31.07) Validation cohort n = 334 Median (IQR) age 67 (58 to 73) years 75% male Ethnicity NR Median (IQR) BO maximum segment length 3 (2 to 6) cm Median (IQR) BO circumferential length 1 (0 to 4) cm Median (IQR) BMI 27.90 (25.20 to 30.81) Prospective cohort analysis: Median age 69 (IQR 60 to 74) years 74% male Ethnicity NR Median (IQR) BO maximum segment length 3 (2 to 6) cm Median (IQR) BO circumferential length 1 (0 to 4) cm Median (IQR) BMI 26.90 (24.12 to 29.30)	Cross-sectional study: <u>Cytosponge alone to detect HGD or intramucosal cancer</u> Training cohort: sensitivity 74% (95% CI 65 to 83%), specificity 86% (95% CI 83 to 89%), AUC 80% (95% CI 75 to 85%) Validation cohort: sensitivity 89% (95% CI 77 to 97%), specificity 84% (95% CI 80 to 88%), AUC 86% (95% CI 81 to 92%) <u>Cytosponge alone to detect any grade of dysplasia</u> Training cohort: AUC 77% (95% CI 73 to 81%) Validation cohort: AUC 80% (95% CI 74 to 86%) <u>Cytosponge combined with clinical risk factors to detect HGD or intramucosal cancer</u> Training cohort: sensitivity 77% (95% CI 68 to 86%) Validation cohort: sensitivity 80% (95% CI 66 to 91%) Prospective cohort analysis: PPV for HGD or intramucosal cancer 31%, PPV for any grade of dysplasia 44%	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Endoscopies were performed on the same day as Cytosponge (BEST2) or within 2 months of Cytosponge (BEST3). • Participants recruited in the prospective cohort analysis had their regular Barrett's surveillance delayed by Covid-19. • Clinical risk factors used in the Cytosponge and risk factors model were age, sex and BO segment length.

Evidence source	Number of participants	Population	Diagnostic accuracy	Comments
Tan et al. (2025) Prospective cohort study	n = 910	n = 910 Consecutive patients undergoing BO surveillance from 13 hospitals in the UK who participated in the DELTA study and the NHS England implementation pilot study. Median (IQR) age 68 (60 to 74) years 76% male Histology at baseline: • Non-dysplastic BO 90% • Indefinite for dysplasia 1% • Crypt dysplasia < 1% • LGD 5% • HGD or intramucosal carcinoma 3% • Adenocarcinoma (≥ T2) 1%	<u>To detect any level of dysplasia or cancer</u> • Sensitivity 87.2% (95% CI 77.9 to 93.1%) (high and moderate risk vs low risk) • PPV 37.7% (95% CI 29.7 to 46.4%) (high-risk group) • PPV 8.1% (95% CI 5.3 to 12.1%) (moderate-risk group) • NPV 97.8% (95% CI 95.9 to 98.8%) (low-risk group) • NPV 91.9% (95% CI 87.9 to 94.7%) (moderate-risk group) <u>To detect HGD or cancer</u> • Sensitivity 94.4% (95% CI 80.0 to 99.0%) (high and moderate risk vs low risk) • PPV 19.6% (95% CI 13.5 to 27.4%) (high-risk group) • PPV 2.5% (95% CI 1.1 to 5.2%) (moderate-risk group) • NPV 99.6% (95% CI 98.4 to 99.9%) (low-risk group) • NPV 97.5% (95% CI 94.8 to 98.9%) (moderate-risk group)	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • The DELTA study and the NHS England implementation pilot study followed the same protocol. • Patients were assigned to low- or moderate-risk groups at baseline based on clinical risk factors and previous BO findings. Patients were escalated to the high-risk group after capsule sponge testing if their results showed any of atypia, atypia of uncertain significance, equivocal p53, or aberrant p53 expression. • Study took place during Covid-19 pandemic when endoscopy services were disrupted. • Some patients had more than one endoscopy follow-up, for example for indefinite for dysplasia or first diagnosis of LGD, which followed the clinical standard of a repeat at 6 months. • These results are, in some respects, prognostic rather than diagnostic, due to the time delay between index and reference standard tests.
Barrett's oesophagus under surveillance after ablative treatment				
Eluri et al. (2022) Prospective cohort study	n = 175 (175 of 234 patients had adequate Cytosponge samples), 142 had endoscopic and histologic data available and were	Mean age 71 ± 9 years 83% male 65% History of endoscopic mucosal resection Median (IQR) time since first ablation 20 (2 to 113) months Median time since last ablation 10 (1 to 111) months	Detection of residual BO after ablative treatment: Sensitivity 74% (95% CI 49 to 91%) Specificity 85% (95% CI 78 to 91%) Overall accuracy 84% (95% CI 77 to 89%) AUC 0.74 When using a definition of BO which included patients with endoscopic	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • All patients had received prior ablative treatment for BO • All patients underwent upper endoscopy approximately 2 hours after Cytosponge administration. • Biopsies were obtained from BO segments in those with residual BO undergoing further

Evidence source	Number of participants	Population	Diagnostic accuracy	Comments
	included in primary analysis		columnar epithelium of any length with concurrent biopsies showing IM: Sensitivity 63% Specificity 87% Overall accuracy 82% AUC 0.75 Sensitivity analysis of all adequate Cytosponge samples (n = 175, 36 BO cases, 139 controls) Sensitivity 69% Specificity 84% Overall accuracy 81% AUC 0.75 Adjusted odds of a positive Cytosponge in BO cases vs. controls OR 17.1 (95% CI 5.2 to 55.9)	endoscopic treatment, and from the cardia, gastroesophageal junction, and neosquamous oesophagus in post-complete eradication of IM patients. A subset of patients (n = 33) undergoing ablation, but had not achieved complete eradication, only had endoscopic evidence of columnar epithelium documented, without concurrent biopsies, due to the endoscopist's concern of biopsies interfering with ablation. • Presence of BO was defined as columnar epithelium of greater than or equal to 1 cm in the tubular oesophagus, with concurrent IM on biopsies or endoscopic mucosal resection specimens of that area.

Abbreviations: AUC: area under the receiver operating characteristic curve; BEST: Barrett's oEsophagus Screening Trial; BO: Barrett's oesophagus; CI: confidence interval; CS: capsule sponge; DELTA: integrateD diagnostic solution for EarLy deTection of oesophageal cAncer; GORD: gastro-oesophageal reflux disease; HGD: high-grade dysplasia; IM: intestinal metaplasia; IMC: intramucosal adenocarcinoma; IQR: interquartile range; ITT: intention-to-treat; LGD: low-grade dysplasia; NPV: negative predictive value; NR: not reported; OAC: oesophageal adenocarcinoma; OR: odds ratio; PPV: positive predictive value; RCT: randomised controlled trial; RR: risk ratio; SCC: squamous cell carcinoma; SR: systematic review; TFF3: Trefoil factor 3; UGI: upper gastrointestinal

5.3 Detection rates

5.3.1 Proactive screening of people with GORD taking medication

In the BEST3 RCT, there were 221 patients in the intervention group that tested TFF3 positive and had a follow-up endoscopy (Fitzgerald et al. 2020). Of these 221 positive capsule sponge tests, endoscopic biopsy confirmed Barrett's oesophagus in 127, oesophago-gastric cancer in four, and no Barrett's oesophagus in 90, of which 33 had intestinal metaplasia. Of the 127 with Barrett's oesophagus, 116 had no dysplasia, seven had indefinite dysplasia, one had low-grade dysplasia, and three had high-grade dysplasia. All four of the oesophago-gastric cancer cases were stage I.

5.3.2 Case finding of Barrett's oesophagus

In the small sample of participants with positive EndoSign results that underwent endoscopy (n = 11) in the cross-sectional study, eight were diagnosed with Barrett's oesophagus (Norton et al. 2025). One of these Barrett's oesophagus cases also showed evidence of low-grade dysplasia. No cases of high-grade dysplasia or cancer were identified.

In a real-world, prospective cohort analysis of patients investigated for chronic reflux symptoms in Scottish hospitals, 1,243 out of 1,385 capsule sponge tests (90%) on 1,305 patients provided sufficient results (Chien et al. 2024a). Of those with a negative TFF3 result who had a subsequent biopsy (n = 112), 102 were found to have no intestinal metaplasia (91.1%). Six patients did have intestinal metaplasia (5.4%), whilst two had gastric adenocarcinoma (1.8%) and there was one case (0.9%) each of oesophageal adenocarcinoma and gastric lymphoma. Patients whose upper gastrointestinal tract appeared macroscopically normal during endoscopy did not have biopsies taken. There were a further 78 patients who were TFF3 negative but did not have a biopsy taken. Therefore, the vast majority of TFF3-negative patients did not have Barrett's oesophagus, however, there were several cases of severe pathology in this patient group. Amongst those with a positive TFF3 result (with or without atypia or p53 positivity), biopsies in 97 participants identified no intestinal metaplasia in 52 (53.6%), intestinal metaplasia in 43 (44.3%), one person with low-grade dysplasia (1.0%), and one person with neuroendocrine carcinoma (1.0%). Fourteen TFF3-positive participants did not have a biopsy. For those positive for atypia or p53, 19 had biopsies. Of these, 14 had no intestinal metaplasia (73.7%), four had intestinal metaplasia (21.1%), and one had neuroendocrine carcinoma (5.3%). These results suggest quite a high number of false positives with the three biomarkers, but also a very low number of false negatives. However, as the majority of those with negative tests did not have endoscopy, it is uncertain how many may have had missed pathology.

Data from 277 patients in English hospitals tested for reflux symptoms with Cytosponge and cellular atypia, p53 and TFF3 testing, and having endoscopic biopsy results were reported in a prospective cohort study by Gourgiotis et al. (2025). Of the 111 patients with a positive capsule sponge result (TFF3, p53, or atypia) who underwent endoscopy, 22 had Barrett's oesophagus and seven had intestinal metaplasia. Twenty-four had other findings, with the remaining 58 having no findings on endoscopy. Other findings include oesophagitis and hiatus hernia, which can lead to atypia findings on capsule sponge tests. From those with equivocal capsule sponge test results, 87 underwent endoscopy and three were found to have Barrett's oesophagus and one had intestinal metaplasia, 21 had other findings and 62 had no findings. Endoscopy was carried out on 79 patients with negative capsule sponge results due to ongoing symptoms. Within this results group none had Barrett's oesophagus or intestinal metaplasia, 25 had other findings and 54 had no findings. These results suggest a low rate of false negative capsule sponge tests, but false positive rates could be quite high. This study also compared the Barrett's oesophagus rates between the capsule sponge group and a counterfactual group where all received routine

endoscopy. They found similar Barrett's oesophagus rates between the groups (capsule sponge group 1.8%, counterfactual group 1.4%); however, the diagnostic yield from the capsule sponge group was higher as this rate was identified from a much smaller group of participants that required endoscopy (307 compared with 1,441).

Another real-world study from England (Angel et al. 2025) found that a small number of patients with normal capsule sponge biomarker results had biopsy findings. Of 163 patients, one (0.6%) had Barrett's oesophagus without intestinal metaplasia, whilst two (1.2%) had Barrett's oesophagus with intestinal metaplasia. Eleven of the 163 patients (6.7%) were found to have reflux oesophagitis and three (1.8%) had gastric atrophy or cancer. The majority of patients with negative biomarkers (75.9%) did not undergo endoscopy. For patients with abnormal biomarker results, 85 underwent endoscopic biopsy. The biopsy results indicated one patient (1.2%) had high-grade dysplasia and one (1.2%) had oesophageal adenocarcinoma. Six patients (7.1%) had gastric diagnoses and 30 (35.3%) had Barrett's oesophagus without dysplasia or indefinite for dysplasia. Of those who had inadequate capsule sponge tests after two attempts, 83 agreed to routine endoscopy. The biopsy findings of these patients showed no cases of high-grade dysplasia or cancer and only six cases (7.2%) of Barrett's oesophagus were identified. Of the six patients who tested positive for p53, atypia, or both, one had oesophageal adenocarcinoma, two had Barrett's oesophagus with intestinal metaplasia, and there was one case each of focal intestinal metaplasia at the gastro-oesophageal junction, reflux oesophagitis, and atrophic gastritis without intestinal metaplasia. These findings, again, suggests high numbers of false positives but also low numbers of false negatives. However, some serious pathology was detected in patients with normal biomarkers on capsule sponge testing.

This study also followed up patients with a negative test that did not undergo endoscopy for a median of 27.24 months (range 12 to 48 months) to see if they returned for endoscopic investigation. It was found that 76% of patients did not return for an endoscopy within this time, and if they did, it usually took place within the first year since their capsule sponge test.

Results of all studies are shown in Table 3.

5.3.3 Barrett's oesophagus under surveillance

Pilonis et al. (2022) concluded that the results of Cytosponge with cellular atypia and p53 testing, and assessing clinical risk factors can be used to risk stratify people under surveillance for Barrett's oesophagus. Those with p53 overexpression or cellular atypia, or both, were considered the high-risk group, whilst those with clinical risk factors (such as increasing Barrett's oesophagus segment length, increasing age, or male sex) were considered moderate risk, and the low-risk group had neither positive biomarkers nor clinical risk factors. In the training and validation cohorts of the cross-sectional study, 52% (68 of 132) and 41% (31 of 75) of participants classified as high risk were found to have high-grade dysplasia or intramucosal cancer, respectively. Of those assigned to the moderate-risk group, 79% (19 of 24) and 50% (2 of 4), respectively, had high-grade dysplasia or intramucosal cancer. Whereas within the low-risk groups, only 2% (5 of 210) in the training cohort and 1% (2 of 185) in the validation cohort were found to have high-grade dysplasia or cancer at endoscopy. Within both these cohorts, the diagnostic yield to detect high-grade dysplasia or cancer was higher in those receiving endoscopy after a positive capsule sponge result, at 47% (97 of 206 endoscopies), than in those that received endoscopic surveillance alone (14%, 125 of 891).

In the prospective cohort analysis part of the above study, 17% (39 of 223) of patients were positive for p53, cellular atypia, or both and classified as high risk. 64% of patients with both aberrant p53 expression and cellular atypia (7 of 11) were found to have high-grade dysplasia or cancer. Using clinical risk factors, 17% (39 patients) were classified as moderate risk in this

cohort and 65% (145 of 223) were classified as low risk. The authors concluded that using the biomarkers p53 and cellular atypia, and assessing clinical risk factors, can be used to risk stratify patients under surveillance for Barrett's oesophagus and inform clinical decision-making on the need for endoscopy and the urgency. This could help relieve pressure on endoscopy services whilst ensuring those in most clinical need are prioritised for investigation.

In a retrospective cohort analysis of people under surveillance for Barrett's oesophagus in Scottish hospitals, 608 patients underwent endoscopy within 12 months of a Cytosponge test and were included in analyses (Chien et al. 2024b). Of the 608 Cytosponge tests, 20% did not provide sufficient results. Of 136 patients that tested TFF3-negative and underwent endoscopic biopsy, the majority had non-dysplastic Barrett's oesophagus (80 patients, 58.8%), 48 had no intestinal metaplasia (35.3%), five were indefinite for dysplasia (3.7%), and three had low-grade dysplasia (2.2%). Forty-eight patients tested positive for TFF3 and underwent endoscopy; the majority also had non-dysplastic Barrett's oesophagus (37 patients, 77.1%), eight patients had no intestinal metaplasia (16.7%), one was indefinite for dysplasia (2.1%), and two had low-grade dysplasia (4.2%). Those with non-dysplastic Barrett's oesophagus also made up the majority of patients that tested positive for cellular atypia only (121 of 179 patients, 67.6%). Eleven patients in this results group (6.1%) had no intestinal metaplasia whilst 15 (8.4%) were indefinite for dysplasia, 20 (11.2%) had low-grade dysplasia, and six (3.4%) had high-grade dysplasia. Six patients that were positive for cellular atypia were found to have cancer; three (1.7%) had adenocarcinoma, two (1.1%) had intramucosal carcinoma, and one (0.6%) had squamous cell carcinoma. Only one case of cancer, an intramucosal carcinoma, was found amongst 24 patients that tested positive for p53 only (4.2%). The majority (17 of 24, 70.8%) were found to have non-dysplastic Barrett's oesophagus, with one patient (4.2%) indefinite for dysplasia, and five having low-grade dysplasia (20.8%). A higher proportion of cancer (16.5%) was found in patients who tested positive for both cellular atypia and p53; out of 97 patients, nine had intramucosal carcinoma and seven had adenocarcinoma. Of the remaining patients in this group, 35 (36.1%) had non-dysplastic Barrett's oesophagus, 11 (11.3%) were indefinite for dysplasia, 18 (18.6%) had low-grade dysplasia, and 17 (17.5%) had high-grade dysplasia. These results suggest cellular atypia and p53 biomarkers provide a notable number of false positive results for dysplasia or cancer, but also very low numbers of false negatives. Cancers were only found in those that were positive for these two markers, however, the majority of people negative for these markers did not undergo endoscopy, so the true number of false negatives is not known.

Similar to Pilonis et al. (2022), hospitals in NHS Scotland used capsule sponge testing with TFF3, p53 and cellular atypia to risk stratify patients under surveillance for Barrett's oesophagus and to help determine clinical management (Chien et al. 2024b). Of 4,204 Barrett's oesophagus cases under surveillance during the study period, 7.8% were classed as high risk, 20.3% as moderate risk, 19.4% as low risk, and 52.5% as ultra-low risk. The criteria for each risk group are shown in Table 2. Only small proportions of patients in the ultra-low, low and moderate risk groups were referred for endoscopy, compared with 98.5% of patients in the high-risk group.

Table 2 – Criteria for risk stratification reported by (Chien et al. 2024b)

Risk group	Criteria
Ultra-low risk	TFF3-negative and/or no previous IM TFF3-positive or known IM, and M < 3 and C < 2
Low risk	TFF3-positive or known IM M3 to 5 or C2 to 3
Moderate risk	TFF3-positive and: • M > 10 or C > 6 • M > 5 or C > 3 and male • M > 5 or C > 3 and age over 60 years
High risk	TFF3-positive, p53 and/or atypia positive

Abbreviations: C, circumferential length of Barrett's oesophagus (cm); IM, intestinal metaplasia; M, maximal length of Barrett's oesophagus (cm); TFF3, trefoil factor 3

In another recent prospective cohort study, patients were risk stratified into low- and moderate-risk groups based on age, sex, and length of Barrett's oesophagus identified at their last surveillance endoscopy (Tan et al. 2025). The moderate risk criteria were the same as those shown in Table 2, except for TFF3-positivity, and those that did not meet these criteria were classed as low risk. Patients were then reassigned to the high-risk group if they were found to have atypia, atypia of uncertain significance, equivocal p53, or aberrant p53 expression on capsule sponge testing. This risk stratification then determined the next course of action: high risk, urgent endoscopy pathway; low or moderate risk, triage for endoscopy at a later date. Endoscopy results from the low-risk group (n = 489) showed 478 (97.8%) had non-dysplastic Barrett's oesophagus, one (0.2%) had crypt dysplasia, eight (1.6%) had low-grade dysplasia and two (0.4%) had high-grade dysplasia or intramucosal carcinoma. For the moderate-risk group (n = 283), 255 (90.1%) were found to have non-dysplastic Barrett's oesophagus, five (1.8%) were indefinite for dysplasia, 15 (5.3%) had low-grade dysplasia, and seven (2.5%) had high-grade dysplasia or intramucosal cancer. Within the high-risk group (n = 138), 84 (60.9%) had non-dysplastic Barrett's oesophagus on endoscopy, two (1.4%) were indefinite for dysplasia, one (0.7%) had crypt dysplasia, 24 (17.4%) had low-grade dysplasia, 22 (15.9%) had high-grade dysplasia or intramucosal carcinoma, and five (3.6%) had adenocarcinoma. Prevalence estimates showed much higher values for both any level of dysplasia, and high-grade dysplasia and cancer, in high-risk groups than low- or moderate-risk groups. The high-risk group was split into tier 1 (positive for both atypia and p53, representing the highest level of risk) and tier 2 (all other positive capsule sponge findings), and the prevalence of dyplasia was 85.2% and 26.1%, respectively. This compared with just 2.2% in the low-risk group. For high-grade dysplasia or cancer, the prevalence was 55.6% and 10.8% in high-risk tiers 1 and 2, respectively, compared with 0.4% and 2.5% for low and moderate risk. These results suggest risk stratifying can effectively assigned patients to an appropriate group for triage of further investigation. However, there may still be significant pathology in those assigned to low- or moderate-risk groups, and many patients sent for urgent endoscopy may not be found to have dysplasia.

In another prospective cohort study, the authors compared the dysplasia rates before and after capsule sponge testing was introduced to a health board in Scotland (Chien & Glen 2025). They found there was no difference in the detection rates for indefinite for dysplasia, high-grade dysplasia, intramucosal cancer or invasive cancer, but there was statistically significantly less detection of low-grade dysplasia (3.4% pre-implementation compared with 2.2% post-implementation, p = 0.033). The study also compared the dysplasia rates after implementation of capsule sponge testing, comparing those that received capsule sponge testing to those that were investigated with endoscopy only. There was a statistically significantly lower detection rate for indefinite for dysplasia within the capsule sponge cohort than the endoscopy only cohort, but

no differences in the detection of low-grade dysplasia, high-grade dysplasia or cancer. Statistically significantly lower detection rates for indefinite for dysplasia and low-grade dysplasia were found when those who received capsule sponge testing were compared to all those who received endoscopy only (the pre-intervention group and the endoscopy only cohort of the implementation group combined). This shows that capsule sponge testing does not negatively impact the detection of high-grade dysplasia or cancer, though it may be worse for detecting indefinite for dysplasia or low-grade dysplasia. However, diagnosing these two conditions is often difficult and can be variable at endoscopy and red flag symptoms were excluded in the capsule sponge cohort, but not endoscopy only, and this may have led to more endoscopy findings in the latter cohort. It is also notable that 763 fewer endoscopies were carried out in the 2-year period after capsule sponge testing was introduced, and only 17.1% of the capsule sponge cohort underwent endoscopy, without a reduction in the number of high-grade dysplasia or cancer.

Results from all studies are shown in Table 3.

Table 3 – Capsule sponge: detection rates

Evidence source	Number of participants	Population	Detection rates	Comments
Proactive screening of people with GORD				
Fitzgerald et al. (2020) RCT	<u>Intervention group</u> n = 6,834. 1,654 successfully swallowed the capsule sponge device, 221 with positive TFF3 result underwent endoscopy	<u>Intervention group</u> Age distribution 50 to 59 years 20%, 60 to 69 years 34%, 70 to 79 years 37%, 80 to 89 years 8%, 90 to 99 years 1% 48% male Median (IQR) Index of Multiple Deprivation decile NR	<u>Intervention group</u> BO: 140 (had Cytosponge test 127, no Cytosponge test 13) • No dysplasia 129 (had Cytosponge test 116, no Cytosponge test 13) • Indefinite dysplasia 7 (had Cytosponge test 7, no Cytosponge test 0) • LGD 1 (had Cytosponge test 1, no Cytosponge test 0) • HGD 3 (had Cytosponge test 3, no Cytosponge test 0) Oesophago-gastric cancer: 7 (had Cytosponge test 4, no Cytosponge test 3) • Stage I 5 (had Cytosponge test 4, no Cytosponge test 1) • Stage IV 2 (had Cytosponge test 0, no Cytosponge test 2) No BO: 90, 33 had IM (all from those who underwent Cytosponge testing)	• Cytosponge test was optional in intervention group, ITT analysis used. • Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Primary care setting. • 24% of participants randomised to intervention group successfully swallowed Cytosponge device. • 150 participants (9%) in intervention arm had low-confidence result after repeat Cytosponge testing.
Case finding of Barrett's oesophagus				
Norton et al. (2025) Cross-sectional study	n = 60 (12 positive Endosign tests, 11 accepted endoscopy offer)	(Of 78 participants invited to undergo EndoSign test): Mean age 57.1 ± 9.4 years 85.9% male Demographics of the 60 participants who successfully swallowed the capsule, and were included in analysis, NR	<u>Endoscopy results of those who had positive EndoSign results (n = 11)</u> BO 8, of which 1 participant with LGD	• Included self-referred individuals who had chronic heartburn who were deemed to be high-risk. • The study was part of a charity campaign that was supported by Cyted. • Unclear whether/how many participants who underwent endoscopy also had biopsies taken.

Evidence source	Number of participants	Population	Detection rates	Comments
				• Proportion of participants positive for each biomarker NR. • Not part of NHS pathways. EndoSign testing carried out in mobile units, those with positive results were sent to a private clinic for confirmatory gastroscopy. Anyone with clinically actionable findings was referred to their GP for ongoing care.
Chien et al. (2024a) Prospective cohort analysis	n = 1,305 patients, 1,385 Cytosponge tests	Median (IQR) age 56 (46 to 65) years 42.4% male Median BMI 28.1 (25 to 32.4) Positive smoking history 37.5% Proton pump inhibitor use 88.2%	142 of 1,385 tests insufficient <u>Biopsy results</u> TFF3 negative (n = 190): no biopsy 78, no IM 102, IM 6, OAC 1, gastric adenocarcinoma 2, gastric lymphoma 1 At least one positive biomarker (n = 124): no biopsy 15, no IM 63, IM 44, LGD 1, neuroendocrine carcinoma 1 TFF3 positive (with or without atypia or p53 positive) (n = 111): no biopsy 14, no IM 52, IM 43, LGD 1, neuroendocrine carcinoma 1 Atypia or p53 positive (with or without TFF3 positive) (n = 20): no biopsy 1, no IM 14, IM 4, neuroendocrine carcinoma 1	• Pilot was conducted during the COVID-19 pandemic when usual endoscopy services were disrupted. • 80 tests were repeat tests performed due to insufficient first samples or assessment of inflammation healing. • If UGI tract appeared macroscopically normal during endoscopy, no biopsy was taken.
Gourgiotis et al. (2025) Prospective cohort study	<u>CS group</u> n = 2,875 (1,549 with sufficient data for detailed analysis) <u>Counterfactual group</u> n = 1,181	<u>CS group</u> Median (IQR) age at referral 52 (40 to 62) years 42.3% male Median time between referral and index date 27 (13 to 70) days 80.4% White, 19.6% non-White Heartburn 14.8% Waterbrash 0.9%	<u>Biopsy results in CS group</u> Positive CS with endoscopy (n = 111): BO 22, IM 7, other 24, no findings 58 Equivocal TFF3 or p53 with endoscopy (n = 87): BO 3, IM 1, other 21, no findings 62 Negative CS with endoscopy (n = 79): BO 0, IM 0, other 25, no findings 54	• Developer of Cytosponge and co-founder of Cyted involved in study. • Patients that were ineligible for Cytosponge or declined were excluded.

Evidence source	Number of participants	Population	Detection rates	Comments
		Reflux 74.2% Use of acid suppressants within last 6 months 84.1% <u>Counterfactual group</u> Demographics not reported but stated to be similar to the intervention group	1,189 of 1,411 provided unequivocal result (94 data missing) <u>BO rates</u> CS group 25 out of 1,411 (1.8%) Counterfactual 17 out of 1,181 (1.4%) Higher diagnostic yield from CS as only 307 participants required endoscopy	
Angel et al. (2025) Prospective cohort analysis	n = 871 (808 successfully swallowed capsule sponge device, 763 adequate samples) 331 patients underwent endoscopy	Median (IQR) age 54 (41.0 to 65.5) years 40.1% male Patients with adequate samples: Median (IQR) age 54 (41.0 to 64.0) years for males, 56 (42.6 to 65.7) years for females	<u>All endoscopic biopsy results (n = 331)</u> BO no IM 7, BO with IM 30, Focal IM at gastro-oesophageal junction 5, BO with IM and indefinite dysplasia 2, BO with IM and HGD 1, Oesophageal adenocarcinoma with BO and IM 1, reflux oesophagitis 20, gastric atrophy/cancer 2 <u>Abnormal biomarker biopsy results (n = 85)</u> BO no IM 4, BO with IM 25, Focal IM at gastro-oesophageal junction 5, BO with IM and indefinite dysplasia 1, BO with IM and HGD 1, Oesophageal adenocarcinoma with BO and IM 1, reflux oesophagitis 4, gastric atrophy/cancer 6 <u>Normal biomarker biopsy results (n = 163)</u> BO no IM 1, BO with IM 2, Focal IM at gastro-oesophageal junction 0, BO with IM and indefinite dysplasia 0, BO with IM and HGD 0, Oesophageal adenocarcinoma with BO and IM 0, reflux oesophagitis 11, gastric atrophy/cancer 3 <u>Inadequate capsule sponge test biopsy results (n = 83)</u> BO no IM 2, BO with IM 3, Focal IM at gastro-oesophageal junction 0, BO with IM and indefinite dysplasia 1, BO with IM and HGD 0, Oesophageal adenocarcinoma with BO and IM 0, reflux oesophagitis 5, gastric atrophy/cancer 1	• Study started during the COVID-19 pandemic when usual endoscopy services were disrupted. • All patients were recruited to the DELTA or NHS England evaluations reported elsewhere. • For those who had a negative capsule sponge test and were not offered endoscopy, a review of the Medilogik EMS database was undertaken at 1, 2 and 3 years from the test to see if they had been referred back to endoscopy and to review subsequent endoscopy findings. • From November 2020 to June 2023, the Cytosponge device was used and from July 2023 onwards, the EndoSign device was used. • Only patients with abnormal, inadequate or failed capsule sponge tests, or ongoing/concerning symptoms had endoscopy.

Evidence source	Number of participants	Population	Detection rates	Comments
			<u>High-risk biomarkers (p53 and/or atypia positive) biopsy results (n = 6)</u> BO no IM 0, BO with IM 2, Focal IM at gastro-oesophageal junction 1, BO with IM and indefinite dysplasia 0, BO with IM and HGD 0, Oesophageal adenocarcinoma 1, reflux oesophagitis 1, gastric atrophy/cancer 1	
Barrett's oesophagus under surveillance				
Chien et al. (2024b) Retrospective cohort analysis	n = 3,745, 4,204 Cytosponge tests. n = 608 underwent UGI endoscopy within 12 months and were included in analysis	Median (IQR) age 67 (60 to 73) years 70.2% male Median follow-up time 14 (8 to 22) months Median time from last endoscopy to Cytosponge test 38 (29 to 48) months 83.7% demonstrated IM on previous endoscopic biopsies	124 of 608 tests insufficient <u>Biopsy results</u> TFF3 negative (n = 136): No IM 48, non-dysplastic BO 80, indefinite for dysplasia 5, LGD 3 TFF3 positive only (n = 48): No IM 8, non-dysplastic BO 37, indefinite for dysplasia 1, LGD 2 Atypia only (n = 179): No IM 11, non-dysplastic BO 121, indefinite for dysplasia 15, LGD 20, HGD 6, intramucosal carcinoma 2, adenocarcinoma 3, SCC 1 p53 only (n = 24): No IM 0, non-dysplastic BO 17, indefinite for dysplasia 1, LGD 5, intramucosal carcinoma 1 Atypia and p53 (n = 97): No IM 0, non-dysplastic BO 35, indefinite for dysplasia 11, LGD 18, HGD 17, intramucosal carcinoma 9, adenocarcinoma 7	• Pilot was conducted during the COVID-19 pandemic when usual endoscopy services were disrupted. • Patients were recruited for capsule sponge testing if previously entered in local Barrett's surveillance programmes, where prior endoscopy demonstrated macroscopic changes consistent with BO. The presence of IM on endoscopic biopsies was not considered a prerequisite for entry into surveillance.
Chien & Glen (2025) Prospective cohort study	n = 3,359 Pre-intervention group, n = 1,568 Implementation group, n = 1,791 (capsule sponge	Pre-intervention group: Median (IQR) age 65 (57 to 72) years 64.3% male Proton-pump inhibitor use 95.6% IM on last endoscopic pathology results 82.1%	<u>Dysplasia rates in pre-intervention group vs implementation group</u> Indefinite for dysplasia 54 (3.4%) vs 66 (3.7%), p = 0.707 LGD 53 (3.4%) vs 39 (2.2%), p = 0.033 HGD 15 (1.0%) vs 9 (0.5%), p = 0.151 Intramucosal cancer 3 (0.2%) vs 5 (0.3%), p = 0.731	• Patients were invited to undertake capsule sponge testing in lieu of surveillance endoscopy in the absence of red flag symptoms. • The presence of IM on endoscopic biopsies was not considered a

Evidence source	Number of participants	Population	Detection rates	Comments
	cohort, n = 920; endoscopy only cohort, n = 871)	Median (IQR) time from last endoscopic surveillance 25 (23 to 34) months Implementation group: Median (IQR) age 66 (57 to 73) years 63.9% male Proton-pump inhibitor use 94.8% IM on last endoscopic pathology results 76.8% Median (IQR) time from last endoscopic surveillance 35 (27 to 45) months	Invasive cancer 4 (0.3%) vs 9 (0.5%), p = 0.280 <u>Dysplasia rates in implementation group (endoscopy only cohort vs capsule sponge cohort)</u> Indefinite for dysplasia 46 (5.3%) vs 20 (2.2%), p < 0.001 LGD 25 (2.9%) vs 14 (1.5%), p = 0.051 HGD 6 (0.7%) vs 3 (0.3%), p = 0.331 Intramucosal cancer 1 (0.1%) vs 4 (0.4%), p = 0.375 Invasive cancer 7 (0.8%) vs 2 (0.2%), p = 0.100 <u>Dysplasia rates in pre-intervention group and endoscopy only cohort combined vs capsule sponge cohort</u> Indefinite for dysplasia 100 (4.1%) vs 20 (2.2%), p = 0.007 LGD 78 (3.2%) vs 14 (1.5%), p = 0.008 HGD 21 (0.9%) vs 3 (0.3%), p = 0.101 Intramucosal cancer 4 (0.2%) vs 4 (0.4%), p = 0.151 Invasive cancer 11 (0.5%) vs 2 (0.2%), p = 0.331	prerequisite for entry into surveillance. • Patients with red flag symptoms were excluded from capsule sponge testing, but were included in the endoscopy only group. • The capsule sponge cohort had a longer median time from last endoscopic surveillance than the endoscopy only cohort (38 vs 31 months, p < 0.001). Patients may have been more likely to opt for capsule sponge testing if their surveillance interval was delayed.
Pilonis et al. (2022) Cross-sectional study and prospective cohort analysis	Cross-sectional study (n = 891) Prospective cohort analysis (n = 223)	Cross-sectional study: Training cohort n = 557 Median (IQR) age 65 (59 to 72) years 81% male 98% white, 2% other ethnicity Median (IQR) BO maximum segment length 5 (3 to 8) cm Median (IQR) BO circumferential length 3 (1 to 6) cm Median (IQR) BMI 28.25 (25.61 to 31.07) Validation cohort n = 334 Median (IQR) age 67 (58 to 73) years 75% male Ethnicity NR	Cross-sectional study: Diagnosis of HGD or cancer at endoscopy: positive biomarker 47%, endoscopy surveillance alone 14% Prospective cohort analysis: Diagnosis of HGD or cancer at endoscopy in those with aberrant p53 expression and cellular atypia: 64%	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Endoscopies were performed on the same day as Cytosponge (BEST2) or within 2 months of Cytosponge (BEST3). • Participants recruited in the prospective cohort analysis had their regular Barrett's surveillance delayed by Covid-19.

Evidence source	Number of participants	Population	Detection rates	Comments
		Median (IQR) BO maximum segment length 3 (2 to 6) cm Median (IQR) BO circumferential length 1 (0 to 4) cm Median (IQR) BMI 27.90 (25.20 to 30.81) Prospective cohort analysis: Median age 69 (IQR 60 to 74) years 74% male Ethnicity NR Median (IQR) BO maximum segment length 3 (2 to 6) cm Median (IQR) BO circumferential length 1 (0 to 4) cm Median (IQR) BMI 26.90 (24.12 to 29.30)		
Tan et al. (2025) Prospective cohort study	n = 910	n = 910 Consecutive patients undergoing BO surveillance from 13 hospitals in the UK who participated in the DELTA study and the NHS England implementation pilot study. Median (IQR) age 68 (60 to 74) years 76% male Histology at baseline: • Non-dysplastic BO 90% • Indefinite for dysplasia 1% • Crypt dysplasia < 1% • LGD 5% • HGD or intramucosal carcinoma 3% • Adenocarcinoma (≥ T2) 1%	<u>Endoscopy results</u> Low-risk group (n = 489): non-dysplastic BO 478, crypt dysplasia 1, LGD 8, HGD or intramucosal carcinoma 2 Moderate-risk group (n = 283): non-dysplastic BO 255, indefinite for dysplasia 5, crypt dysplasia 1, LGD 15, HGD or intramucosal carcinoma 7 High-risk group (n = 138): non-dysplastic BO 84, indefinite for dysplasia 2, crypt dysplasia 1, LGD 24, HGD or intramucosal carcinoma 22, adenocarcinoma (≥ T2) 5 <u>Prevalence estimates by capsule sponge and clinical risk groups</u> Any dysplasia: Low risk 2.2% (95% CI 1.2 to 4.1%), moderate risk 8.1% (95% CI 5.3 to 12.1%), high risk tier 2 26.1% (95% CI 18.5 to 35.5%), high risk tier 1 85.2% (95% CI 65.4 to 95.1%)	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • The DELTA study and the NHS England implementation pilot study followed the same protocol. • Patients were assigned to low- or moderate-risk groups at baseline based on clinical risk factors and previous BO findings. Patients were escalated to the high-risk group after capsule sponge testing if their results showed any of atypia, atypia of uncertain significance, equivocal p53, or aberrant p53 expression. • Study took place during Covid-19 pandemic when endoscopy services were disrupted. • Some patients had more than one endoscopy follow-up, for example for

Evidence source	Number of participants	Population	Detection rates	Comments
			HGD or cancer: Low risk 0.4% (95% CI 0.1 to 1.6%), moderate risk 2.5% (95% CI 1.1 to 5.2%), high risk tier 2 10.8% (95% CI 6.0 to 18.5%), high risk tier 1 55.6% (95% CI 35.6 to 74.0%)	indefinite for dysplasia or first diagnosis of LGD, which followed the clinical standard of a repeat at 6 months. • 'High risk tier 1' defined as those positive for both glandular atypia and p53 (not including uncertain or equivocal results, respectively). All other patients with positive biomarkers defined as 'high risk tier 2'.

Abbreviations: BEST: Barrett's oEsophagus Screening Trial; BO: Barrett's oesophagus; CI: confidence interval; CS: capsule sponge; DELTA: integrateD diagnostic solution for EarLy deTection of oesophageal cAncer; HGD: high-grade dysplasia; IM: intestinal metaplasia; IQR: interquartile range; ITT: intention-to-treat; LGD: low-grade dysplasia; NR: not reported; OAC: oesophageal adenocarcinoma; OR: odds ratio; RCT: randomised controlled trial; RR: risk ratio; SCC: squamous cell carcinoma; SR: systematic review; TFF3: Trefoil factor 3; UGI: upper gastrointestinal

5.4 Time to diagnosis

The only reporting of time from a capsule sponge test to diagnosis is in the HTA by SHTG (2023). This figure comes from evaluation of primary data from NHS Scotland, which does not appear to have been reported in the subsequent peer-reviewed publications. SHTG reports that the mean time to diagnosis for patients under surveillance for Barrett's oesophagus, with an urgent endoscopy referral triggered by their Cytosponge result, was 110.84 ± 85.23 days (n = 261). No comparative data on time to diagnosis was identified.

5.5 Time to treatment

The only reporting of time to treatment also came from analysis of primary data from NHS Scotland by SHTG (2023). SHTG reports that 'for high risk patients with Barrett's oesophagus under surveillance (n = 299), the average time from last endoscopy to treatment (1,538 days) was longer than the time from Cytosponge to treatment (244 days)'. However, these data should not be interpreted to imply people received treatment quicker when they had been tested using Cytosponge, because these are data from the same cohort of patients, who have received Cytosponge testing as a triage test instead of receiving their next routine endoscopic screening as standard. SHTG state that they cannot comment on whether Cytosponge leads to quicker diagnosis and treatment as they do not have comparator data, however, the above statement could be interpreted as a comparison. These data were also not reported in subsequent peer-reviewed publications.

5.6 Safety and adverse events

In their RCT, Fitzgerald et al. (2020) reported 142 participants (9%) out of 1,654 that successfully swallowed the Cytosponge experienced an adverse event. This included 63 participants (4%) reporting sore throat requiring medication or causing eating problems and one serious adverse event of sponge detachment, which required endoscopic retrieval. Adverse events in the control group were not reported.

Chien et al. (2024a) reported two sponge detachments out of 1,385 Cytosponge tests. In a real-world study from England (Angel et al. 2025), there was one sponge detachment with the Cytosponge device; this prompted a switch to using EndoSign instead and no further detachments occurred. There were no other reported complications.

The report by SHTG mentioned the urgent field safety notice issued by the MHRA for Cytosponge in June 2023 (MHRA 2023). Fifteen batches of Cytosponge were recalled due to increased risk of sponge detachment. In the prior six months (December 2022 to June 2023), 13 patients globally had reported sponge detachment during the Cytosponge procedure. Urgent endoscopy was performed for all to remove the sponge and there were no further adverse events related to this. Experts have highlighted that all detachments reported in this EAR occurred with the Cytosponge device and that EndoSign was developed to address the issue of sponge detachment.

Ross-Innes et al. (2015) reported 16.7% of patients were found to have bleeding from a Cytosponge abrasion during endoscopy. Three serious adverse events were also reported, however none of these were due to the Cytosponge device. Kadri et al. (2010) and Gourgiotis et al. (2025) reported no serious adverse events, and Norton et al. (2025) also reported no adverse events.

5.7 Quality of life

No evidence on quality of life (QoL), using validated QoL measures, was identified.

5.8 Ongoing studies

One potentially relevant ongoing study was identified that is due to complete within the next 12 months. The intervention is not specified in the trial registry, therefore, we contacted the lead researcher to ask what device and biomarkers are being used. It was confirmed that Cytosponge is the device being utilised. The biomarker analysis is being conducted by Exact Sciences Corporation and includes methylation markers; however, the lead researcher was unsure whether TFF3 and p53 are also included.

Table 4 – Summary of ongoing primary studies

Study information	Status	Research question and outcome measures
Registration: NCT06335966 Barrett's Esophagus Screening Towards Rural Referral Pathways: Screening for Esophageal Cancer in Rural Oregon Without Endoscopy **Country:** USA **Target recruitment**: 110 participants **Follow-up:** 8 months	Active, not recruiting **Last updated:** 07 March 2025	The BEST-RPP study aims to evaluate the acceptability and feasibility of using swallowable oesophageal cell-collection devices to screen for Barrett's oesophagus and oesophageal carcinoma in rural primary care clinic settings in Oregon, USA. **Population:** Patients with suspected Barrett's oesophagus or at risk for oesophageal cancer. Patients who receive primary care in a rural settings and are in need of screening for Barrett's oesophagus or oesophageal cancer. **Intervention:** Screening with swallowable oesophageal cell-collection devices **Comparator:** N/A **Primary Outcome Measures:** Feasibility of the use of swallowable cell-collection devices, patient acceptability **Secondary Outcome Measure:** Access (time to full diagnostic work up for patients with positive cell-collection device results)

5.9 Certainty of the evidence

- The majority of evidence is from studies that involved employees, developers, or founders of Cytosponge or Cyted. There is potential for some unknown level of bias in these studies.
- Most of the identified studies examined Cytosponge, with three studies examining EndoSign. Experts have indicated evidence using the biomarkers TFF3, p53 and cellular atypia can be generalised between these two devices. However, the evidence cannot be generalised to other device types or biomarkers.
- One RCT was identified and all the remaining evidence was from observational studies. There is, therefore, little evidence from randomised trials, though RCTs may not be the most

appropriate trials for assessing diagnostic accuracy. The comparator arm in this RCT did not match the PICO of this review, which was endoscopy in all.

- Less than 25% of the RCT's intervention group provided a sufficient Cytosponge test result and having a Cytosponge test was optional, leading to possible selection bias in this study. Another limitation is that those with negative Cytosponge results were not offered endoscopies.
- There is the potential for double reporting of patients from the studies involved in the DELTA trial and NHS England evaluation (Angel et al. 2025, Gourgiotis et al. 2025, Pilonis et al. 2022, Tan et al. 2025).
- There is a lack of longer-term outcomes in the literature, such as mortality and survival. The BEST4 trial is underway, which aims to investigate whether capsule sponge-biomarker technology reduces mortality from oesophageal cancer (NIHR135565). The end date for this study is September 2035.
- Due to some studies not performing endoscopies on those with negative capsule sponge test results, PPV was the only diagnostic accuracy measure that could be calculated. In other studies, there may have been some selective reporting of diagnostic accuracy, as they did not report all of sensitivity, specificity, PPV, and NPV when participant data would have allowed this.
- Ideally, participants should receive the index test and reference test at the same time when assessing diagnostic accuracy; this does not appear to have happened in several studies. In the real-world evidence studies, it is not always clear how soon after capsule sponge testing the endoscopies were performed, but this was often several weeks to months later.
- The lack of follow-up endoscopies on patients with negative capsule sponge test results in some studies means it is not clear how many of these were true negative results.

6. Cost effectiveness

6.1 Economic literature review

Appendix 4 summarises the selection of articles for inclusion in the evidence review. The titles and abstracts of 1,513 records identified in the search for this research question were screened and 10 records were deemed potentially relevant. The full texts of these records were reviewed against the inclusion/exclusion criteria and six were excluded.

The NICE guideline NG231 on the monitoring and management of Barrett's oesophagus and stage 1 OAC (NICE 2023a) was identified and excluded as no economic analysis was conducted for non-endoscopic surveillance techniques. A study which conducted a cost-effectiveness analysis (Aoki et al. 2024) was excluded as the patient population was focused on the general population which did not meet the inclusion criteria. Another study (Swart et al. 2021) was excluded as the analysis compared Cytosponge to referral for endoscopy as deemed necessary by a primary care physician. This did not meet the inclusion criteria as not all patients in the comparator arm were offered an endoscopy.

Two studies were selectively excluded. A cost-utility analysis considering 2007/08 costs (Benaglia et al. 2013) was excluded as more recent studies conducting analyses of this type were available. A budget impact analysis (BIA) conducted within an SHTG assessment (SHTG 2023) was also excluded due to the availability of several economic studies conducting a cost-utility analysis.

One identified study (Matchett et al. 2025) was a systematic review on cost-effectiveness analyses of Barrett's oesophagus screening strategies. The study itself was excluded from the economic review and studies within the systematic review were compared against the inclusion/exclusion criteria. Four studies met the inclusion/exclusion criteria, all of which were already identified in the search for this research question (Benaglia et al. 2013, Heberle et al. 2017, Sami et al. 2021, Swart et al. 2021).

Four studies were included and are summarised in Table 5. All studies conducted a cost-utility analysis: three focused on capsule sponge devices used for initial diagnostic screening, and one focused on using the capsule sponge device for surveillance. One study was directly applicable to the research question, and three were partially applicable. All studies had potentially serious limitations.

6.1.1 Capsule sponge device used for initial diagnostic screening

The study directly applicable to the research question (IQVIA 2023) considered a cohort of low-risk GORD patients waiting for an endoscopy via referral through usual care. The study used epidemiological and statistical techniques on clinical data to evaluate the real-world impact of Cytosponge. A Markov model was developed to evaluate the cost effectiveness of Cytosponge when used as a diagnostic triage tool in secondary care, comparing Cytosponge testing to endoscopy-only. Cytosponge patients with a positive result received a confirmatory endoscopy. The analysis took a UK NHS perspective and evaluated outcomes over a lifetime horizon.

The model comprised of two phases. The first phase represented the short-term diagnostic pathway, and the second phase represented the post-diagnostic lifetime pathway where all patients underwent appropriate monitoring and surveillance. Baseline characteristics, time to diagnosis and adherence were based on the clinical findings of this study. Performance characteristics of Cytosponge were informed by the BEST2 trial (Ross-Innes et al. 2015), and endoscopy was assumed to be perfectly accurate. The cost of resources associated with

Cytosponge and endoscopy were sourced from published data (NHS England 2022, NICE 2020). Health state costs and utilities were sourced from previous studies.

Results of their base case analysis estimated cost savings of £422 per patient triaged using Cytosponge compared with endoscopy alone. However, this was also associated with a reduction of 0.0041 quality-adjusted life years (QALYs). In monetary terms, the Cytosponge approach corresponded to a net monetary benefit (NMB) of £339, at a cost-effectiveness threshold of £20,000 per QALY. At this threshold, the study concluded that these results indicate endoscopy-only screening was not cost effective compared to Cytosponge.

Uncertainties were explored in sensitivity and scenario analyses. In probabilistic sensitivity analysis (PSA), Cytosponge had a probability of ~65% of being cost effective. In one-way sensitivity analysis, costs of endoscopy and Cytosponge testing were identified as key model drivers. In scenario analysis, a scenario assuming equal efficacy estimated that the Cytosponge approach has cost savings of £526 per patient. A further scenario assuming full adherence to endoscopy referral estimated the NMB would reduce to £300. Following this assumption, the NMB would increase to £361 if clinicians followed the guidance precisely in assigning subsequent actions within the Cytosponge arm.

Potentially serious limitations of this study were identified. These limitations include possible biases from the clinical data used to inform the diagnostic pathway and comparator arm, uncertainties in how representative the clinical data is to the modelled population and concerns in the quality of health state utility values taken from published literature. Further details of the studies limitations are provided in Table 5.

Another study (Sami et al. 2021) developed a Markov model to compare the cost effectiveness of six screening strategies with each other, and with no screening, from a third-party payer perspective based on Medicare reimbursement rates in the US, over a 40-year horizon. One of the populations considered in their analysis consisted of a cohort of white men aged 50 years with chronic GORD symptoms (GORD-based population). Screening strategies included sedated endoscopy, transnasal endoscopy (hospital-based and mobile-based), Cytosponge + TFF3, EsophaCap + MDMs, and exhaled volatile organic compounds, where it was assumed a positive finding in the latter five strategies were confirmed by sedated endoscopy. Test performance characteristics, participation rate, and Barrett's oesophagus prevalence were informed from a range of published literature and assumptions by the authors (Benaglia et al. 2013, Iyer et al. 2018, Kadri et al. 2010, Peters et al. 2020, Rubenstein et al. 2010, Sami et al. 2015, Sami et al. 2019, Shariff et al. 2016, Visrodia et al. 2018). Sedated endoscopy was assumed to be perfectly accurate. Direct costs were based on Medicare reimbursement rates (Russell et al. 1996), and the cost of testing and treatments were sourced from a range of previous economic studies (Heberle et al. 2017, Hur et al. 2012, Moriarty et al. 2018), assumptions by the authors and the GI Endoscopy Coding and Reimbursement Guide (Cook Medical 2018). Health state utilities were informed by a previous NICE clinical guideline (CG 106) for Barrett's oesophagus ablative therapy (NICE 2010), which is now obsolete.

Total costs and QALYs of their base case results are presented in Table 3. Based on their findings, incremental cost-effectiveness ratios (ICERs) comparing Cytosponge + TFF3 to strategies using endoscopy have been calculated by HTW. Cytosponge + TFF3 was estimated to be less costly and more effective (i.e. dominant) compared to sedated and hospital-based transnasal endoscopy, and when compared to mobile-based transnasal endoscopy, the ICER was £12,539. When these are further explored in a scenario analysis assuming 100% participation, no scenario comparing Cytosponge + TFF3 to strategies using endoscopy were estimated to be cost effective.

This study was not directly applicable as the UK perspective was not considered and was associated with potentially serious limitations. As test performance characteristics are informed by multiple sources, the comparability of screening strategies used may be limited, as well as

the generalisability to the modelled population. Furthermore, the prevalence of Barrett's oesophagus is based on a 2010 study which may be outdated due to changes in population health. Additional details of the studies limitations are provided in Table 5.

A third study (Heberle et al. 2017) used two validated models of OAC progression to estimate the cost effectiveness of using Cytosponge in first-line screening compared to endoscopy-only screening, from a US societal perspective. The model evaluated lifetime outcomes of men aged 60 years, with GORD symptoms and without an OAC diagnosis, from a 1950 US birth cohort. Performance characteristics of Cytosponge were derived from the BEST2 trial (Ross-Innes et al. 2015) and estimated from the literature for endoscopy (Provenzale et al. 1999). Rates of endoscopy complications, post-treatment recurrence and dysplasia eradication were sourced from the literature (Falk et al. 1997, Silvis et al. 1976, Wolf et al. 2014a, Wolf et al. 2014b). Endoscopy and radiofrequency ablation (RFA) treatment costs were estimated from Medicare reimbursement rates, and Cytosponge costs were estimated from the Centers for Medicare and Medicaid Services (2017) and information from the manufacture. QoL utility values were estimated from the literature; however, the sources of these estimates are not reported.

Base case results estimated the ICER for endoscopy-only screening compared with Cytosponge screening was £75,507 from one model, and £228,792 from the other, deeming endoscopy-only screening not cost effective based on the cost-effectiveness threshold established in this study. A PSA was performed on one of the models which estimated the ICER for this comparison ranged from £162,585 to £293,509. A one-way sensitivity analysis identified endoscopy-only screening becomes cost effective when Cytosponge testing costs exceed £418 in one model, and £155 in the other.

This study was not directly applicable as the UK perspective was not considered and was associated with potentially serious limitations. The authors noted a significant limitation was the uncertainty of parameters, however, it was noted that this is mitigated via use of best available parameter estimates from the literature and sensitivity analysis. The comparability between strategies is a concern as diagnostic performance outcomes of endoscopy and Cytosponge were estimated from different studies. Furthermore, the performance characteristics and complication rates of endoscopy were estimated using studies from 1999, 1997 and 1974, which may not accurately reflect current practices due to technological advancements and may lack representativeness of the modelled population. Additional details of the studies limitations are provided in Table 5.

6.1.2 Capsule sponge device used for surveillance

A prospective study (Eluri et al. 2022), conducted across US and UK tertiary care referral centres, determined diagnostic outcomes of Cytosponge for the detection of residual or recurrent Barrett's oesophagus in patients post-complete eradication of intestinal metaplasia (CEIM) scheduled for further therapy or surveillance. Using these outcomes, a microsimulation model was developed to evaluate the cost effectiveness of various surveillance strategies over a lifetime. A cohort of male patients aged 68 years was considered and surveillance strategies included endoscopy-only, Cytosponge-only, and strategies where endoscopy and Cytosponge were alternated. Patients with a positive Cytosponge result received a confirmatory endoscopy two months later. Endoscopy misdiagnosis probabilities were sourced from literature (Pasricha et al. 2014) and surveillance frequency was informed by a clinical practice review and ACG guidelines (Shaheen et al. 2016, Wani et al. 2016). The source of other resource use, costs, utilities, and various treatment-related inputs were not reported. We have assumed the analysis takes the perspective of the US healthcare system; however, this is unclear.

Results of their base case analysis estimated all strategies using Cytosponge are less costly and more effective (i.e. dominant) when compared to endoscopy-only, with Cytosponge-only being the most dominant strategy. In one-way sensitivity analysis, the Cytosponge-only strategy remained dominant when sensitivity and specificity was set to a lower threshold of 50%.

This study was not directly applicable as the UK perspective was not considered and was associated with potentially serious limitations. As endoscopy misdiagnosis rates were derived from a different study, the comparability between surveillance strategies may be reduced. Additionally, the generalisability of the economic findings is a concern as baseline characteristics (i.e. age, proportion male) applied to the model are not fully aligned to the characteristics of patients who participated in the clinical study. Model uncertainties were not fully explored, where only a one-way sensitivity analysis of Cytosponge performance characteristics was performed. Additional details of the studies limitations are provided in Table 5.

Table 5 – Summary of included economic studies (Eluri et al. 2022, Heberle et al. 2017, IQVIA 2023, Sami et al. 2021)

Study details	Study population and design	Data sources	Results	Quality assessment
Capsule sponge device used for initial diagnostic screening				
Author and year: IQVIA (2023) **Country:** United Kingdom **Type of economic analysis:** Cost-utility analysis **Perspective:** UK NHS **Currency:** UK pounds **Price year:** 2020 – 2022 (assumed) **Time horizon:** Lifetime **Discounting:** No discounting reported **Potential conflict of interest:** None declared	**Population:** Low-risk gastro-oesophageal reflux disease (GORD) patients waiting for an endoscopy. **Cohort settings:** Based on the clinical findings of this study: • Mean age 52 years. • 42.4% of the cohort were men. **Intervention:** Cytosponge followed by confirmatory endoscopy for patients with positive screening results **Comparator:** Endoscopy-only **Study design** NHS England piloted Cytosponge as a triaging pathway in secondary care where epidemiological and statistical techniques were used to evaluate the real-world impact of Cytosponge. A Markov model, consisting of two phases, was developed to evaluate cost effectiveness of Cytosponge when used as a diagnostic triage tool in secondary care for the diagnosis	**Source of baseline and effectiveness data:** Baseline characteristics and time to diagnosis were based on the clinical findings of this study. Performance characteristics of Cytosponge was based on the BEST2 trial (Ross-Innes et al. 2015). The sensitivity of endoscopy-only was assumed to be 100%. Assumptions regarding protocol adherence were based on the observed outcomes of this study. Transition probabilities were sourced from published literature (Sami et al. 2021). **Source of resource use and cost data:** The model included the following direct costs: diagnostic test acquisition and administration costs, subsequent treatment acquisition and administration costs, and adverse event costs. Cytosponge testing resource use was based the clinical findings of this study, with costs based on	**Base case results** Total costs and quality-adjusted life years (QALYs) are presented per patient. **Total costs** Intervention: £9,858 Comparator: £10,280 Incremental: -£422 **Total QALYs** Intervention: 14.5537 Comparator: 14.5578 Incremental: -0.0041 The corresponding net monetary benefit (NMB) was £339 at a willingness to pay threshold of £20,000. The reported ICER for endoscopy-only screening compared with Cytosponge was £102,188, implying endoscopy-only is not cost effective at a willingness to pay threshold of £20,000. **Scenario analysis** A cost-minimisation scenario analysis was conducted assuming equal efficacy (i.e. equal time-to-diagnosis and sensitivity as in the endoscopy-only programme).	**Applicability** Directly applicable **Limitations** Some potentially serious limitations were identified: • Data informing the short-term diagnostic pathway may include confounding, selection, information, and triaging biases. • Due to the sample size in the comparator arm, a propensity score weighting approach was undertaken which was subject to challenges and potential bias. • As the study took place over the COVID-19 pandemic, it's findings may not be fully representative of usual healthcare settings. • There is uncertainty related to the representativeness of the performance characteristics based on the BEST2 trial, as this was a different use case. • Health state utility values, taken from a previously published analysis, are not based on directly elicited values from patient reports, • There are uncertainties in transition probabilities considered. • A full list of the inputs, sources, assumptions, and alternate

Study details	Study population and design	Data sources	Results	Quality assessment
	of Barrett oesophagus (BO). Phase 1 represented the short-term diagnostic pathway. Phase 2 represented the post-diagnostic lifetime pathway where all patients underwent appropriate monitoring and surveillance.	published data from the NHS, NICE and PSSRU (NICE 2020). Endoscopy costs were sourced from the national schedule of NHS costs (NHS England 2022). A band 7 nurse was considered for both Cytosponge and endoscopy. Health state costs were sourced from previous studies, which included treatment costs for BO and oesophageal adenocarcinoma (OAC). The sources of these costs are not reported. **Source of resource quality of life data:** Health state utility values were sourced on previous studies. The sources of these values are not included in this report.	This resulted in a cost saving of £526 in favour of Cytosponge. In a scenario assuming full adherence to endoscopy referral, the NMB would reduce to £300. Following this assumption, if clinicians followed the guidance precisely in assigning subsequent actions within the Cytosponge arm, the NMB would increase to £361. **Sensitivity analysis** The probabilistic sensitivity analysis (PSA) estimated that the Cytosponge programme has ~65% probability of being cost effective at a willingness to pay threshold of £20,000. A one-way sensitivity analysis was performed varying model parameters by +/- 20%. This identified the costs of endoscopy and Cytosponge testing as key model drivers.	scenarios were not included in this report. They are described in an unpublished economic evaluation technical report.
Author and year: Sami et al. (2021) **Country:** United States **Type of economic analysis:**	**Population:** Two populations were considered: one in patients with GORD symptoms and another independent of GORD symptoms. **Cohort settings:**	**Source of baseline and effectiveness data:** For the GORD-based patient population, an 8% BO prevalence was assumed (Rubenstein et al. 2010).	Results presented in this table only relate to the GORD-based patient population. Comparative results are only presented for Cytosponge + TFF3 (strategy 4) compared to strategies using endoscopy.	**Applicability** Partially applicable because non-UK perspective taken. **Limitations** Some potentially serious limitations were identified:

Study details	Study population and design	Data sources	Results	Quality assessment
Cost-utility analysis **Perspective:** Third-party payer based on Medicare reimbursement rates **Currency:** US dollars (converted to UK pounds [a]) **Price year:** 2020 (assumed) **Time horizon:** 40 years **Discounting:** 3% per year **Potential conflict of interest:** Some authors received research funding and/or consulting fees from companies including Exact Sciences and Medtronic. Some authors are associated with a medical centre that holds a minor equity investment in Exact Sciences.	The model simulated hypothetical cohorts of 500,000 individuals for the following populations: • GORD-based: white men aged 50 years with chronic GORD symptoms. • GORD-independent: general US population aged 50 years. **Comparators:** Six screening strategies were included: 1. Sedated endoscopy 2. Hospital transnasal endoscopy 3. Mobile transnasal endoscopy 4. Cytosponge + trefoil factor 3 (TFF3) 5. EsophaCap + methylated DNA markers (MDMs) 6. Exhaled volatile organic compounds These strategies were compared with no screening and compared with each other. For strategies 2 – 6, it was assumed a positive finding was confirmed by sedated endoscopy. **Study design** A Marko model was developed to evaluate the cost effectiveness of BO screening tests in GORD-based and GORD-independent testing scenarios.	Health state transitions were primarily taken from a study which carried out a systematic review and workshop with experts regarding the surveillance of BO (Garside et al. 2006). Transitions from no BO and mortality rates were taken from published literature or sources referenced in previous economic analyses (Benaglia et al. 2013, Inadomi et al. 2003, NICE 2010, Rubenstein et al. 2007, Wu et al. 2014). Test performance characteristics were sourced from published literature (Iyer et al. 2018, Kadri et al. 2010, Peters et al. 2020, Sami et al. 2019, Shariff et al. 2016, Visrodia et al. 2018). Sedated endoscopy was considered as the gold standard test. Following a diagnosis of BO, patients would undergo surveillance using sedated endoscopy. Age-specific mortality probabilities was sourced from the National Vital Statistics Report (US Department of Health and Human Services 2017). Inputs related to the treatment efficacy of endotherapy was sourced from published literature	HTW have calculated ICERs for these comparisons where necessary. **Base case results** Total costs and QALYs are presented per patient. **Total costs** Strategy 1: £325 Strategy 2: £210 Strategy 3: £146 Strategy 4: £192 Strategy 5: £186 Strategy 6: £496 No screening: £56 **Total QALYs** Strategy 1: 18.3768 Strategy 2: 18.3768 Strategy 3: 18.3768 Strategy 4: 18.3805 Strategy 5: 18.4203 Strategy 6: 18.396 No screening: 18.3575 **ICER** Cytosponge + TFF3 (strategy 4) compared to: Strategy 1: Dominant Strategy 2: Dominant Strategy 3: £12,539 **Scenario analysis** A scenario analysis assuming equal 100% participation across all strategies was performed to demonstrate the	• A PSA was not explored. Authors state that this was due to concerns regarding limited data availability for certain model parameters. • Test performance characteristics are informed by multiple sources which may limit the comparability of screening strategies used in the analysis. • Variability in the sources used to inform test performance may limit the generalisability to the modelled population. • The prevalence of BO is based on a 2010 study exploring the age-specific yield of endoscopy for BO. Due to the age of the study, this prevalence estimate may be outdated due to changes in population health. • Possible quality of life reductions due to invasive testing and false positive diagnoses were not considered. However, authors noted that these reductions would have been over periods of time shorter than a single cycle.

Study details	Study population and design	Data sources	Results	Quality assessment
		(Phoa et al. 2014, Shaheen et al. 2011). The proportion of those with symptomatic cancer suitable for surgery and the five-year survival after surgery were sourced from published literature (Garside et al. 2006, NICE 2010, Ovrebo et al. 2012). The participation rate of each testing strategy was based on Sami et al. (2015), Benaglia et al. (2013) and assumptions by the authors. Subtype distribution was sourced from Garside et al. (2006) and Sami et al. (2015). **Source of resource use and cost data:** Direct costs were based on Medicare reimbursement rates estimates (Russell et al. 1996). The cost of endoscopy included procedure costs only and was sourced from the GI Endoscopy Coding and Reimbursement Guide (Cook Medical 2018). Sedation costs were not included. Hospital and mobile transnasal endoscopy testing costs were sourced from a previous economic analysis comparing	relative maximal effectiveness of each strategy. All scenarios comparing Cytosponge + TFF3 (strategy 4) to strategies using endoscopy were not cost effective. **Sensitivity analysis** A one-way sensitivity analysis was performed on all parameters with a reported range. For the GORD-based population, all ICER values (comparing strategies to no screening) remained cost effective.	

Study details	Study population and design	Data sources	Results	Quality assessment
		types of endoscopies (Moriarty et al. 2018). The test cost of Cytosponge + TFF3 was sourced from a previous cost effectiveness analysis (Heberle et al. 2017). The test costs for EsophaCap + MDMs and exhaled volatile organic compounds were based on assumptions from the authors. The cost of endotherapy and cancer-related surgery were sourced from a study focused on the cost effectiveness of radiofrequency ablation (RFA) for BO (Hur et al. 2012). **Source of resource quality of life data:** Health state utilities have been informed by the NICE clinical guideline for BO ablative therapy (NICE 2010).		
Author and year: Heberle et al. (2017) **Country:** United States **Type of economic analysis:** Cost-utility analysis **Perspective:** Societal perspective	**Population:** Patients with GORD symptoms who have not been diagnosed with OAC. **Cohort settings:** A 1950 US birth cohort of men starting at age 20 was simulated. At age 60, the population was restricted to those with GORD symptoms without an OAC diagnosis.	**Source of baseline and effectiveness data:** For Cytosponge, performance characteristics conditional on dysplastic grade, and the bleed rate, were derived from the BEST2 trial (Ross-Innes et al. 2015). For endoscopy, performance characteristics and complication rates were estimated from the literature (Falk et al. 1997,	**Base case results** Total costs and QALYs are presented per 1,000 GORD patients. Results are presented as a range based on the results of the two models. **Total costs** Strategy 1: £1.0M – £1.1M Strategy 2: £1.4M - £1.5M Strategy 3: £0.49M - £0.52M	**Applicability** Partially applicable because non-UK perspective taken. **Limitations** Some potentially serious limitations were identified: • The authors noted a significant limitation of the analysis surrounded around the uncertainty of parameters

Study details	Study population and design	Data sources	Results	Quality assessment
Currency: US dollars (converted to UK pounds [a]) **Price year:** 2015 (assumed) **Time horizon:** Lifetime **Discounting:** 3% per year **Potential conflict of interest:** None declared	**Comparators:** The following screening strategies were included: 1. Cytosponge followed by confirmatory endoscopy for patients with positive screening results 2. Endoscopy-only 3. No screening **Study design** Clinical trial data (BEST2) was incorporated into two validated microsimulation models of OAC progression to estimate the cost effectiveness of using Cytosponge in first-line screening compared with endoscopy-only screening. The two models were: • Model 1: The OAC model from Massachusetts General Hospital • Model 2: The microsimulation screening analysis model from Erasmus University Medical Center and the University of Washington	Provenzale et al. 1999, Silvis et al. 1976). The rates of post-treatment recurrence and dysplasia eradication have been sourced from the literature (Wolf et al. 2014a, Wolf et al. 2014b). **Source of resource use and cost data:** The cost of endoscopy and RFA treatment have been estimated from Medicare reimbursement rates. A cost for Cytosponge was estimated based on communication with the manufacture and Medicare facility payments for comparable diagnostic tests Centers for Medicare and Medicaid Services (2017). **Source of resource quality of life data:** Quality of life utility values by OAC stage and utility decrements for endoscopy, endoscopic eradication therapy (EET), and complications were estimated from the literature. The sources of these values are not reported.	**Total QALYs** Strategy 1: 15,099 – 15,110 Strategy 2: 15,101 – 15,116 Strategy 3: 15,076 – 15,078 The ICER for endoscopic screening (strategy 2) compared with Cytosponge (strategy 1) was £75,507 - £228,792 and deemed not cost effective by the studies willingness to pay threshold. **Sensitivity analysis** A PSA was performed using Model 1. This estimated that endoscopic screening (strategy 2) was not cost effective compared to Cytosponge (strategy 1), with an ICER ranging from £162,585 to £293,509. A one-way sensitivity analysis was performed on key parameters. Endoscopic screening (strategy 2) is cost effective compared to strategy 1, when the total cost of Cytosponge exceeds £418 in Model 1, and £155 in Model 2. In other parameters explored, Model 1 found endoscopic screening (strategy 2) was not cost effective compared to Cytosponge in any scenario. Model 2 found endoscopic screening (strategy 2) to be	including test performance characteristics, complications, quality-of-life adjustments and the natural history of OAC. However, the authors note that this limitation is mitigated via use of best available parameter estimates from the literature and sensitivity analysis. • Cytosponge-based surveillance strategies were not considered in this analysis. The reason this was not considered is because surveillance requires discrimination between nondysplastic BO, low-grade dysplasia (LGD) and high-grade dysplasia (HGD) which requires endoscopic diagnosis. • The authors noted that the base case cost used for Cytosponge could be significantly different once implemented in clinical practice. • A full list of input values used is not reported. • Diagnostic performance outcomes of Cytosponge are based on a case control study (BEST2), which may be subject to selection bias as participants were not randomly allocated. • Diagnostic performance outcomes of endoscopy and Cytosponge were estimated from different studies, which may reduce the comparability between strategies.

Study details	Study population and design	Data sources	Results	Quality assessment
			cost effective compared to Cytosponge in three scenarios explored (lower bound Cytosponge performance characteristics, upper bound RFA effectiveness and lower bound recurrence after RFA).	• The performance characteristics and complication rates of endoscopy were estimated using studies from 1999, 1997 and 1974. This may not accurately reflect current practices due to technological advancements and may lack representativeness of the current patient population.
Capsule sponge device used for surveillance				
Author and year: Eluri et al. (2022) **Country:** United Kingdom/ United States **Type of economic analysis:** Cost-utility analysis **Perspective:** US healthcare system (assumed) **Currency:** US dollars (converted to UK pounds [a]) **Price year:** 2021 (assumed) **Time horizon:** Lifetime **Discounting:**	**Population:** Patients aged 18 years or over with dysplastic BO, LGD, HGD or intramucosal adenocarcinoma, who had undergone at least one round of EET. These patients were scheduled for further ablative therapy or endoscopic surveillance after complete eradication of intestinal metaplasia (CEIM). **Cohort settings:** A hypothetical cohort of 1,000,000 male patients, aged 68 years, assumed to have achieved CEIM after RFA for dysplastic BO, was modelled. **Comparators:** The following surveillance strategies were included: 1. Endoscopy-only surveillance 2. Alternating Cytosponge and endoscopy at each surveillance	**Source of baseline and effectiveness data:** Cytosponge false positive and false negative rates were calculated from this study. The same false negative rate was assumed for nondysplastic BO, LGD, HGD, and OAC. Endoscopy misdiagnosis probabilities were obtained from prior literature (Pasricha et al. 2014). The source of complication rates, EET touch up efficacy and recurrence rates are not reported. **Source of resource use and cost data:** The frequency of surveillance was informed by American College of Gastroenterology (ACG) guidelines for HGD patients (Shaheen et al. 2016) and a clinical practice	**Base case results** Total costs and QALYs are presented per 1,000 patients. HTW have calculated ICER results where the endoscopy-only surveillance strategy (strategy 1) is considered as the comparator. **Total costs** Strategy 1: £8.0M Strategy 2: £7.2M Strategy 3: £7.0M Strategy 4: £6.6M No surveillance: £5.7M **Total QALYs** Strategy 1: 11,839 Strategy 2: 11,842 Strategy 3: 11,843 Strategy 4: 11,844 No surveillance: 11,734 **ICERs (vs. strategy 1)** Strategy 2: Dominant	Applicability Partially applicable because non-UK perspective taken. Limitations Some potentially serious limitations were identified: • Baseline characteristics (i.e. age, proportion male) applied to the model are not fully aligned to the characteristics of patients who participated in the clinical study, which may reduce the generalisability of the findings. • Endoscopy misdiagnosis probabilities were derived from a different study, which may reduce the comparability between surveillance strategies. • It is unclear if endoscopy misdiagnosis probabilities are applied to those having a confirmatory endoscopy following a positive Cytosponge result.

Study details	Study population and design	Data sources	Results	Quality assessment
No discounting **Potential conflict of interest:** Several authors received research funding and/or served as consultants for medical companies, and some are named on related patents and hold shares in Cyted Ltd.	3. Alternating Cytosponge and endoscopy at every third surveillance 4. Cytosponge-only surveillance A natural history comparator with no post-treatment surveillance was also modelled. Patients with a positive Cytosponge result received a confirmation endoscopy two months later. **Study design** A prospective study was conducted in five tertiary care referral centres across the United Kingdom and United States, where diagnostic outcomes of Cytosponge to detect residual or recurrent BO after RFA was determined. A microsimulation model assessed cost effectiveness outcomes of various surveillance strategies.	review for LGD patients (Wani et al. 2016). The source of other resource use and US costs used are not reported. **Source of resource quality of life data:** No disutility for Cytosponge surveillance was assumed. The source of quality of life utility values used are not reported.	Strategy 3: Dominant Strategy 4: Dominant Compared to the endoscopy-only surveillance strategy, strategies where Cytosponge is used are estimated to be less costly and more effective (i.e. dominant). The Cytosponge-only strategy was estimated to be the most dominant strategy. **Sensitivity analysis** A one-way sensitivity analysis was performed to explore the uncertainty related to the diagnostic accuracy of Cytosponge. The Cytosponge-only strategy remained dominant when sensitivity and specificity was set to a lower threshold of 50%.	• No disutility is assumed for Cytosponge surveillance whilst disutility related to endoscopy is considered, potentially biasing results in favour of Cytosponge. • Cost, utility and some resource use sources are not reported. • Complication rates, EET touch up efficacy, recurrence rates sources are not reported. • One-way sensitivity analysis only explores the uncertainly of the performance characteristics for Cytosponge. The uncertainty of other model parameters was not explored. • A PSA was not explored.

Abbreviations: ACG: American College of Gastroenterology; BO: Barrett Oesophagus; EET: endoscopic eradication therapy; GORD: gastro-oesophageal reflux disease; HGD: high-grade dysplasia; ICER: incremental cost-effectiveness ratio; LGD: low-grade dysplasia, NMB: net monetary benefit; OAC: oesophageal adenocarcinoma; PPI: proton pump inhibitor; PSA: probabilistic sensitivity analysis; QALY: quality-adjusted life year; RFA: radiofrequency ablation; TFF3: trefoil factor 3

[a] Costs converted to UK pounds using purchasing power parities (OECD 2024) for the price year of each study. Costs have not been inflated to current values.

6.2 HTW cost utility analysis

HTW researchers developed an economic model to evaluate the cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage OAC in people with chronic reflux, compared to endoscopic biopsy.

A separate model for the surveillance population was not developed. While diagnostic accuracy evidence for this population exists, studies have only assessed single-timepoint performance, and it is uncertain whether accuracy would be maintained across repeated rounds of surveillance. In addition, there is uncertainty in the long-term disease progression following endotherapy treatment for Barret's oesophagus. The need for additional assumptions around surveillance intervals, disease progression risks, and repeat test performance led to a focus on the chronic reflux population for this evaluation, where available disease progression models are more established.

For the chronic reflux population, a hybrid decision tree and Markov model was developed to estimate the incremental costs and QALYs between intervention and comparator arms. The decision tree captured short-term diagnostic outcomes, and the Markov model captured long-term outcomes related to disease progression, costs, QoL and mortality. The structure of the Markov model closely follows previous cost-utility analyses developed in this disease area (Benaglia et al. 2013, Sami et al. 2021, Swart et al. 2021). The model took the perspective of NHS Wales and personal social services (PSS). Analyses were conducted over a lifetime horizon and future costs and benefits were discounted at a rate of 3.5% per annum. The following diagnostic strategies are included in the base case economic model:

1. Cytosponge testing where those with positive biomarker results receive an endoscopic biopsy (intervention arm)
2. Endoscopic biopsy (comparator arm)

All endoscopic procedures are carried out in secondary care. In the model's base case, patients in the intervention arm undergo Cytosponge testing in primary care. This care setting is explored in scenario analysis.

An overview of the model structure is shown in Figure 2. A cohort of 1,000 people with chronic reflux enter the model, where they undergo diagnostic testing. In the intervention arm, those with negative results (i.e. true and false negatives) are assumed not to have any further testing and enter the Markov model. Patients with positive results (i.e. true and false positives) undergo confirmatory endoscopic biopsy before entering the Markov model. Any false positive cases from the capsule sponge test are confirmed not to have Barrett's oesophagus at this stage. It is assumed those who are unable to swallow the capsule sponge device or experience sponge detachment receive an endoscopic biopsy in secondary care. In the comparator arm, endoscopic biopsy is performed to directly confirm cases of Barrett's oesophagus before they enter the Markov model.

Following diagnostic testing, patients are distributed between their corresponding health states in the Markov model (described in Appendix 6), including no Barrett's oesophagus, nondysplastic Barrett's oesophagus (NDBO), low-grade dysplasia (LGD), high-grade dysplasia (HGD), and early-stage OAC. Long-term disease progression to more severe health states, including late-stage OAC and death, is tracked across annual cycles, with associated costs and utilities captured throughout the modelled time horizon. Treatment costs for proton pump inhibitor (PPI) therapy are not considered as it is assumed all patients receive this due to their underlying chronic reflux. Patients diagnosed with LGD, HGD or early-stage OAC are treated with endotherapy which aims to completely eradicate dysplasia. Endotherapy may also result in complete eradication of any concurrent intestinal metaplasia. Due to this treatment effect, patients could transition to the no Barrett's oesophagus or NDBO health states. Patients progressing to late-stage OAC are

assumed to lead directly to clinical intervention due to the presence of symptoms. These patients receive oesophagectomy or palliative cancer treatments depending on if they are suitable for surgery. Late-stage OAC patients not suitable for oesophagectomy are assumed to transition to death in the subsequent model cycle after entering this health state.

Patients diagnosed with Barrett's oesophagus are assumed to receive endoscopic surveillance every three years, based on NICE recommendations (NICE 2023a), which stops if patients progress to late-stage OAC. Patients identified with LGD, HGD, or early-stage OAC through surveillance, who have not previously undergone endotherapy treatment, proceed to receive endotherapy.

Patients in the intervention arm with a false negative result are assigned to their corresponding true health state following diagnostic testing, and do not receive endotherapy treatment or endoscopic surveillance.

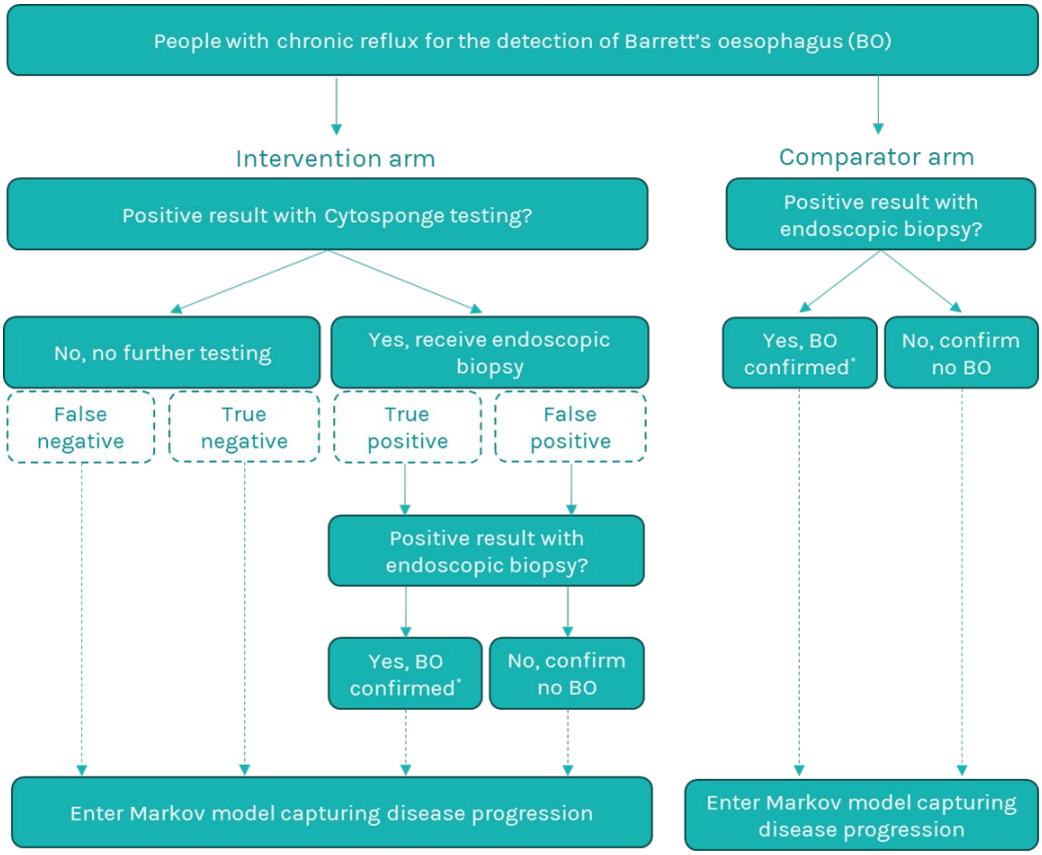


Figure 2 – Model structure overview

The sensitivity and specificity of the Cytosponge test were sourced from a prospective cohort study undertaken in 12 practices in the UK by Kadri et al. (2010), as described in Section 5.2, with values used in the model presented in Table 6. Whilst Kadri et al. (2010) was considered an appropriate source for diagnostic outcomes, limitations of this study should be noted, as described in Appendix 6.

As the study used endoscopic biopsy as the reference standard, the comparator arm is assumed to be perfectly accurate with a sensitivity and specificity of 100%. While experts contacted by HTW noted that this is not perfectly accurate in reality, this assumption was necessary to remain consistent with the evidence base and to allow a relative comparison to test performance.

Table 6 – Diagnostic accuracy inputs

Diagnostic outcome	Mean[a] (%)	SE[b] (%)	Source
Sensitivity	73.3 (44.9 – 92.2)	12.1	Kadri et al. (2010)
Specificity	93.8 (91.3 – 95.8)	1.1	Kadri et al. (2010)

Abbreviations: SE, standard error

[a] 95% CIs displayed in brackets
[b] Sampled from a beta distribution.

Baseline characteristics for age and sex were aligned to the Kadri et al. (2010) study. The reported median age of all participants in the study was 62 years and 45.7% of the participants were male. The baseline prevalence and sub-distribution for Barrett's oesophagus were estimated from two systematic reviews and meta-analyses (Eusebi et al. 2021, Saha et al. 2024) and data from Cancer Research UK (2024). The baseline prevalence of Barrett's oesophagus was estimated to be 8.6% for use in the base case model.

The cost of diagnostic testing with Cytosponge has been informed by the Medtech innovation briefing (MIB240) for Cytosponge for detecting abnormal cells in the oesophagus (NICE 2020). The cost is reported as £280 which includes the cost of the device itself, the immunohistochemical assay test (TFF3), and haematoxylin and eosin stain. The base case analysis assumes the test is administered in primary care by a general practice (GP) nurse, with costs informed by the 2024 PSS Research Unit (PSSRU) report (Jones et al. 2025).

The costs related to endoscopy, endotherapy and oesophagectomy have been sourced from the 2023/24 National Cost Collection data (NHS England 2024). The cost and resource use of palliative cancer treatments have been informed from a previous economic model by Swart et al. (2021). The frequency of endotherapy sessions are based on assumptions and data from an RCT exploring RFA in Barrett's oesophagus with dysplasia (Shaheen et al. 2009). Further resource use considerations are based on previous economic modelling studies in this disease area (Benaglia et al. 2013, Sami et al. 2021, Swart et al. 2021).

Comparative evidence for adverse events related to Cytosponge testing was not identified. However, events which would not apply to the comparator arm were considered in the economic model (i.e. the ability to swallow the capsule sponge and sponge detachment). The proportion of patients failing to swallow the capsule sponge was sourced from a retrospective study performing a patient-level review of five prospective trials assessing Cytosponge (Januszewicz et al. 2019). In the economic model, it is assumed patients failing to swallow the capsule sponge would receive an endoscopic biopsy. The proportion of patients experiencing sponge detachment leading to endoscopic retrieval was sourced from a study reporting outcomes from the BEST3 RCT (Fitzgerald et al. 2020). Patients experiencing sponge detachment are also assumed to receive endoscopic biopsy.

The model estimates effectiveness in terms of QALYs, estimated by combining life year estimates with QoL utility values associated with being in a particular health state. QoL utility values used in the model are closely aligned with values used in previous economic studies (Benaglia et al. 2013, Sami et al. 2021, Swart et al. 2021) and in historic modelling by NICE (NICE CG106, now obsolete (NICE 2010, cited in Sami et al. 2021)). The model also incorporated general population age-adjusted QoL utilities, sourced from the NICE Decision Support Unit (Hernández Alava et al. 2022).

Transitions between health states were aligned with values used in Swart et al. (2021), who utilises values used in previous economic models. For patients receiving endotherapy, a

proportion of patients will experience a treatment effect whereby dysplasia or intestinal metaplasia is completely eradicated and would transition to the no Barrett's oesophagus or NDBO health states. This treatment effect is informed two clinical studies (Phoa et al. 2014, Shaheen et al. 2011) reporting eradication outcomes of RCTs exploring RFA in Barrett's oesophagus with dysplasia.

Mortality rates published by the Office for National Statistics (2024) were used to calculate the annual probability of mortality from any cause and applied to each modelled cycle. Late-stage OAC patients not suitable for oesophagectomy are assumed to transition to death in the subsequent model cycle after entering this health state. The annual probability of mortality for late-stage OAC patients following oesophagectomy is based on a study investigating the long-term survival from OAC after oesophagectomy (Ovrebo et al. 2012).

Base case results are presented in Table 7. Over a lifetime horizon, the results show that use of Cytosponge in primary care, followed by endoscopic biopsy in those with a positive result, is expected to reduce costs by [redacted] per patient with a loss of 0.02 QALYs, compared to endoscopic biopsy in all patients. In this context, where the intervention is less costly and less effective than the comparator, an ICER above the commonly accepted cost effectiveness threshold of £20,000 per QALY is considered cost effective as cost savings outweigh the reduction in health outcomes. All ICERs in this evaluation should be interpreted using this framework.

The base case outcomes correspond to an ICER of [redacted], representing the cost savings per QALY lost. Therefore, Cytosponge is estimated to be cost effective, with results of the PSA indicating a 65.8% probability of being cost effective at this threshold.

Table 7 – Base case health economic results (per-patient)

	Intervention	Comparator	Incremental
Total Costs	[redacted]	£1,403	[redacted]
Total QALYs	11.74	11.76	-0.02
ICER (cost savings per QALY lost)			[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

Deterministic sensitivity analysis identified Cytosponge sensitivity, age and Barrett's oesophagus prevalence as key influential drivers, with the ICER ranging from [redacted] Furthermore, a threshold analysis on these parameters revealed that, independently, a minimum Cytosponge sensitivity of [redacted], a population aged [redacted] years or over and a maximum Barrett's oesophagus prevalence of [redacted] is required to achieve cost effectiveness.

Scenario analyses explored a range of alternative assumptions relating to capsule sponge delivery care setting, intervention costs, population characteristics, methodological approaches and structural features of the model. Scenarios exploring the capsule sponge administered in secondary and community-based care settings had minimal impact on health economic outcomes and no impact on cost effectiveness conclusions, as well as scenarios involving Endosign, endoscopy costs, sponge detachment rate, a male only population and utility assumptions around late-stage OAC. Scenarios exploring a shorter time horizon, no endoscopic surveillance, lower prevalence from Kadri et al. (2010), and diagnostic accuracy using a segment length of 2 cm or more produced stronger cost effective ICERs than the base case.

Two explored scenarios changed the cost effectiveness conclusions of the base case analysis. One was a scenario where age and sex inputs were adjusted to match the study used to inform Barrett's oesophagus prevalence in the base case (Saha et al. 2024). This led to a population

younger than the base case, resulting in a lower ICER which was not cost effective. The other scenario was when age-adjusted utilities were not included. This increased QALY losses and produced an ICER which was not cost effective.

A scenario considering insourcing costs for endoscopic biopsy was not included due to the lack of a robust cost estimate. However, if insourced procedures are expected to be more costly than in-house procedures, this would increase the cost of the comparator. As a result, cost effectiveness conclusions would remain unchanged from the base case analysis.

Full details of the methods and results are available in Appendix 6.

7. Organisational considerations

Reporting from six real-world cohort studies and by SHTG highlighted some of the impacts adoption of capsule sponge testing had on organisations.

Gourgiotis et al. (2025) found that using capsule sponge testing to triage patients referred for investigation of chronic reflux symptoms in England resulted in 78% of patients being removed from endoscopy waiting lists. They also found that patients with negative capsule sponge results could be discharged sooner, as endoscopy was not required as often, and that these timelines were similar to the counterfactual group that all received endoscopies. The pathway was longer for patients with positive capsule sponge results due to the need for follow-up investigation. Another cohort study in England found that 62% of patients avoided having endoscopy due to the capsule sponge triage pathway and 82% of patients were discharged (Angel et al. 2025). In Scotland, only 27.2% of patients under investigation for reflux symptoms underwent endoscopy after capsule sponge testing (Chien et al. 2024a). 70% of patients were discharged from secondary care after capsule sponge testing and did not require further investigation. However, not all patients with negative capsule sponge tests were discharged, with some referred for endoscopy based on the clinicians' judgment, and 10 patients with negative results were found to have significant pathology on endoscopy. The authors noted that there were four cancer diagnoses in the cohort, three of which were gastric cancers, and highlighted that capsule sponge testing is primarily for oesophageal conditions and should not be used in isolation. The authors suggested that 'all patients undergoing capsule sponge testing for reflux symptoms should undergo consultation with an [upper gastrointestinal] specialist nurse or medical practitioner within secondary care to assess the need for additional investigation before discharge is instigated'. Angel et al. (2025) also made a similar recommendation to safety net patients.

For patients under surveillance for Barrett's oesophagus in Scotland, Chien et al. (2024b) found that 18.6% were discharged from surveillance based on capsule sponge testing results and other clinical details. People within the ultra-low risk group with two tests negative for intestinal metaplasia and those aged over 80 years were discharged as well as 75 individuals in the moderate risk group, due to advanced age or comorbidities. Before the introduction of capsule sponge testing, all Barrett's oesophagus patients would have received surveillance endoscopies, whereas only 16.2% of the cohort in this retrospective study required endoscopy within 12 months of their capsule sponge test. Similarly, in a later cohort study in Scotland, only 17.1% of patients in the capsule sponge cohort required endoscopy (Chien & Glen 2025). SHTG (2023) reported that capsule sponge testing led to reductions in delays to Barrett's oesophagus surveillance in Scotland. The median length of delay was nine months in the first year of capsule sponge testing, which reduced to five months in the second year ($p < 0.001$). The proportion of patients experiencing delays of more than three months also significantly reduced in this time period, going from 72.5% to 57.0% ($p < 0.001$). By risk stratifying patients under surveillance for Barrett's oesophagus using clinical risk factors and capsule sponge results, those at highest risk can be prioritised for endoscopic investigation and a cohort study in England (Tan et al. 2025) found that the median time from capsule sponge to endoscopy for patients stratified as high risk was 1.5 months (interquartile range [IQR] 1.1 to 2.7) compared with 13.1 months (IQR 7.1 to 23.6) for those in the low-risk group.

Overall, real-world evidence indicates that the use of capsule sponge triage testing may lead to greatly reduced demand on endoscopy services. However, it is important to note that these evaluations took place during the Covid-19 pandemic when endoscopy services were greatly disrupted, and this will have had an effect on some of the findings.

The Small Business Research Initiative (SBRI) Centre of Excellence funded the Celtic Capsule Project, which was a pilot study of introducing EndoSign testing for patients on endoscopy

waiting lists for chronic reflux symptoms and Barrett's oesophagus surveillance across three centres in Wales and two in Northern Ireland (Cyted Health 2025). The project ran from November 2024 to March 2025 and 196 capsule sponge tests (101 in Wales, 95 in Northern Ireland) were successfully carried out. Within Wales, 66 Barrett's oesophagus surveillance patients and 35 chronic reflux patients were tested. At the time the Celtic Capsule Project evaluation report was written, 188 tests had been processed (113 Barrett's oesophagus surveillance, 75 chronic reflux) and were included in analyses. For Barrett's surveillance, 76% of patients avoided having endoscopy based on their capsule sponge test results and 56% of patients being investigated for chronic reflux symptoms did not require endoscopy. This shows a notably reduced demand on endoscopy services and may have allowed those in need of urgent endoscopy to undergo this investigation sooner than they would have otherwise. The report estimates that 29.25 hours of endoscopists' time was released due to the number of endoscopies avoided based on capsule sponge test results, as well as 39 hours of theatre time. CVUHB reported a reduction in their number of overdue endoscopies from 3,108 to 2,924 during the period of the project, and the wait time for an endoscopy went down from 49 weeks to 42 weeks. However, CVUHB also insourced resources during this time so it is not possible to determine how much of these reductions was due to capsule sponge testing.

Comments from experts showed agreement with the findings of these real-world studies, feeling that triage and risk stratification with capsule sponge testing could reduce the demand on endoscopy services. Experts also agreed that 'safety netting' of patients would be needed to reduce the risk of patients being discharged inappropriately, as well as very clear inclusion and exclusion criteria for capsule sponge testing. Much of this work has already been done by SBRI's Celtic Capsule project, including developing patient pathways.

Multiple experts stated capsule sponge testing could, or should, be performed in primary care settings. This could lead to improved access to this service, though it was also raised that links with secondary care would be required to handle any adverse events, such as sponge detachments. Experts also said that capsule sponge testing services could be nurse-led with senior clinical oversight.

Cyted have stated that they would provide training in the use of the EndoSign device and would provide all pathology services for capsule sponge devices. This, therefore, would not add pressure to NHS pathology services and could reduce pressure on services if fewer endoscopic biopsies are also performed in response to capsule sponge results.

There are also equity of access considerations. Capsule sponge testing is currently used in Scotland and England, and in some areas of Wales as discussed earlier in this report (Section 3). Only one health board in Wales (BCUHB) has implemented capsule sponge testing, as part of its Barrett's oesophagus surveillance service. This is performed at only one hospital within the health board and so there may still be equity of access issues on a local level. Provision of capsule sponge testing in primary care settings could help address this inequity, particularly in rural areas of Wales.

8. Patient, carer and family considerations

HTW collaborated with patient organisations to gather perspectives and experiences on capsule sponges for detecting Barrett's oesophagus. Reponses were received from Barrett's Patient Support and from Heartburn Cancer UK. Barrett's Patient Support sent published literature that they were involved in producing and agreed to host an online survey. Heartburn Cancer UK also sent patient experiences from their Demanding Hope Campaign and agreed to circulate the same survey.

In addition, the 'patient and social aspects' section of STHG's review of capsule sponge devices (2023) is summarised here.

8.1 Responses from Barrett's Patient Support

8.1.1 Qualitative papers

Barrett's Patient Support sent three papers for consideration. Only one paper was relevant and insights from this are summarised below.

8.1.1.1 Living with Barrett's oesophagus

In "Learning to Live with Barrett's Oesophagus", Davies (2024) explores people's responses to getting a diagnosis of Barrett's oesophagus. Despite reassurances from healthcare professionals, speculations of cancerous futures can be 'terrifying' and, at times, 'all-consuming' for some. These patients describe 'feeling like a ticking time bomb' and becoming 'obsessed' by the possibility of developing cancer. These patients may struggle to sleep as their thoughts become fixed on worrying about the future. This can place a significant strain on various aspects of a person's life, as they struggle to be 'present' at work and at home, leading to breakdowns in family relationships and the workplace. It can also 'break people's trust in their knowledge of their health', which can lead to patients seeking more frequent surveillance out of a sense of loss of control. Patients may seek help from their GPs and may have their fears dismissed or be treated for anxiety.

Trust in surveillance technologies can be key for patients to reestablish hope for the future. Having the support of an online community can also be key to how patients manage their diagnosis and receive and share information. Patients often share the difficulties and challenges they face amongst themselves only, due to feeling that they need to 'stay strong' for family and friends and 'not get upset' in front of them. Family members can also experience significant distress on a loved one's diagnosis and the potential for future cancer, with some describing themselves as 'devastated' and 'worried sick'.

However, not all patients respond this way. Some patients report being less concerned about potential cancerous futures and not experiencing the same sense of loss of control.

8.2 Responses from Heartburn Cancer UK

Heartburn Cancer UK shared patient feedback on the capsule sponge test from their HCUK Demanding Hope February 2024 OC Awareness Month campaign.

Patients reported that they were satisfied with the information they were given prior to having the capsule sponge test. 63% of patients who took part had previously had an endoscopy as part

of their Barrett's oesophagus monitoring/diagnosis. 81% of these patients rated the capsule sponge test as 'five stars' for 'overall experience' and 70% advised that they would have the test again and recommend it to family and friends. Some patients found swallowing the capsule to be 'difficult' and 'unpleasant' and one patient reported being unable to swallow the capsule and having to abandon the test. When invited to provide comments, patient feedback centred on the attitude and helpfulness of the clinicians administering the test (including reassurance, empathy, friendliness and kindness). One patient commented on the overall process from capsule sponge to endoscopy diagnosis:

> *"as a results [of the capsule test] I have been diagnosed with Barrett's and underwent endoscopy. The whole process has been seamless and to have the endoscopy conducted so quickly after the sponge test was incredible"*
>
> Patient quote from HCUK Demanding Hope campaign

More detailed insights were not provided.

8.3 Survey Results

8.3.1 Survey respondent demographics

The online survey ran from5 May to 6 June 2025 and 58 responses were received. Respondents were a mixture of people with suspected or confirmed Barrett's oesophagus (21 people), chronic acid reflux (CAR) (18 people), both Barrett's oesophagus and CAR (6 people), GORD (1 person) and suspected oesophageal cancer (1 person). Others preferred not to disclose this information.

Thirty-three respondents were invited to take the capsule sponge test as part of investigating their CAR or suspected Barrett's oesophagus, two were taking the capsule sponge test as part of monitoring their already diagnosed Barrett's oesophagus, and the remaining 23 respondents had endoscopy but no experience of the capsule sponge.

Symptoms reported by respondents included acid reflux (particularly at night), heartburn, chest pain, cough, stomach pain, indigestion, sore throat, weight loss, choking and pain under the ribs.

8.3.2 Pre-test experiences

Most of the respondents who had the capsule sponge test advised that they had no concerns before taking the test. For those who did have concerns, these included:

- not be able to swallow the capsule,
- choking,
- string breaking,
- the capsule getting stuck,
- not being able to retrieve the capsule and needing it retrieved by surgery,
- anxiety and feeling scared.

One respondent discussed their pre-test anxiety regarding the results of the test itself.

8.3.3 Capsule sponge device experiences

Respondents' ability to swallow the capsule was varied. Most reported that they had 'no issues' swallowing the capsule and that it was 'easy'. Some reported that it was more difficult, but with help (such as having 'a lot' of water or using warm water) they were able to swallow it on the first attempt. A few respondents were able to swallow the capsule after two or more attempts. One was unable to swallow the capsule despite several attempts.

> *"Very easy swallowed first time"*
>
> *"Yes, swallowed at first try, but it took a lot of drinking water to get it down. That was OK."*
>
> *"I swallowed it first time but it was very big and difficult to do so"*
>
> *"It took me two or three attempts to swallow it down. It felt a bit awkward."*
>
> *"Impossible to swallow it after many attempts, but I think I might have been able to swallow it if I had been allowed to do it my own way"*
>
> Patient quotes from survey

Twenty-four respondents reported no adverse effects when swallowing the capsule. Some reported mild discomfort, while others reported gagging. Gagging was associated with the presence of the string, rather than the capsule itself. Some respondents advised they experienced gagging for 'a few seconds', 'a little' and 'slight' gagging. One respondent struggled with the gagging they experienced for the duration of the test.

> *"Very slight feeling of discomfort as the capsule was swallowed."*
>
> *"Gagging for a few seconds"*
>
> *"No issues when swallowing but gagging sensation from the string. I had to constantly try to distract myself from the string while waiting for the time to elapse. Talking to the health care professionals really helped me stay calm."*
>
> Patient quotes from survey

There were no reported instances of more serious adverse effects such as choking, vomiting or difficulty breathing.

Most of the respondents reported no issues in the retrieval of the sponge, advising that it was 'easy' if 'a little strange'. For others, difficulties retrieving the sponge ranged from some discomfort and unpleasantness to choking, coughing, gagging and one instance of 'bringing back up the water to swallow it down'. One responded described this part of the test as 'violent' as they 'weren't able to breathe'. Of these, gagging was the highest reported experience.

> *"Very easy, over in seconds"*
>
> *"Very easy, a little strange but no problems"*
>
> *"Fine but not very pleasant"*

"It was ok until it reached the gag reflex, coughed it up along with the water I drank to take it down"

"This was the worst part as it felt quite violent and for a moment I was worried because I couldn't breathe or talk. But it was quickly over."

Patient quotes from survey

Some of the respondents advised having a 'sore throat' following the retrieval of the sponge, feeling sick and some discomfort.

"I really thought I would be sick but followed advice and had an empty stomach. Slight sore throat for an hour"

"Just cough for a few seconds after"

"A slight soreness in the throat which lasted a couple of days and then went away. Otherwise OK."

"Gagging, pain as the Brillo pad type sponge scratched my throat as it was being pulled out, feeling of choking and being unable to breathe."

Patient quotes from survey

Eighteen of the respondents who had the capsule sponge test advised that they were progressed on to endoscopy. One advised that this was because the results of the capsule sponge test were indeterminate. Time from the capsule sponge test to endoscopy varied. Some respondents advised they had an endoscopy 'immediately' or within '2 to 3 weeks' to '5 or 6 weeks' later. Endoscopy took four to five hours at a hospital setting. Sedation was varied. Respondents needed the help of family members, took time off work and had to report to a hospital.

"Was told that there were insufficient cells to give a result."

"Yes, I did, and not very long to wait, 2-3 weeks."

"Yes. Can't remember exactly but there was a gap of around 4-5 weeks between the result of the sponge test and the endoscopy"

"Had an endoscopy approx 6 weeks after sponge test"

Patient quotes from survey

For those respondents who were not referred to endoscopy, they advised that the reasons why were effectively explained to them, and they were happy with the decision.

"Very pleased with explanation"

Patient quote from survey

8.3.4 Capsule sponge test vs endoscopy

Respondents who had experience of both the capsule sponge test and endoscopy were asked to compare both procedures. Their responses predominantly agreed that the capsule sponge test is 'easier' (both to prepare for and to undertake) and 'more efficient', but that endoscopy is seen as 'more accurate' and 'reassuring'.

> *"The sponge is much the easier option. It is obviously quicker and is not as uncomfortable as an endoscopy. The endoscopy is the most reassuring."*
>
> *"I would say sponge test is much better and less intrusive, also much quicker"*
>
> *"If it had picked up more cells the capsule sponge would be preferable"*
>
> *"The sponge was the easiest I had no problem with it. but better results with the endoscopy that took about two hours but as I was under anaesthetic every time"*
>
> *"Sponge test much more efficient and a lot less discomforting. However and endoscopy is much more detailed with the results ie for my case, they found small acid burns on my oesophagus using endoscopy"*
>
> Patient quotes from survey

Those respondents who were satisfied with the level of accuracy of the capsule sponge test did not consider endoscopy superior. Equally, those who had negative experiences with endoscopy considered the capsule sponge to be superior.

> "*I've had many endoscopy's over the years. There is no comparison this [CST] is quick and virtually pain free. I hated endoscopy's. I tried endoscopy with and without sedation both were awful for me/"*
>
> *"No comparison. Capsule test only took 10 minutes to complete and I was reassured of the accuracy of the negative test result by the same result of an endoscopy later in the year."*
>
> *"Way prefer the sponge, 10mins done! Less prep, picks up more i think compared to endoscopy."*
>
> *"Endoscopy was vile. I have anxiety about the second one I will need"*
>
> *"The Capsule sponge test is much easier to have done than the endoscopy. I wouldn't have an endoscopy unless the doctors thought that it was really necessary."*
>
> *"After the endoscopy and burning off of abnormalities it does leave you uncomfortable for a few days and a little anxious until you get results of tests with photo's of what has been done."*
>
> Patient quotes from survey

Similarly, those respondents who did not consider the capsule sponge test to be as accurate as the endoscopy showed a preference for endoscopy despite it being more challenging. Few respondents advised that the endoscopy was actually less challenging than the capsule sponge test.

"It [CST] was quicker but much more traumatic. Similar preparation time. The endoscopy was easier for me, more reassuring and less traumatic!"

"Endoscopy as goes further down"

"Endoscopy was easier, more comfortable for me"

"The endoscopy was very easy as I was completely sedated."

Patient quotes from survey

Preferences also took into consideration surrounding circumstances. The ability to have endoscopy under sedation was identified by several respondents as the key difference between how 'uncomfortable' it is and whether they would prefer the capsule sponge test. Length of procedure, needing help from family, and costs were also factors patients considered.

"Endoscopy feels much more invasive than swallowing the sponge, is much more uncomfortable if you don't have sedation and take's considerably longer. Requires a visit to hospital, more waiting around, someone to accompany you if you have sedation (both to ensure you get home safely and to take note of anything your clinician reported as you won't remember the conversation). Capsule sponge is much easier."

"Endoscopies cost a lot of money so the capsule sponge is an excellent, quick and easy alternative. I would have faith in either procedure."

Patient quotes from survey

Respondents were also asked to state which procedure they would choose to have, if both were available for them. Responses were mixed. Twelve respondents clearly stated they preferred endoscopy, 23 respondents clearly stated they preferred the capsule sponge test. The rest gave more nuanced responses that considered various factors, and one respondent advised that they would leave this decision to their healthcare advisor.

"Would rather have an endoscopy due to feel it's the gold standard of surveillance. Less traumatic"

"Capsule sponge … quick and easy".

"Would be happy with long term monitoring by way of sponge test But would still want an endoscopy after a few sponge tests for a more detailed result"

"I can not swallow tablets so I don't think I would be able to do this [CPT] due to fear of choking"

"Would prefer the sponge test as it's less intrusive. But an endoscopy does shows up any visible issues"

"Not the capsule sponge - the entire process was unpleasant and scary. I was pleased it was done at the hospital as would have been worse elsewhere because of how scary it was, I was very worried something might go wrong or I'd choke."

"I should prefer monitoring by the sponge test. The sponge test was much easier to have done than the arrangements for an endoscopy. Also I felt that the sponge test was less invasive than an endoscopy and recovery was much easier."

Patient quotes from survey

When asked to consider what the benefits were of having a choice between endoscopy or capsule sponge testing, respondents considered that the capsule sponge test would encourage people to get tested as it is quicker and less off-putting than endoscopy, which requires much more preparation. It was also considered a good way to get early indications of the presence of disease, particularly as it could be conducted at a patient's GP surgery, which could lead to faster investigations by endoscopy if necessary. Potential cost savings were also identified as a benefit, if the capsule sponge test could be used in early detection.

"I think it would make all the difference to some people, it is an easier procedure to prepare for, it is much easier to sit in chair and swallow the sponge than having to go through an endoscopy."

"Easy, quick, no sedation, could get on with the rest of my day. Would put less people off getting checked out."

"Early detection is key my husband died 5 weeks after diagnosis"

"If they have not been diagnosed, the capsule sponge should always take preference, especially if it can be done in their local surgery."

"The sponge would I think give a first indication of any issues. It's quick and would save the NHS a lot of money,"

Patient quotes from survey

Some respondents felt it would also benefit those who require ongoing monitoring, as it can be delivered quicker in local settings. However, respondents acknowledged that it 'may not be for everyone'

"A regular monitor basis for people with other conditions like myself with GERD to make sure I don't suffer BO oesophageal cancer in the future"

"It is a much less stressful test, must be less costly than endoscopy and gives reliable results. Possibly could mean you may be tested more frequently"

"I think for some people it gives another option . But for myself I don't think I could do it"

"I assume that the capsule sponge test would involve less resource to administer and would enable more people to be tested more often to detect and treat any problem at an earlier stage."

Patient quotes from survey

Final comments were predominantly positive for the capsule sponge test.

> *"I can't see any reason why the capsule sponge test shouldn't be more widely available"*
>
> *"I feel very grateful to have had the sponge test and I would recommend anyone with long-standing acid reflux to have it done if that is possible."*
>
> *"Since treatment for OC I have had regular surveillance endoscopies and also some capsule sponge tests. I would recommend the latter for their speed and accuracy and the former because you can immediately discuss any obvious inflammation/ lesions with the endoscopist. The capsule sponge has the advantage of collecting cells over a wide surface area compared with the small biopsy samples collected during an endoscopy."*
>
> *"The sponge was like a Brillo pad, incredibly sharp and large as it's being pulled out. Having to sit for several minutes with the string hanging out of my mouth was not pleasant. I gagged a lot and felt very scared waiting to have it pulled out. I wouldn't have one again as the experience was quite traumatising!"*
>
> *"If anybody is suffering from heartburn or indigestion this procedure could ultimately save your life, be in no doubt"*
>
> Patient quotes from survey

8.4 SHTG report: patient and social aspects

The SHTG (2023) review's 'patient and social aspects' section comprised of patient experiences with capsule sponge devices and public perception of capsule sponge devices reported in the included studies.

8.4.1 Patient experiences with capsule sponge devices

SHTG identified studies exploring the experiences of patients with chronic reflux, who had a capsule sponge test, were identified by SHTG. Overall results showed that patients were satisfied with their experience of the Cytosponge test and 80% would be willing to have the test again. Patients preferred having the test in a primary care setting. The lowest rated part of the procedure in terms of patient satisfaction was for the retrieval stage of the sponge from the oesophagus. Patients described feeling anxious about being able to complete the test or about the test itself. For some, their anxiety resolved after they received their test result. Some patients reporting having difficulty in swallowing the sponge capsule (gagging, retching or heaving) because the string was uncomfortable, or because it was difficult to drink water to swallow the capsule with the string attachment.

Patients who had high or very high anxiety levels were more likely to have a poor experience compared with participants with normal anxiety. The odds of having a poor experience were also greater for individuals who drank alcohol on most days compared with individuals who never drank alcohol, for those who struggled to swallow the capsule on the first attempt, and for men compared to women.

8.4.2 Public perception of capsule sponge devices

SHTG identified a qualitative analysis exploring the acceptability of the Cytosponge test in a sample of people from the UK who were living with GORD. Concerns from the participants included worries about swallowing and extracting the sponge, such as the possibility of swallowing the string, the string getting stuck, gagging/vomiting while trying to swallow the capsule, the string detaching, and discomfort.

Participants with previous experience of endoscopy felt that the capsule sponge device would be preferable to endoscopy physically, practically and economically. Participants were enthusiastic about having the test at their local general practice, not needing an anaesthetic and being able to return to everyday activities immediately.

8.5 Equality, diversity and equity considerations

No information on inequalities/inequities or considerations for patients with protected characteristics were identified during the evidence review.

9. Conclusions

This evidence review summarised published evidence on the effectiveness and cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer.

The literature search identified four clinical guidelines and 17 studies: one HTA, one RCT, 12 observational studies, and four economic studies (one of which was one of the 12 observational studies).

The evidence included in this review suggests the diagnostic accuracy of capsule sponge devices with TFF3 testing for proactive screening of Barrett's oesophagus in those with chronic reflux is good, with high sensitivity and specificity. Where reported, PPV is low whilst NPV is high. This is important as it suggests the likelihood of capsule sponge testing missing Barrett's oesophagus is low. For case finding of Barrett's oesophagus using capsule sponge testing with TFF3, p53, and cellular atypia, detection rates suggest potentially high rates of false positives but, importantly, very low rates of false negatives as well. The diagnostic accuracy for case finding also appears to be good, with sensitivity above 90%, and PPV and NPV findings supporting the findings from detection rates. Capsule sponge testing with p53 and cellular atypia for Barrett's oesophagus under surveillance also shows good accuracy for detecting dysplasia or cancer, however, the two biomarkers in isolation may not be sufficiently accurate. Again, the data suggest the number of false positives is quite high in this indication, but false negatives are very low. The evidence also suggests that using capsule sponge testing, in combination with assessing clinical risk factors, is effective in risk stratifying Barrett's oesophagus patients.

There was no reporting of longer-term outcomes, such as mortality and survival, however this is currently under investigation in the BEST4 trial. Time to diagnosis and time to treatment were reported in one evaluation of real-world data, with no comparisons to standard care. No data on health-related QoL were identified.

The safety of capsule sponge devices appears to be good, and the incidence of adverse events is low. Though a field safety notice was issued for several batches of Cytosponge in 2023 due to higher risk of sponge detachment, the rates of sponge detachment reported in the literature are very low. The rates of other adverse events were also very low.

Based on the evidence available, it is not possible to say whether there is any difference in outcomes depending on whether capsule sponge testing takes place in primary or secondary care. However, the evidence suggests it is a viable option in either setting.

Most of the evidence was related to the device Cytosponge, however, evidence is generalisable across Cytosponge and EndoSign but not to other non-endoscopic cell collection devices. The majority of studies involved people who were involved in the development of the examined devices, or were employees or founders of the companies that manufacture them. There is therefore the possibility for some level of bias in these studies, however all interests were appropriately declared. More research is needed on the effect capsule sponge testing has on cancer outcomes and a long-term trial is underway to collect data on mortality, however this trial is not due to end until 2035. Evidence comparing outcomes and patient experiences of capsule sponge testing in primary and secondary care settings would also be beneficial.

Four studies were included in the economic review. Only one study took the perspective of the UK NHS and concluded that endoscopy-only screening was not cost effective compared to using Cytosponge. Their base case analysis estimated cost savings of £422 with a reduction of 0.0041 QALYs per patient triaged using Cytosponge compared with endoscopy alone. However, potentially serious limitations of this study were identified including possible biases in the data

used to inform the diagnostic pathway and comparator arm, as well as uncertainties in how representative the clinical data is to the modelled population.

Therefore, HTW researchers developed a cost-utility analysis from the NHS Wales perspective to estimate the cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage OAC in people with chronic reflux, compared to endoscopic biopsy. Over a lifetime horizon, results estimated that Cytosponge use in primary care, followed by endoscopic biopsy in those with a positive result, is expected to reduce costs by [redacted] per patient with a loss of 0.02 QALYs, corresponding to an ICER of [redacted] representing the cost savings per QALY lost. This is above the £20,000 cost-effectiveness threshold, indicating that the use of Cytosponge is cost effective in the context where the intervention is less costly and less effective than the comparator. Probabilistic sensitivity analysis suggested a 65.8% probability of cost effectiveness at this threshold. Capsule sponge sensitivity, age and Barrett's oesophagus prevalence were identified as influential drivers of cost effectiveness. Scenarios exploring capsule sponge delivery in secondary and community-based care settings, as well as the use of the Endosign device, had minimal impact on health economic outcomes, with no change in cost effectiveness conclusions. However, conclusions did change in scenarios exploring younger populations and where age-related utility decline is not considered.

Real-world evidence and feedback from subject experts indicated introducing capsule sponge testing could significantly reduce demand on endoscopy services, which are currently under pressure. This testing could also ensure those most in need have quicker access to endoscopic investigation. However, safety netting and clear patient pathways with defined eligibility criteria would also be needed to ensure patients do not receive unnecessary investigations or inappropriate discharges as serious pathology has been identified in patients with negative capsule sponge results. Introduction of capsule sponge testing could also address equity of access issues both within Wales and across the UK.

Overall, the evidence suggests that false positives may be quite high with capsule sponge testing, but triaging based on this testing can still significantly reduce the number of endoscopies needed to be performed. The number of false negatives also appears to be very low, meaning that the risk of missing pathology is very low. However, the lack of endoscopic biopsy results on the majority of patients that were negative on capsule sponge testing means the number of true/false negative results is not known and this is a limitation of the evidence.

10. Contributors

This topic was proposed by Jeff Turner, Clinical Lead, National Cancer Recovery Programme.

The HTW staff involved in producing this report were:

- G Hammond, Health Services Researcher - Effectiveness author
- N Bromham, Senior Health Services Researcher – Effectiveness quality assurance
- R Miller, Senior Health Economist – Cost effectiveness author
- M Prettyjohns, Principal Researcher – Cost effectiveness quality assurance
- J Washington, Information Specialist – Literature search & information management
- A Evans, Patient and Public Involvement (PPI) Manager – PPI author
- R Shepherd, Project Manager – Project Management

The HTW Assessment Group advised on methodology throughout the scoping and development of the report.

We are grateful to the following subject experts, who also contributed to this appraisal:

- A Christian, Consultant Pathologist, Cardiff and Vale University Health Board, National Pathology Programme
- S Dolwani, Professor of Gastroenterology, Screening, Prevention and Early Detection of Cancer, School of Medicine, Cardiff University
- R Fitzgerald, Professor of Cancer Prevention and Director of the Early Cancer Institute, University of Cambridge. Honorary Consultant Gastroenterology
- P Glen, Consultant Upper GI Surgeon, NHS Scotland
- H Haboubi, Consultant Gastroenterologist and Upper GI Medicine Lead, Cardiff & Vale University Health Board; Barrett's Endotherapy Lead for South Wales; Upper GI Clinical Implementation Lead, Cancer Recovery Programme
- N Hopwood, Upper GI / General Surgical Advanced Nurse Practitioner / Lead Nurse for Capsule Sponge Test Implementation Pilot, Betsi Cadwaladr University Health Board
- B Norton, Doctoral Research Fellow & gastroenterology registrar, Cleveland Clinic London, University College London, University College London Hospitals
- M O'Donovan, Consultant Histopathologist, Cambridge University Hospitals NHS Trust; Associate Affiliated professor, Cambridge University Hospital; Lead Pathologist for Capsule Sponge, Cyted Health Ltd
- D Spackman, Vice President and UK General Manager, Cyted Health Ltd
- J Turner, Consultant Gastroenterologist; Clinical Lead, National Cancer Recovery Programme

Subject experts contributed to this appraisal by commenting on a draft of this report, and in some cases providing other advice to HTW's staff and decision-making groups. All contributions from reviewers were considered by HTW's Assessment Group and actioned accordingly. However, subject experts had no role in authorship or editorial control, and the views expressed are those of Health Technology Wales.

References

Alexandropoulou K, van Vlymen J, Reid F, et al. (2013). Temporal trends of Barrett's oesophagus and gastro-oesophageal reflux and related oesophageal cancer over a 10-year period in England and Wales and associated proton pump inhibitor and H2RA prescriptions: a GPRD study. European Journal of Gastroenterology and Hepatology. 25(1): 15-21. doi: https://doi.org/10.1097/MEG.0b013e3283595086

Angel D, Kumar R, Shaw K, et al. (2025). Evaluation of a novel capsule sponge triage pathway for patients routinely referred with reflux symptoms: safety, long-term outcomes and impact on endoscopy from a large volume single site cohort over 4 years. Frontline Gastroenterology. flgastro-2025-103154. doi: https://doi.org/10.1136/flgastro-2025-103154

Aoki T, Watson DI, Bulamu NB. (2024). Cost-effective identification of Barrett's esophagus in the community: a first step towards screening. Journal of Gastroenterology and Hepatology. 39(12): 2654-63. doi: https://doi.org/10.1111/jgh.16762

Benaglia T, Sharples LD, Fitzgerald RC, et al. (2013). Health benefits and cost effectiveness of endoscopic and nonendoscopic Cytosponge screening for Barrett's esophagus. Gastroenterology. 144(1): 62-73.e6. doi: https://doi.org/10.1053/j.gastro.2012.09.060

Cancer Research UK. (2024). Early cancer diagnosis data hub: incidence by stage (gold standard). Available at: https://crukcancerintelligence.shinyapps.io/EarlyDiagnosis/ [Accessed 16 Apr 2025].

Centers for Medicare and Medicaid Services. (2017). 2017 NPRM OPPS cost statistics files. CMS.gov. Available at: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HospitalOutpatientPPS/Hospital-Outpatient-Regulations-and-Notices-Items/CMS-1656-P.html.

Chien S, Glen P. (2025). Impact on Barrett's dysplasia yield following the introduction of capsule sponge testing versus traditional endoscopic surveillance. Diseases of the Esophagus. 38(3): doaf033. doi: https://doi.org/10.1093/dote/doaf033

Chien S, Glen P, Bryce G, et al. (2024a). Clinical application of capsule sponge testing in symptomatic reflux disease: a national prospective cohort study. BMC Gastroenterology. 24(1): 431. doi: https://doi.org/10.1186/s12876-024-03503-5

Chien S, Glen P, Penman I, et al. (2024b). Oesophageal cell collection device and biomarker testing to identify high-risk Barrett's patients requiring endoscopic investigation. British Journal of Surgery. 111(5): znae117. doi: https://doi.org/10.1093/bjs/znae117

Cook Medical. (2018). GI endoscopy coding and reimbursement guide.

Cyted Health. (2024). EndoSign: frequently asked questions. Cyted Ltd. Available at: https://www.endosign.com/hcp/faqs [Accessed 10 Mar 2025].

Cyted Health. (2025). Celtic Capsule Project: evaluation report [Unpublished report]. Cyted Ltd.

Davies E. (2024). 'This has given me the peace of mind I needed': ethnographic insights into Barrett's oesophagus screening using the capsule sponge test. Journal of Medical Ethics. 51(6): 391-5. doi: https://doi.org/10.1136/jme-2024-109928

Eluri S, Paterson A, Lauren BN, et al. (2022). Utility and cost-effectiveness of a nonendoscopic approach to Barrett's esophagus surveillance after endoscopic therapy. Clinical Gastroenterology and Hepatology. 20(2): e51-e63. doi: https://doi.org/10.1016/j.cgh.2021.02.013

ESGE. (2023). Diagnosis and management of Barrett esophagus: European Society of Gastrointestinal Endoscopy (ESGE) guideline. Available at: https://www.esge.com/diagnosis-and-management-of-barrett-esophagus [Accessed 28 Jan 2025].

Eusebi LH, Cirota GG, Zagari RM, et al. (2021). Global prevalence of Barrett's oesophagus and oesophageal cancer in individuals with gastro-oesophageal reflux: a systematic review and meta-analysis. Gut. 70(3): 456-63. doi: https://doi.org/10.1136/gutjnl-2020-321365

Falk GW, Chittajallu R, Goldblum JR, et al. (1997). Surveillance of patients with Barrett's esophagus for dysplasia and cancer with balloon cytology. Gastroenterology. 112(6): 1787-97. doi: https://doi.org/10.1053/gast.1997.v112.pm9178668

Fitzgerald RC, di Pietro M, O'Donovan M, et al. (2020). Cytosponge-trefoil factor 3 versus usual care to identify Barrett's oesophagus in a primary care setting: a multicentre, pragmatic, randomised controlled trial. Lancet. 396(10247): 333-44. doi: https://doi.org/10.1016/S0140-6736(20)31099-0

Garside R, Pitt M, Somerville M, et al. (2006). Surveillance of Barrett's oesophagus: exploring the uncertainty through systematic review, expert workshop and economic modelling. Health Technology Assessment. 10(8). doi: https://doi.org/10.3310/hta10080

Gourgiotis V, Graham C, Foerster K, et al. (2025). Use of a non-endoscopic capsule-sponge triage test for reflux symptoms: results from the NHS England prospective real-world evaluation. Alimentary Pharmacology and Therapeutics. 10(5): 876-85. doi: https://doi.org/10.1111/apt.18472

Heberle CR, Omidvari AH, Ali A, et al. (2017). Cost effectiveness of screening patients with gastroesophageal reflux disease for Barrett's esophagus with a minimally invasive cell sampling device. Clinical Gastroenterology and Hepatology. 15(9): 1397-404.e7. doi: https://doi.org/10.1016/j.cgh.2017.02.017

Hernández Alava M, Pudney S, Wailoo A. (2022). Estimating EQ-5D by age and sex for the UK. NICE DSU report. NICE Decision Support Unit, University of Sheffield. Available at: https://www.sheffield.ac.uk/nice-dsu/methods-development/estimating-eq-5d [Accessed 25 Oct 2023].

Hur C, Choi SE, Rubenstein JH, et al. (2012). The cost effectiveness of radiofrequency ablation for Barrett's esophagus. Gastroenterology. 143(3): 567-75. doi: https://doi.org/10.1053/j.gastro.2012.05.010

Inadomi JM, Sampliner R, Lagergren J, et al. (2003). Screening and surveillance for Barrett esophagus in high-risk groups: a cost-utility analysis. Annals of Internal Medicine. 138(3): 176-86. doi: https://doi.org/10.7326/0003-4819-138-3-200302040-00009

Inadomi JM, Somsouk M, Madanick RD, et al. (2009). A cost-utility analysis of ablative therapy for Barrett's esophagus. Gastroenterology. 136(7): 2101-14.e1-6. doi: https://doi.org/10.1053/j.gastro.2009.02.062

IQVIA. (2023). NHS England cytosponge evaluation for routine referrals: final report. Available at: https://www.canceralliance.co.uk/application/files/7617/1223/6984/Cytosponge_Evaluation_Final_Report_4.pdf [Accessed 27 Jan 2025].

Iyer PG, Lansing R, Johnson ML, et al. (2018). Accurate non-endoscopic detection of Barrett's esophagus in a multicenter prospective validation cohort: the SOS 2 trial. Gastroenterology. 154(6, Supplement 1): S175-6 [Conference abstract: 878]. doi: https://doi.org/10.1016/S0016-5085(18)30998-3

Januszewicz W, Tan WK, Lehovsky K, et al. (2019). Safety and acceptability of esophageal cytosponge cell collection device in a pooled analysis of data from individual patients. Clinical Gastroenterology and Hepatology. 17(4): 647-56.e1. doi: https://doi.org/10.1016/j.cgh.2018.07.043

Jones KC, Weatherly H, Birch S, et al. (2025). Unit costs of health and social care 2024 manual. Technical report. Personal Social Services Research Unit (University of Kent) & Centre for Health Economics (University of York). Available at: https://kar.kent.ac.uk/109563/ [Accessed 22 April 2025].

Kadri SR, Lao-Sirieix P, O'Donovan M, et al. (2010). Acceptability and accuracy of a non-endoscopic screening test for Barrett's oesophagus in primary care: cohort study. BMJ. 341: c4372. doi: https://doi.org/10.1136/bmj.c4372

Matchett CL, Sahinkoc M, Chhatwal J, et al. (2025). Cost effectiveness analysis of Barrett's esophagus screening: a systematic review. Best Practice & Research: Clinical Gastroenterology. 75: 101980. doi: https://doi.org/10.1016/j.bpg.2025.101980

MHRA. (2023). Medtronic: Cytosponge Cell Collection Device. Field Safety Notices (FSNs) from 5 to 9 June 2023 MHRA reference: 5296419. Medicines and Healthcare products Regulatory Agency. Available at: https://www.gov.uk/drug-device-alerts/list-of-field-safety-notices-fsns-from-5-to-9-june-2023#medtronic-cytosponge-cell-collection-device [Accessed 22 Jan 2025].

Moriarty JP, Shah ND, Rubenstein JH, et al. (2018). Costs associated with Barrett's esophagus screening in the community: an economic analysis of a prospective randomized controlled trial of sedated versus hospital unsedated versus mobile community unsedated endoscopy. Gastrointestinal Endoscopy. 87(1): 88-94.e2. doi: https://doi.org/10.1016/j.gie.2017.04.019

Muthusamy VR, Wani S, Gyawali CP, et al. (2022). AGA Clinical Practice Update on new technology and innovation for surveillance and screening in Barrett's esophagus: expert review. Clinical Gastroenterology and Hepatology. 20(12): 2696-706.e1. doi: https://doi.org/10.1016/j.cgh.2022.06.003

NHS England. (2022). National Cost Collection Data for the NHS 2020/2021 National Health Service. Available at: https://www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/ [Accessed 22 Apr 2025].

NHS England. (2024). National Cost Collection Data for the NHS 2023/2024. National Health Service. Available at: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/ [Accessed 16 Apr 2025].

NICE. (2010). Barrett's oesophagus: ablative therapy. National Institute for Health and Care Excellence. Available at: https://www.nice.org.uk/guidance/cg106/evidence/ablative-therapy-full-guideline-appendix-6-health-economics-134747683.

NICE. (2018). Oesophago-gastric cancer: assessment and management in adults. NICE guideline NG83. National Institute for Health and Care Excellence. Available at: https://www.nice.org.uk/guidance/ng83 [Accessed 16 Apr 2025].

NICE. (2020). Cytosponge for detecting abnormal cells in the oesophagus. Medtech innovation briefing MIB240. National Institute for Health and Care Excellence. Available at: https://www.nice.org.uk/advice/mib240/chapter/The-technology [Accessed 22 Jan 2025].

NICE. (2023a). Barrett's oesophagus and stage 1 oesophageal adenocarcinoma: monitoring and management NICE guideline NG231. National Institute for Health and Care Excellence. Available at: https://www.nice.org.uk/guidance/ng231 [Accessed 22 Jan 2025].

NICE. (2023b). Dyspepsia - proven GORD. Clinical Knowledge Summary. National Institute of Health and Care Excellence. Available at: https://cks.nice.org.uk/topics/dyspepsia-proven-gord/ [Accessed 17 Mar 2025].

Norton BC, Telese A, Papaefthymiou A, et al. (2025). Targeted screening of at-risk patients with chronic heartburn using a non-endoscopic capsule sponge device can enhance the detection of Barrett's oesophagus. Frontline Gastroenterology. 16(4): 285-91. doi: https://doi.org/10.1136/flgastro-2024-102854

OECD. (2024). Purchasing power parities (PPP): PPPs for GDP, for private consumption and for actual individual consumption, long time series for OECD countries. Organisation for Economic Co-operation and Development. Available at: https://www.oecd.org/en/data/datasets/purchasing-power-parities.html [Accessed 24 Apr 2024].

Office for National Statistics. (2024). National life tables: Wales. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetableswalesreferencetables [Accessed 16 Apr 2025].

Old O, Jankowski J, Attwood S, et al. (2025). Barrett's Oesophagus Surveillance Versus Endoscopy at Need Study (BOSS): a randomized controlled trial. Gastroenterology. [Article in press]. doi: https://doi.org/10.1053/j.gastro.2025.03.021

Ovrebo KK, Lie SA, Laerum OD, et al. (2012). Long-term survival from adenocarcinoma of the esophagus after transthoracic and transhiatal esophagectomy. World Journal of Surgical Oncology. 10: 130. doi: https://doi.org/10.1186/1477-7819-10-130

Pasricha S, Bulsiewicz WJ, Hathorn KE, et al. (2014). Durability and predictors of successful radiofrequency ablation for Barrett's esophagus. Clinical Gastroenterology and Hepatology. 12(11): 1840-7.e1. doi: https://doi.org/10.1016/j.cgh.2014.04.034

Peters Y, Schrauwen RWM, Tan AC, et al. (2020). Detection of Barrett's oesophagus through exhaled breath using an electronic nose device. Gut. 69(7): 1169-72. doi: https://doi.org/10.1136/gutjnl-2019-320273

Phoa KN, van Vilsteren FG, Weusten BL, et al. (2014). Radiofrequency ablation vs endoscopic surveillance for patients with Barrett esophagus and low-grade dysplasia: a randomized clinical trial. JAMA. 311(12): 1209-17. doi: https://doi.org/10.1001/jama.2014.2511

Pilonis ND, Killcoyne S, Tan WK, et al. (2022). Use of a Cytosponge biomarker panel to prioritise endoscopic Barrett's oesophagus surveillance: a cross-sectional study followed by a real-world prospective pilot. Lancet Oncology. 23(2): 270-8. doi: https://doi.org/10.1016/S1470-2045(21)00667-7

Pollit V, Graham D, Leonard C, et al. (2019). A cost-effectiveness analysis of endoscopic eradication therapy for management of dysplasia arising in patients with Barrett's oesophagus in the United Kingdom. Current Medical Research and Opinion. 35(5): 805-15. doi: https://doi.org/10.1080/03007995.2018.1552407

Provenzale D, Schmitt C, Wong JB. (1999). Barrett's esophagus: a new look at surveillance based on emerging estimates of cancer risk. American Journal of Gastroenterology. 94(8): 2043-53. doi: https://doi.org/10.1111/j.1572-0241.1999.01276.x

Ross-Innes CS, Chettouh H, Achilleos A, et al. (2017). Risk stratification of Barrett's oesophagus using a non-endoscopic sampling method coupled with a biomarker panel: a cohort study. Lancet Gastroenterology & Hepatology. 2(1): 23-31. doi: https://doi.org/10.1016/S2468-1253(16)30118-2

Ross-Innes CS, Debiram-Beecham I, O'Donovan M, et al. (2015). Evaluation of a minimally invasive cell sampling device coupled with assessment of trefoil factor 3 expression for diagnosing Barrett's esophagus: a multi-center case-control study. PLoS Medicine. 12(1): e1001780. doi: https://doi.org/10.1371/journal.pmed.1001780

Rubenstein JH, Inadomi JM, Brill JV, et al. (2007). Cost utility of screening for Barrett's esophagus with esophageal capsule endoscopy versus conventional upper endoscopy. Clinical Gastroenterology and Hepatology. 5(3): 312-8. doi: https://doi.org/10.1016/j.cgh.2006.12.008

Rubenstein JH, Mattek N, Eisen G. (2010). Age- and sex-specific yield of Barrett's esophagus by endoscopy indication. Gastrointestinal Endoscopy. 71(1): 21-7. doi: https://doi.org/10.1016/j.gie.2009.06.035

Russell LB, Gold MR, Siegel JE, et al. (1996). The role of cost-effectiveness analysis in health and medicine. JAMA. 276(14): 1172-7. doi: https://doi.org/10.1001/jama.1996.03540140060028

Saha B, Vantanasiri K, Mohan BP, et al. (2024). Prevalence of Barrett's esophagus and esophageal adenocarcinoma with and without gastroesophageal reflux: a systematic review and meta-analysis. Clinical Gastroenterology and Hepatology. 22(7): 1381-94.e7. doi: https://doi.org/10.1016/j.cgh.2023.10.006

Sami SS, Dunagan KT, Johnson ML, et al. (2015). A randomized comparative effectiveness trial of novel endoscopic techniques and approaches for Barrett's esophagus screening in the community. American Journal of Gastroenterology. 110(1): 148-58. doi: https://doi.org/10.1038/ajg.2014.362

Sami SS, Iyer PG, Pophali P, et al. (2019). Acceptability, accuracy, and safety of disposable transnasal capsule endoscopy for Barrett's esophagus screening. Clinical Gastroenterology and Hepatology. 17(4): 638-46.e1. doi: https://doi.org/10.1016/j.cgh.2018.07.019

Sami SS, Moriarty JP, Rosedahl JK, et al. (2021). Comparative cost effectiveness of reflux-based and reflux-independent strategies for Barrett's esophagus screening. American Journal of Gastroenterology. 116(8): 1620-31. doi: https://doi.org/10.14309/ajg.0000000000001336

Shaheen NJ, Falk GW, Iyer PG, et al. (2016). ACG Clinical Guideline: diagnosis and management of Barrett's esophagus. American Journal of Gastroenterology. 111(1): 30-50. doi: https://doi.org/10.1038/ajg.2015.322

Shaheen NJ, Falk GW, Iyer PG, et al. (2022). Diagnosis and management of Barrett's esophagus: an updated ACG Guideline. American Journal of Gastroenterology. 117(4): 559-87. doi: https://doi.org/10.14309/ajg.0000000000001680

Shaheen NJ, Overholt BF, Sampliner RE, et al. (2011). Durability of radiofrequency ablation in Barrett's esophagus with dysplasia. Gastroenterology. 141(2): 460-8. doi: https://doi.org/10.1053/j.gastro.2011.04.061

Shaheen NJ, Sharma P, Overholt BF, et al. (2009). Radiofrequency ablation in Barrett's esophagus with dysplasia. New England Journal of Medicine. 360(22): 2277-88. doi: https://doi.org/10.1056/NEJMoa0808145

Shariff MK, Varghese S, O'Donovan M, et al. (2016). Pilot randomized crossover study comparing the efficacy of transnasal disposable endosheath with standard endoscopy to detect Barrett's esophagus. Endoscopy. 48(2): 110-6. doi: https://doi.org/10.1055/s-0034-1393310

SHTG. (2023). Capsule sponge technologies for the detection of Barrett's oesophagus and early stage oesophageal cancer. SHTG Assessment. Scottish Health Technologies Group. Available at:

https://shtg.scot/our-advice/capsule-sponge-technologies-for-the-detection-of-barrett-s-oesophagus-and-early-stage-oesophageal-cancer/ [Accessed 22 Jan 2025].

Silvis SE, Nebel O, Rogers G, et al. (1976). Endoscopic complications: results of the 1974 American Society for Gastrointestinal Endoscopy survey. JAMA. 235(9): 928-30. doi: https://doi.org/10.1001/jama.235.9.928

Swart N, Maroni R, Muldrew B, et al. (2021). Economic evaluation of Cytosponge®-trefoil factor 3 for Barrett esophagus: a cost-utility analysis of randomised controlled trial data. EClinicalMedicine. 37: 100969. doi: https://doi.org/10.1016/j.eclinm.2021.100969

Tan WK, Ross-Innes CS, Somerset T, et al. (2025). Biomarker risk stratification with capsule sponge in the surveillance of Barrett's oesophagus: prospective evaluation of UK real-world implementation. Lancet. 406(10500): 271-82. doi: https://doi.org/10.1016/S0140-6736(25)01021-9

US Department of Health and Human Services. (2017). National vital statistics reports: United States life tables, 2014. Available at: https://www.cdc.gov/nchs/products/nvsr.htm.

Visrodia K, Zakko L, Allen J, et al. (2018). Ongoing development of a screening test for Barrett's esophagus using electronic-nose device analysis of exhaled volatile organic compounds. Gastroenterology. 154(6): S894-5 [Conference abstract: Tu1111]. doi: https://doi.org/10.1016/S0016-5085(18)33018-X

Wani S, Rubenstein JH, Vieth M, et al. (2016). Diagnosis and management of low-grade dysplasia in Barrett's esophagus: expert review from the Clinical Practice Updates Committee of the American Gastroenterological Association. Gastroenterology. 151(5): 822-35. doi: https://doi.org/10.1053/j.gastro.2016.09.040

Wolf WA, Overholt BF, Li N, et al. (2014a). Durability of Radiofrequency Ablation (RFA) in Barrett's esophagus with dysplasia: the AIM Dysplasia trial at five years. Gastroenterology. 146(5): S131 [Conference abstract: 778]. doi: https://doi.org/10.1016/S0016-5085(14)60466-2

Wolf WA, Pruitt RE, Ertan A, et al. (2014b). Predictors of esophageal adenocarcinoma in patients with prior Radiofrequency Ablation (RFA) for treatment of Barrett's esophagus: results from the U.S. RFA Registry. Gastrointestinal Endoscopy. 79(5): AB217 [Conference abstract: Sa1452]. doi: https://doi.org/10.1016/j.gie.2014.05.071

Wu J, Pan YM, Wang TT, et al. (2014). Endotherapy versus surgery for early neoplasia in Barrett's esophagus: a meta-analysis. Gastrointestinal Endoscopy. 79(2): 233-41.e2. doi: https://doi.org/10.1016/j.gie.2013.08.005

Appendix 1 – Evidence review methods

We searched for evidence that could be used to answer the review question: what is the clinical effectiveness and cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer?

The criteria used to select evidence for the appraisal are outlined in Appendix 2. These criteria were developed following comments from the Health Technology Wales (HTW) Assessment Group and UK experts.

The systematic search followed HTW's standard rapid review methodology. A search was undertaken of Medline, Embase, CINAHL, KSR Evidence, Cochrane Library, and the International Network of Agencies for Health Technology Assessment (INAHTA) HTA database. Additionally, searches were conducted of key websites and clinical trials registries. The searches were carried out in January 2025 with update searches of Medline, Embase, CINAHL, KSR Evidence, Cochrane Library, and INAHTA HTA database and forward citation searching of already included studies in Scopus conducted on 11 June 2025 and again on 30 September 2025.

Appendix 3 gives details of the search strategy used for Medline. Search strategies for other databases are available on request.

Appendix 4 summarises the selection of articles for inclusion in the review.

Appendix 2 – Inclusion and exclusion criteria for evidence included in the review

	Inclusion criteria	Exclusion criteria
Population	People with chronic acid reflux/gastroesophageal reflux disease (GORD) and suspected to have Barrett's oesophagus or under investigation for oesophageal cancer People on surveillance for Barrett's oesophagus	People with confirmed oesophageal cancer
Intervention	Capsule sponges (for example, Cytosponge and EndoSign) with TFF3, cellular atypia, and p53 testing followed by endoscopy, if indicated by capsule sponge results	
Comparison/ Comparators	Endoscopy in all	
Outcome measures	Diagnostic accuracy outcomes (e.g., sensitivity, specificity, positive predictive value, negative predictive value) with endoscopic biopsy as the reference standard Detection rates Time to diagnosis Time to treatment Safety and adverse events Health related QoL Resource use Economic outcomes	
Study design	We will prioritise the following study types, in the order listed: • Systematic reviews of randomised controlled trials. • Randomised controlled trials. • Diagnostic accuracy studies. • Non-randomised comparative trials. • Single-arm (no control group) trials that report any relevant outcome. We will only include evidence from "lower priority" sources where this is not reported by a "higher priority" source. This could be because higher priority evidence: • Does not cover all relevant populations • Does not compare the technology of interest to all relevant comparators • Does not cover all outcomes of interest • Reports over short-term follow up periods, and longer follow up data is required to facilitate decision making. Where relevant and well-conducted systematic reviews exist we will use these by:	

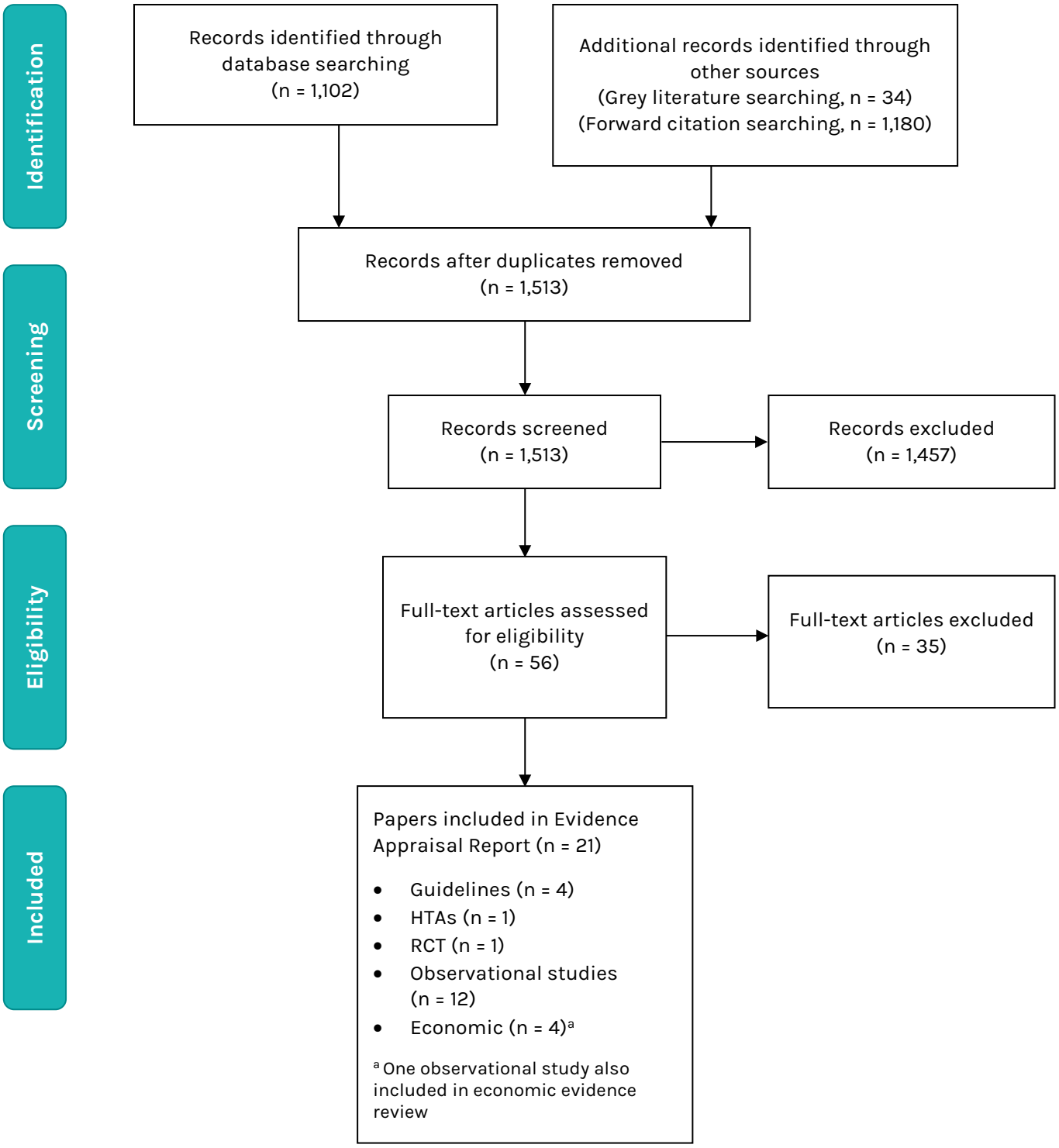
	Inclusion criteria	Exclusion criteria
	• Reporting or adapting their reported outcome measures where these are fully relevant to the scope of our review, and appropriate synthesis methods have been used • Using these reviews as a source of potentially relevant studies where the review cannot be used as a source of outcome data We will prioritise systematic reviews in terms of the sources of evidence they include, using the order described above.	
Search limits	English language only	
Other factors	Where the evidence allows, we will report outcomes separately according to list any factors identified as potentially influential on outcomes such as: • Use in primary care or secondary care	

Appendix 3 – Medline strategy

Ovid MEDLINE(R) ALL <1946 to September 29, 2025>		
Barrett's oesophagus (population)		
1	Barrett Esophagus/	9067
2	(barrett* adj3 (esophag* or oesophag* or epitheli* or metaplasia* or syndrome* or surveillanc*)).tw,kf.	11102
3	exp Esophageal Neoplasms/	64265
4	((esophag* or oesophag*) adj3 (cancer* or neoplasm* or lesion* or tumo?r* or carcinoma* or adenocarcinoma* or adenoma* or sarcoma* or malignan*)).tw,kf.	74852
5	exp Gastroesophageal Reflux/	30923
6	(gastro-oesophag* reflux* or gastrooesophag* reflux* or GORD or gastro-esophag* reflux* or gastroesophag* reflux* or GERD).tw,kf.	32712
7	((acid or acidic or gastr* or esophag* or oesophag*) adj3 reflux*).tw,kf.	38445
8	exp Esophagitis/	14233
9	(esophagitis or oesophagitis).tw,kf.	20496
10	((esophag* or oesophag* or barrett*) adj3 dysplasia*).tw,kf.	1924
11	(esophag* or oesophag* or barrett*).tw,kf. and dysplasia*.kf.	484
12	or/1-11	147039
13	Esophagus/	45804
14	(esophagus or oesophagus or esophageal or oesophageal or barrett*).tw,kf.	195069
15	or/13-14	204284
16	Precancerous Conditions/	30532
17	"Early Detection of Cancer"/	45483
18	or/16-17	75058
19	15 and 18	3076
20	((precancer* or pre-cancer*) adj3 (barrett* or esophag* or oesophag*)).tw,kf.	328
21	((cancer* or neoplasm* or lesion* or tumo?r* or carcinoma* or adenocarcinoma* or adenoma* or sarcoma* or malignan*) adj3 (screen* or diagnos* or detect*) adj3 (barrett* or esophag* or oesophag*)).tw,kf.	3331
22	or/19-21	6215
23	12 or 22	147220
Capsule Sponge (intervention)		
24	Trefoil Factor-3/	736
25	("trefoil factor 3" or TFF3 or "TFF-3" or "TFF 3").tw,kf.	1013
26	(intestinal adj2 trefoil factor).tw,kf.	235
27	(capsule adj3 (sponge* or balloon* or swallow*)).tw,kf.	446
28	(sponge adj3 (string* or cytolog* or test*)).tw,kf.	183
29	((esophag* or oesophag*) adj3 (sponge* or string*)).tw,kf.	73
30	(((esophag* or oesophag*) adj3 cell collection device*) or OCCD or ECCD).tw,kf.	125
31	((esophag* or oesophag*) adj3 cell sampling device*).tw,kf.	4
32	((nonendoscop* or non-endoscop*) adj3 (screen* or diagnos* or detect* or test* or surveillanc* or cytolog*)).tw,kf.	98
33	(minimally invasive and (nonendoscop* or non-endoscop*)).tw,kf.	75
34	or/24-33	2286
Barrett's oesophagus AND capsule sponge		
35	23 and 34	220
Brands & final check		
36	"sponge on a string".tw,kf.	4
37	cytosponge*.tw,kf.	85
38	endosign*.tw,kf.	0
39	cytoprime*.tw,kf.	0
40	esophacap*.tw,kf.	6
41	esocheck*.tw,kf.	12
42	or/36-41	99
43	(Barrett Esophagus/di or exp *Esophageal Neoplasms/di or exp Gastroesophageal Reflux/di or exp Esophagitis/di) and sponge*.tw,kf.	32
Final set combination		
44	35 or 42 or 43	267

45	limit 44 to english language	252
46	exp animals/ not humans/	5380029
47	(baboon*1 or bovine*1 or canine*1 or cat or cats or chimpanzee*1 or cow*1 or dog*1 or feline*1 or goat*1 or hens or macque*1 or mice or monkey*1 or (mouse adj2 model*1) or murine*1 or ovine or pig*1 or porcine or (non-human adj2 primate*1) or sheep or rabbit*1 or rat or rats or rattus or rhesus or rodent*1 or zebrafish).ti.	2274568
48	or/46-47	5820542
49	45 not 48	249

Appendix 4 – Flow diagram outlining selection of relevant evidence sources



Appendix 5 – Full sources of evidence and outcome data

Table A1 – Included health technology assessments and reviews: design and characteristics

Review	Design, search period	Eligibility criteria	Trial/patient characteristics	Outcome measures	Comments
SHTG (2023)	Limited systematic review of SRs, HTAs and other evidence-based reports Search dates 3 to 7 July 2023 Includes primary data from NHS Scotland and England. Some Scottish data now published in Chien et al. (2024a), Chien et al. (2024b), and English data published in Gourgiotis et al. (2025)	Not explicitly reported	1 SR of 13 studies (n = 3,786), 1 cross-sectional study (n = 891) and prospective cohort analysis (n = 223), 1 RCT (n = 13,514), 1 retrospective cohort analysis (n = 10,577), 1 patient survey (n = 1,458), 1 retrospective analysis of 5 prospective cohort analyses (n = 2,418)	• Time to diagnosis • Time to treatment	• We have only extracted analysis of data from NHS Scotland, included in the report, that has not been published elsewhere.

Abbreviations: HTA: health technology assessment; RCT: randomised controlled trial; SR: systematic review

Table A2 – Randomised controlled trial: design and characteristics

Study reference	Study details	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
Fitzgerald et al. (2020)	109 GP clinics (England) Study dates: 20 March 2017 to 21 March 2019 Inclusion criteria: aged 50 years or older and prescribed acid-suppressant therapy (proton-pump inhibitor or histamine-2 receptor antagonists) for at least 6 months in the previous year. Exclusion criteria: Patients prescribed non-steroidal anti-inflammatory drugs together with acid-suppressant therapy, suggesting that their reflux symptoms were not the primary basis for the proton-pump inhibitor prescription, and patients who had undergone an endoscopy in the previous 5 years or with a previous diagnosis of BO.	n = 13,222 Intervention: n = 6,834. 1,654 successfully swallowed the capsule sponge device, demographics for these reported below. Age distribution 50 to 59 years 20%, 60 to 69 years 34%, 70 to 79 years 37%, 80 to 89 years 8%, 90 to 99 years 1% 48% male Median (IQR) Index of Multiple Deprivation decile NR Control: n = 6,388 Age distribution NR % male NR Median (IQR) Index of Multiple Deprivation decile 6 (4 to 9)	Intervention: Standard management and offered Cytosponge testing, with a subsequent endoscopy if the procedure identified TFF3-positive cells Control: Standard management of their symptoms, only referred for an endoscopy if required	• Diagnostic accuracy to detect BO, dysplasia, or oesophago-gastric cancer • Detection rates • Safety and adverse events	Passive follow-up from 8 to 18 months	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Initially cluster randomised by GP clinic. Approximately two-thirds of the way through recruitment this switched to individual randomisation and Cytosponge was available at all clinics. • Pathologists analysing endoscopic biopsies were blinded to Cytosponge results. • Standard management included prescriptions for acid-suppressant medication, lifestyle advice from their GP, referral for an endoscopy depending on the severity of their symptoms. • Cytosponge testing was optional in the intervention group, could introduce some selection bias as those who agree could ha had more problematic symptoms. ITT analysis used to mitigate this bias. • Participants were not offered Cytosponge if they had dysphagia, were at increased risk of bleeding because of known cirrhosis or varices, or if they were unable to stop taking anticoagulants.

Abbreviations: BO: Barrett's oesophagus; IQR: interquartile range; ITT: intention-to-treat; NR: not reported; TFF3: Trefoil factor 3;

Table A3 – Observational studies: design and characteristics

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
Angel et al. (2025)	Prospective cohort study Multicentre (England) Study dates: November 2020 to October 2023 Inclusion criteria: patients with symptoms of reflux including heartburn, regurgitation, and waterbrash. Exclusion criteria for Cytosponge (absolute contraindications): • Alarm symptoms (dysphagia, dyspepsia and weight loss, dyspepsia and anaemia) • previous cancer of the oesophagus • diagnosis of an oropharyngeal, oesophageal or gastro- oesophageal tumour • had treatment to the oesophagus for example, photodynamic therapy, endoscopic mucosal resection, radio frequency ablation, surgery • known to have oesophageal varices or cirrhosis of the liver	n = 871 (808 successfully swallowed capsule sponge device, 763 adequate samples) Median (IQR) age 54 (41.0 to 65.5) years 40.1% male Patients with adequate samples: Median (IQR) age 54 (41.0 to 64.0) years for males, 56 (42.6 to 65.7) years for females	Intervention: Cytosponge/EndoSign with H&E staining, and TFF3 and p53 testing Comparator/reference standard: Endoscopic biopsy	• Diagnostic accuracy to detect BO • Detection rates • Safety and adverse events	12 to 48 months	• Investigation of reflux symptoms. • Study started during the COVID-19 pandemic when usual endoscopy services were disrupted. • All patients were recruited to the DELTA or NHS England evaluations reported elsewhere. • For those who had a negative capsule sponge test and were not offered endoscopy, a review of the Medilogik EMS database was undertaken at 1, 2 and 3 years from the test to see if they had been referred back to endoscopy and to review subsequent endoscopy findings. • From November 2020 to June 2023, the Cytosponge device was used and from July 2023 onwards, the EndoSign device was used. • Only patients with abnormal, inadequate or failed capsule sponge tests or ongoing/concerning symptoms had endoscopy.

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	• known anomaly of the oesophagus for example, webbing, pouch, stricture and so forth • patients who are pregnant • unable to give consent • patients who have had a stroke or any other neurological disorder where their swallowing has been affected. • patients who have had a myocardial infarction in the last 3 months. Patients to consider as having relative contraindications to Cytosponge use: Patients who have had fundoplication may be candidates for Cytosponge but may have reflux symptoms post procedure					
Chien et al. (2024a)	Prospective cohort study Multicentre (Scotland) Study dates: 14 September 2020 to 30 April 2023 No specific inclusion criteria. All patients referred from primary care, in the absence of red flag symptoms (i.e. dysphagia, weight loss, anaemia), on the routine reflux pathway were considered eligible: reflux symptoms	n = 1,305 patients, 1,385 Cytosponge tests Median (IQR) age 56 (46 to 65) years 42.4% male Median BMI 28.1 (25 to 32.4) Positive smoking history 37.5% Proton pump inhibitor use 88.2%	Intervention: Cytosponge with H&E staining, and TFF3 and p53 testing Comparator/reference standard: Endoscopy (with or without biopsy)	• Detection rates • Safety and adverse events	NR	• Part of CytoScot analysis. • Investigation of reflux symptoms. • Pilot was conducted during the COVID-19 pandemic when usual endoscopy services were disrupted. • Databases were prospectively maintained. • Data on BMI missing from 411 patients. Data on smoking history missing from 26 patients.

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	generally included burning sensation, acid taste, waterbrash, and/or regurgitation. Exclusion criteria: inability to tolerate capsule sponge testing; capsule sponge testing for Barrett's surveillance; those with outstanding histopathology results at time of analysis; paediatric population (age less than 18 years). Contraindications to capsule sponge testing were specified by the manufacturer and included: pregnancy; liver disease including cirrhosis; oesophageal varices; significant dysphagia; previous oesophageal tumour; oesophageal surgery (including endoscopic therapy).					• 80 tests were repeat tests performed due to insufficient first samples or assessment of inflammation healing. • If UGI tract appeared macroscopically normal during endoscopy, no biopsy was taken.
Chien et al. (2024b)	Retrospective cohort analysis Multicentre (Scotland) Study dates: 14 September 2020 to 30 April 2023 Patients were recruited for capsule sponge testing if previously entered in local Barrett's surveillance programmes, where prior endoscopy demonstrated	n = 3,745, 4,204 Cytosponge tests. n = 608 underwent UGI endoscopy within 12 months and were included in analysis. Median (IQR) age 67 (60 to 73) years 70.2% male Median follow-up time 14 (8 to 22) months	Intervention: Cytosponge with H&E staining, and TFF3 and p53 testing Comparator/reference standard: Endoscopic biopsy	• Detection rates	NR	• Part of CytoSCOT analysis. • People under surveillance for BO • Pilot was conducted during the COVID-19 pandemic when usual endoscopy services were disrupted. • Databases were prospectively maintained. • The presence of IM on endoscopic biopsies was not considered a prerequisite for entry into surveillance.

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	macroscopic changes consistent with BO. All patients who subsequently underwent UGI endoscopy within 12 months of capsule sponge test with available histopathology results were identified and included in this analysis. Exclusion criteria: capsule sponge test for reflux symptoms; patients who did not undergo UGI endoscopy within 12 months of capsule sponge test; outstanding histopathology results at the time of analysis; previous endoscopic treatment for dysplasia.	Median time from last endoscopy to Cytosponge test 38 (29 to 48) months 83.7% demonstrated IM on previous endoscopic biopsies				
Chien & Glen (2025)	Prospective cohort study Multicentre (Scotland) Study dates: 1 January 2018 to 31 December 2022 Eligibility criteria: all patients undergoing endoscopic surveillance for BO in a single Scottish health board with a minimum of one previous endoscopy with biopsies showing biopsies either confirmed the presence of IM and/or prior endoscopy demonstrated macroscopic changes consistent with BO.	n = 3,359 Pre-intervention group (n = 1,568): all patients undergoing endoscopic surveillance from 1 January 2018 to 31 December 2019 Median (IQR) age 65 (57 to 72) years 64.3% male Proton-pump inhibitor use 95.6% IM on last endoscopic pathology results 82.1% Median (IQR) time from last endoscopic surveillance 25 (23 to 34) months	Intervention: Cytosponge with H&E staining, and TFF3 and p53 testing Comparator/reference standard: Endoscopic biopsy	• Detection rates	NA	• People under surveillance for BO. • Patients were invited to undertake capsule sponge testing in lieu of surveillance endoscopy in the absence of red flag symptoms. • The presence of IM on endoscopic biopsies was not considered a prerequisite for entry into surveillance. • Capsule sponge testing was introduced to the health board in September 2020, due to the temporary halting of routine endoscopy services in response to the Covid-19 pandemic. Surveillance

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	Exclusion criteria: under 18 years old; previous endoscopic eradication therapy with RFA or EMR; previous oesophagectomy; presence of dysplasia in last endoscopic biopsies, intramucosal adenocarcinoma, or invasive cancer in previous oesophageal biopsies; EMR or oesophagectomy specimens; slides referred from other health boards; squamous dysplasia; repeat endoscopy for mapping biopsies in cases of HGD or intramucosal adenocarcinoma	Implementation group (n = 1,791): patients who underwent surveillance with both endoscopy and capsule sponge testing from 1 January 2021 to 31 December 2022. This group was then split into two cohorts: capsule sponge testing with or without endoscopy (n = 920) and endoscopic surveillance only (n = 871). Median (IQR) age 66 (57 to 73) years 63.9% male Proton-pump inhibitor use 94.8% IM on last endoscopic pathology results 76.8%[a] Median (IQR) time from last endoscopic surveillance 35 (27 to 45) months[a] [a] $p < 0.001$ compared to pre-intervention group				undertaken in 2020 was discarded from analyses. • Patients with red flag symptoms were excluded from capsule sponge testing, but were included in the endoscopy only group.
Eluri et al. (2022)	Prospective cohort study 4 tertiary care referral centres in UK and 1 in USA Study dates: NR	n = 175 (175 of 234 patients had adequate Cytosponge samples), 142 had endoscopic and histologic data available and were included in primary analysis	Intervention: Cytosponge with H&E staining and TFF3 testing Comparator/reference standard: Endoscopy	• Diagnostic accuracy to detect residual BO	NR	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • All patients had received prior ablative treatment for BO

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	Eligibility criteria: adults over 18 years with dysplastic BO, LGD, HGD, or intramucosal adenocarcinoma, confirmed by a second expert gastrointestinal pathologist, who had undergone at least one round of endoscopic eradication therapy and were scheduled for further ablative therapy or endoscopic surveillance after complete eradication of IM	Mean age 71 ± 9 years 83% male 65% History of endoscopic mucosal resection Median (IQR) time since first ablation 20 (2 to 113) months Median time since last ablation 10 (1 to 111) months	(with or without biopsy)			• All patients underwent upper endoscopy approximately 2 hours after Cytosponge administration. • Biopsies were obtained from BO segments in those with residual BO undergoing further endoscopic treatment, and from the cardia, gastroesophageal junction, and neosquamous oesophagus in post-complete eradication of IM patients. A subset of patients (n = 33) undergoing ablation, but had not achieved complete eradication, only had endoscopic evidence of columnar epithelium documented, without concurrent biopsies, due to the endoscopist's concern of biopsies interfering with ablation. • Presence of BO was defined as columnar epithelium of greater than or equal to 1 cm in the tubular oesophagus, with concurrent IM on biopsies or endoscopic mucosal resection specimens of that area.
Gourgiotis et al. (2025)	Prospective cohort study 23 hospitals (England)	<u>Intervention group</u> n = 2,875 (1,549 with sufficient data for detailed analysis)	Intervention group: Cytosponge with H&E staining and TFF3 testing	• Detection rates • Safety and adverse events	NR	• Triage for reflux symptoms • Developer of Cytosponge and co-founder of Cyted involved in study.

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	Study dates February 2021 to August 2022 Eligibility criteria: Patients with symptoms of reflux including heartburn, regurgitation, waterbrash Exclusion criteria for Cytosponge (absolute contraindications): • Alarm symptoms (dysphagia, dyspepsia and weight loss, dyspepsia and anaemia) • previous cancer of the oesophagus • diagnosis of an oropharyngeal, oesophageal or gastro- oesophageal tumour • had treatment to the oesophagus for example, photodynamic therapy, endoscopic mucosal resection, radio frequency ablation, surgery • known to have oesophageal varices or cirrhosis of the liver • known anomaly of the oesophagus for example, webbing, pouch, stricture and so forth • patients who are pregnant	Median (IQR) age at referral 52 (40 to 62) years 42.3% male Median time between referral and index date 27 (13 to 70) days 80.4% White, 19.6% non-White Heartburn 14.8% Waterbrash 0.9% Reflux 74.2% Use of acid suppressants within last 6 months 84.1% <u>Counterfactual group</u> n = 1,181 (demographics not reported but stated to be similar to the intervention group)	Counterfactual group: Routine endoscopy			• Patients that were ineligible for Cytosponge or declined were excluded.

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	(Relative contraindication, Cytosponge not harmful but may not be appropriate): • unable to give consent • patients who have had a stroke or any other neurological disorder where their swallowing has been affected. • patients who have had a myocardial infarction in the last 3 months. Patients to consider as having relative contraindications to Cytosponge use: Patients who have had fundoplication may be candidates for Cytosponge but may have reflux symptoms post procedure					
Kadri et al. (2010)	Prospective cohort study 12 GP clinics (England) Study dates May 2008 to December 2009 Inclusion criteria: adults aged 50 to 70 years with a previous prescription for an acid suppressant (H_2 receptor antagonist or proton pump inhibitor) for more than three months in the past five years Exclusion criteria: previous diagnosis of BO, gastroscopy	n = 501 Median (range) age 62 (56 to 66) years 45.7% male 95.8% White, 4.2% other ethnicity GORD impact scores: 7.0% very well controlled, 19.8% fairly well controlled, uncontrolled 27.1%, poorly controlled 38.9%, very poorly controlled 7.2% Current use of acid suppressants: 13.4% antacids, 7.6% H_2 antagonists, 57.0% proton	Intervention: Cytosponge with TFF3 testing Comparator/reference standard: Endoscopy with biopsy	• Diagnostic accuracy to detect BO • Safety and adverse events	NA	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted, however they had no interests to declare at the time of publication. • GP clinics sent eligible participants an invitation to take part in the study. Those who agreed were sent an appointment for the Cytosponge test at the practice. • Participants who successfully swallowed the Cytosponge were invited to

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	within the past year, dysphagia, known portal hypertension, drug or pathophysiological abnormality of coagulation, important physical or psychological comorbidity precluding gastroscopy, inability to provide informed consent	pump inhibitors, 1.8% H_2 and proton pump inhibitors, 20.2% none				attend for a gastroscopy within three weeks of the Cytosponge procedure. BO was defined as endoscopically visible columnar lined epithelium arising at least 1 cm circumferentially above the gastro-oesophageal junction with IM. If BO was present, four biopsies every 2 cm were collected according to surveillance guidelines. • Endoscopists and histopathologists were blinded to the result of the Cytosponge test. • 32 participants did not attend for gastroscopy and were considered not to have BO in analyses.
Norton et al. (2025)	Cross-sectional study UK Study date February 2024 High-risk criteria: age greater than or equal to 40 years old, current or ex- smoker, no prior investigations, regular acid suppression use and family history of BO or OAC. Exclusion criteria: symptoms of dysphagia, known oesophago-gastric varices, previous upper	n = 60 (12 positive EndoSign tests, 11 accepted endoscopy offer) (Of 78 participants invited to undergo EndoSign test): Mean age 57.1 ± 9.4 years 85.9% male Demographics of the 60 participants who successfully swallowed the capsule, and were included in analysis, NR	Intervention: EndoSign with H&E staining, and TFF3 and p53 testing Comparator/reference standard: Endoscopy (with or without biopsy)	• Diagnostic accuracy to detect BO • Detection rates • Safety and adverse events	NA	• The study was part of a charity campaign that was supported by Cyted. • Members of the public from Greater London were invited to complete an online self-referral screening questionnaire between December 2023 to February 2024 through an advertising campaign. • Individuals who had chronic heartburn who were deemed to be high-risk were subsequently invited to undergo a free heartburn

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	gastrointestinal surgery, pregnancy, prior diagnosis of BO or OAC, and the use of anti-coagulants that could not be stopped.					health check with EndoSign. Once all individuals meeting four or more of the high-risk criteria were contacted, individuals meeting two or more of these criteria were invited on a first come, first serve basis until all appointments were allocated. • Not part of NHS pathways. EndoSign testing carried out in mobile units, those with positive results were sent to a private clinic for confirmatory gastroscopy. Anyone with clinically actionable findings was referred to their GP for ongoing care.
Pilonis et al. (2022)	Cross-sectional study and prospective cohort analysis Multicentre (England) Study dates 7 July 2011 to 1 April 2019 (cross-sectional), participants recruited from 27 August 2020 (prospective) Cross-sectional study inclusion criteria: all available consecutive adult patients with a confirmed diagnosis of BO (with IM confirmed by TFF3 and a minimum Barrett's segment length of 1 cm) who were having endoscopic surveillance as part	Cross-sectional study (n = 891): Training cohort n = 557 Median (IQR) age 65 (59 to 72) years 81% male 98% white, 2% other ethnicity Median (IQR) BO maximum segment length 5 (3 to 8) cm Median (IQR) BO circumferential length 3 (1 to 6) cm Median (IQR) BMI 28.25 (25.61 to 31.07) Validation cohort n = 334	Intervention: Cytosponge with cellular atypia and p53 testing Comparator/reference standard: Endoscopy with biopsy	• Diagnostic accuracy to detect dysplasia or intramucosal cancer • Detection rates		• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Eligible participants were split into training and validation cohorts on the basis of date of recruitment (training cohort 2011–13, validation cohort 2013 onwards). • Endoscopies were performed on the same day as Cytosponge (BEST2) or within 2 months of Cytosponge (BEST3). • At the time of publication, endoscopy data for the

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	of the BEST2 and BEST3 clinical trials. Prospective cohort analysis inclusion criteria: patients who had BO surveillance delayed due to COVID-19 pandemic. Endoscopy results were available for patients categorised as high risk by Cytosponge.	Median (IQR) age 67 (58 to 73) years 75% male Ethnicity NR Median (IQR) BO maximum segment length 3 (2 to 6) cm Median (IQR) BO circumferential length 1 (0 to 4) cm Median (IQR) BMI 27.90 (25.20 to 30.81) Prospective cohort analysis (n = 223): Median age 69 (IQR 60 to 74) years 74% male Ethnicity NR Median (IQR) BO maximum segment length 3 (2 to 6) cm Median (IQR) BO circumferential length 1 (0 to 4) cm Median (IQR) BMI 26.90 (24.12 to 29.30)				prospective cohort were still being collected for low-risk and moderate-risk groups. • Prospective cohort is part of the DELTA trial.
Ross-Innes et al. (2015)	Case-control study 11 centres (UK) Study dates NR	n = 1,110 (647 cases, 463 controls) Cases were individuals with a previous diagnosis of BO attending for their monitoring endoscopy. Controls were individuals referred to endoscopy	Intervention: Cytosponge with TFF3 testing Comparator/reference standard: Endoscopy with biopsy	• Diagnostic accuracy to detect BO • Safety and adverse events	NA	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • BO was defined as endoscopically visible columnar-lined oesophagus that measured at least 1 cm circumferentially or at least 3 cm in non-circumferential

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
		because of dyspepsia and/or reflux symptoms. <u>Cases</u> Median (IQR) age 66 (58 to 73) years Male:female ratio 4:1 96.8% white, 1.8% other ethnicity Median (IQR) BMI 28.1 (25.6 to 31.2) Maximum length of BO (median [IQR]) 5 (3 to 8) cm <u>Controls</u> Median (IQR) age 56 (44 to 66) years Male:female ratio 1.0:1.3 92.5% white, 7.3% other ethnicity Median (IQR) BMI 26.8 (24.0 to 30.2) Maximum length of BO NA Exclusion criteria: patients with bleeding diatheses or on anticoagulant medication, known cirrhosis, varices, or dysphagia.				tongues with documented histopathological evidence of intestinal metaplasia (IM) on at least one biopsy in the course of their endoscopic history. • Participants who were initially enrolled as controls but then diagnosed with BO at endoscopy were crossed over to the case arm. • Four tertiary referral centres for BO were included to enrich for cases of BO with dysplasia, in case dysplasia adversely affected the sensitivity of the assay. • Endoscopy was performed within one hour of Cytosponge testing. • Participants under surveillance for BO who happened to undergo a second surveillance endoscopy for clinical purposes during the study period were invited to take a Cytosponge test again. • Those scoring Cytosponge samples were blinded to the patient's diagnosis and histocytopathologists reviewing biopsy results were blinded to Cytosponge results.

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
Ross-Innes et al. (2017)	Case-control study Multicentre (UK) Study dates NR	<u>Discovery cohort (n = 468)</u> Non-dysplastic BO (n = 376): Median (IQR) age 64 (56 to 71) years Male:female ratio 3.8:1 97% white, 2% other ethnicity, less than 1% refused to disclose Median (IQR) BMI 28.1 (25.5 to 30.8) BO with HGD or IMC (n = 92): Median (IQR) age 69 (63 to 74) years Male:female ratio 7.4:1 99% white, 1% other ethnicity Median (IQR) BMI 28.8 (26.1 to 31.1) Inclusion criteria: all BO patients with IM and a TFF3-positive Cytosponge test. No minimum BO segment length was required provided participants had a least one TFF3-positive cell on Cytosponge.	Intervention: Cytosponge with TFF3 and p53 testing Comparator/reference standard: Endoscopy with biopsy	• Diagnostic accuracy to detect HGD or IMC	NA	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted. • Endoscopy was performed within one hour of Cytosponge testing. • Biopsy samples were taken from any visible lesions and from each quadrant, every 2 cm. • Pathologists reviewing biopsy results were blinded to Cytosponge results.
Tan et al. (2025)	Prospective cohort study Multicentre (UK)	n = 910 Consecutive patients undergoing BO surveillance from 13	Intervention: Cytosponge/EndoSign with cellular atypia and p53 testing	• Diagnostic accuracy to detect HGD or cancer and	NR	• Several authors were involved in the development of Cytosponge and founding / employed by Cyted.

Study reference	Methods, setting	Participants	Intervention(s)	Outcomes	Follow-up, months	Comments
	Study dates: August 2020 to December 2024 Inclusion criteria: at least 18 years old with a diagnosis of non-dysplastic BO at their last endoscopy and undergoing surveillance. All patients had to successfully swallow a capsule sponge and have a confirmatory endoscopy. Exclusion criteria: any previous history of HGD, previous endoscopic or surgical intervention therapy to the oesophagus, contraindications according to the device manufacturer instructions, lack of capacity to provide informed consent.	hospitals in the UK who participated in the DELTA study and the NHS England implementation pilot study. Median (IQR) age 68 (60 to 74) years 76% male Histology at baseline: • Non-dysplastic BO 90% • Indefinite for dysplasia 1% • Crypt dysplasia < 1% • LGD 5% • HGD or intramucosal carcinoma 3% • Adenocarcinoma (≥ T2) 1%	Comparator/reference standard: Endoscopy with biopsy	any level of dysplasia • Detection rates		• The DELTA study and the NHS England implementation pilot study followed the same protocol. • Patients were assigned to low- or moderate-risk groups at baseline based on clinical risk factors and previous BO findings. Patients were escalated to the high-risk group after capsule sponge testing if their results showed any of atypia, atypia of uncertain significance, equivocal p53, or aberrant p53 expression. • Study took place during Covid-19 pandemic when endoscopy services were disrupted. • Some patients had more than one endoscopy follow-up, for example for indefinite for dysplasia or first diagnosis of LGD, which followed the clinical standard of a repeat at 6 months.

Abbreviations: BEST: Barrett's oEsophagus Screening Trial; BMI: body mass index; Barrett's oesophagus; DELTA: integrateD diagnostic solution for EarLy deTection of oesophageal cAncer; EMR: endoscopic mucosal resection; H&E: haematoxylin and eosin; HGD: high-grade dysplasia; IM: intestinal metaplasia; IMC: intramucosal adenocarcinoma; IQR: interquartile range; LGD: low-grade dysplasia; NA: not applicable; NR: not reported; OAC: oesophageal adenocarcinoma; RFA: radiofrequency ablation; TFF3: Trefoil factor 3; UGI: upper gastrointestinal

Appendix 6 – HTW cost utility analysis

1. Background and objective

An economic model was developed to estimate the cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer in people with chronic reflux.

A separate evaluation for the surveillance population was not conducted. While diagnostic accuracy evidence for this population exists, studies have only assessed single-timepoint performance, and it is uncertain whether accuracy would be maintained across repeated rounds of surveillance. In addition, there is uncertainty in the long-term disease progression following endotherapy treatment for Barret's oesophagus. The need for additional assumptions around surveillance intervals, disease progression risks, and repeat test performance led to a focus on the chronic reflux population for this evaluation, where available disease progression models are more established.

2. Methods

2.1 Model approach

Our modelling approach, developed in Microsoft Excel, combines a cohort-level decision tree and Markov model to estimate the cost effectiveness of capsule sponge devices to detect Barrett's oesophagus and early-stage oesophageal cancer, aligning with approaches taken in economic studies identified in the economic review (Section 6.1 of the main report). The decision tree captures short-term diagnostic outcomes, and the Markov model captures long-term disease progression, costs, quality of life (QoL) and mortality. The structure of the Markov model closely follows previous cost-utility analyses developed in this disease area (Benaglia et al. 2013, Sami et al. 2021, Swart et al. 2021).

The baseline population considers people with chronic reflux for the detection of Barrett's oesophagus. The analysis compares the diagnostic strategies described in Table A4.

Table A4 Diagnostic strategies included in the base case economic model

Diagnostic strategy	Description
Intervention	
Cytosponge	All patients receive a Cytosponge test in primary care. Biomarkers are then tested where patients with a positive result receive an endoscopic biopsy in secondary care.
Comparator	
Endoscopy	All patients receive an endoscopic biopsy in secondary care.

The model takes the perspective of NHS Wales and personal social services (PSS). Analyses are conducted over a lifetime horizon and future costs and benefits are discounted at a rate of 3.5% per annum.

Figures 3 and 4 illustrate the diagnostic (decision tree) and Markov model components of the economic model, respectively.

People with chronic reflux enter the diagnostic component of the model, where they undergo diagnostic testing. In the intervention arm, patients receive a capsule sponge test in primary

care. Those with negative results (i.e. true and false negatives) are assumed not to have any further testing and enter the Markov model. Patients with positive results (i.e. true and false positives) undergo confirmatory endoscopic biopsy in secondary care before entering the Markov model. Any false positive cases from the capsule sponge test are confirmed not to have Barrett's oesophagus at this stage. It is assumed those who are unable to swallow the capsule sponge device or experience sponge detachment receive an endoscopic biopsy in secondary care. In the comparator arm, endoscopic biopsy is performed in secondary care to directly confirm cases of Barrett's oesophagus before they enter the Markov model. All patients where Barrett's oesophagus or progressed disease has been confirmed via endoscopy receive treatment based on their health state within the Markov model. The cost of diagnostic testing is captured for all patients. The care setting at which the capsule sponge test is administered is explored is scenario analysis.

Following diagnostic testing, patients are distributed between their corresponding health states in the Markov model, including no Barrett's oesophagus, nondysplastic Barrett's oesophagus (NDBO), low-grade dysplasia (LGD), high-grade dysplasia (HGD), and early-stage oesophageal adenocarcinoma (OAC). Over time, patients may progress to more severe health states, including late-stage OAC and death. An annual cycle length was considered, where natural disease progression was aligned to the modelling approach taken in Swart et al. (2021).

Similar to previous economic models, patients diagnosed with LGD, HGD or early-stage OAC were treated with endotherapy which aims to completely eradicate dysplasia. Endotherapy may also result in complete eradication of any concurrent intestinal metaplasia. Due to this treatment effect, patients could transition to the no Barrett's oesophagus or NDBO health states in the cycle following endotherapy. Aligned to Sami et al. (2021), patients in these health states are assumed to only have a single instance of endotherapy treatment regimens. Therefore, patients re-entering these health states following improvement from initial endotherapy would not undergo subsequent endotherapy treatment. Patients with late-stage OAC received oesophagectomy or palliative cancer treatments depending on if they were suitable for surgery. Progression to late-stage OAC is assumed to directly lead to clinical intervention due to the presence of symptoms. In all health states, patients are at risk of mortality from any cause. Closely aligning to Swart et al. (2021), those who enter the OAC (late-stage) health state who are not suitable for oesophagectomy are assumed to transition to death in the subsequent model cycle.

Those diagnosed with Barrett's oesophagus receive endoscopic surveillance which stops if patients progress to late-stage OAC. Patients identified with LGD, HGD, or early-stage OAC through surveillance, who have not previously undergone endotherapy treatment, proceed to receive endotherapy.

Patients in the intervention arm with a false negative result are assigned to their corresponding true health state following diagnostic testing, and do not receive endotherapy treatment or endoscopic surveillance.

Costs and utilities are applied to patients within each health state accordingly. Treatment costs for proton pump inhibitor (PPI) therapy are not considered as it is assumed all patients receive this due to their underlying chronic reflux. Treatment costs for patients undergoing endotherapy, oesophagectomy or palliative care are applied accordingly. Health state utility values are applied to patients regardless of if the disease has been diagnosed.

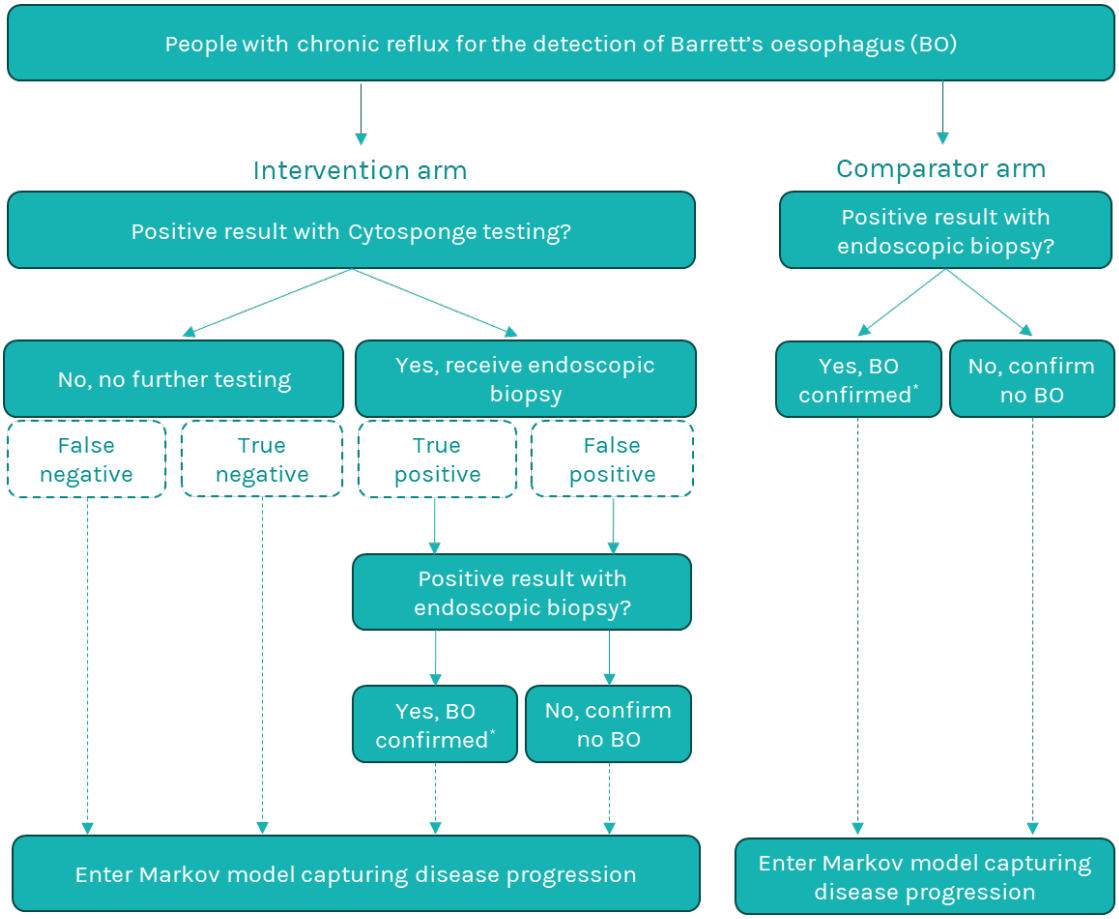
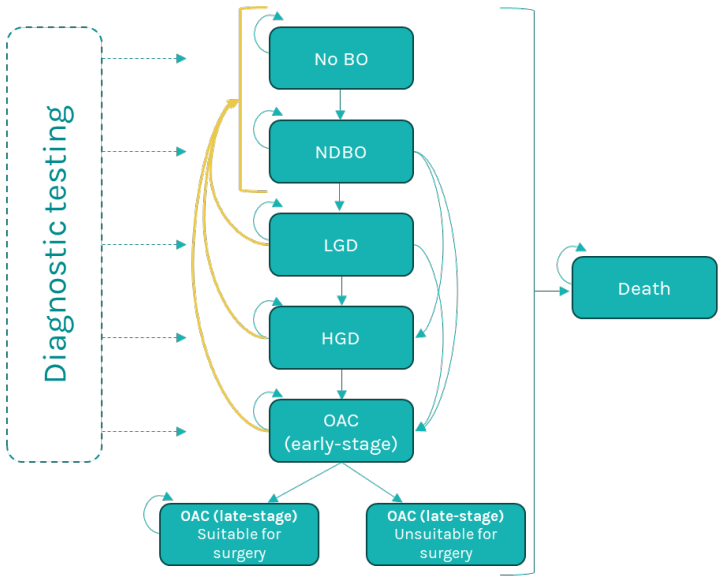


Figure 3 – Model schematic: diagnostic component (decision tree)



Green arrows = natural disease progression between health states; Yellow arrows = treatment effect due to endotherapy
BO, Barrett's oesophagus; HGD, high-grade dysplasia; LGD: low-grade dysplasia; NDBO, nondysplastic Barrett's oesophagus; OAC, oesophageal adenocarcinoma

Figure 4 – Model schematic: Markov model component

2.2 Diagnostic accuracy

The sensitivity and specificity of the Cytosponge test were sourced from a prospective cohort study undertaken in 12 practices in the UK by Kadri et al. (2010). The values used in the model are presented in Table A5 and are based on a cut-off segment length of 1 cm or more. Diagnostic accuracy values are applied uniformly across all severities of Barrett's oesophagus.

As described in Section 5.2 of the main report, this study evaluated outcomes in patients who underwent a Cytosponge-TFF3 test in primary care. Whist this study reflects diagnostic accuracy outcomes in the primary care setting, it is assumed these outcomes can be applied to other care settings explored in this economic evaluation, where the population remains consistent.

This study was considered an appropriate source for diagnostic outcome estimates as its study population was most generalisable to the population of interest for this economic evaluation (i.e. the chronic reflux population), compared to other diagnostic studies where some of which were enriched for dysplasia. However, limitations of this study should be noted. As the patient population includes those who had received reflux medication for more than three months in a five year period, it is possible that this group is broader than our target population and may include patients whose reflux has resolved. Additionally, as the study was undertaken in 2008 – 2009, it's outcomes may not accurately reflect current practices.

As the study used endoscopic biopsy as the reference standard, the comparator arm is assumed to be perfectly accurate with a sensitivity and specificity of 100%. While experts contacted by HTW noted that this is not perfectly accurate in reality, this assumption was necessary to remain consistent with the evidence base and to allow a relative comparison to test performance.

The diagnostic accuracy of the intervention is explored in sensitivity analyses.

Table A5 – Diagnostic accuracy inputs

Diagnostic outcome	Mean[a] (%)	SE[b] (%)	Source
Sensitivity	73.3 (44.9 - 92.2)	12.1	Kadri et al. (2010)
Specificity	93.8 (91.3 - 95.8)	1.1	Kadri et al. (2010)

Abbreviations: SE, standard error

[a] 95% CIs displayed in brackets
[b] Sampled from a beta distribution.

2.3 Baseline characteristics

Inputs related to the baseline characteristics, used for the base case model, are presented in Table A6.

Baseline characteristics for age and sex were aligned to the Kadri et al. (2010) study. The reported median age of all participants in the study was 62 years, ranging from 56 to 66 years, and 45.7% of participants were male. It should be noted that previous economic studies identified in the economic review considered baseline age from as low as 50 years and the proportion male as high as 100%. These values are explored in sensitivity and scenario analyses.

The baseline prevalence and sub-distribution for Barrett's oesophagus were estimated from two systematic reviews and meta-analyses (Eusebi et al. 2021, Saha et al. 2024) and data from Cancer Research UK (2024).

The overall pooled prevalence of Barrett's oesophagus for those with gastroesophageal reflux in Europe is reported as 8.6% (95% confidence interval [CI] 5.0% to 14.4%) by Saha et al. (2024). The study also reports a pooled prevalence for Barrett's oesophagus with dysplasia and OAC of 0.6% (95% CI 0.3% to 1.1%) and 0.6% (95% CI 0.4% to 1.0%), respectively, although this is not specific to Europe. As the model is focused on the detection of Barrett's oesophagus and early-stage OAC, late-stage OAC should be removed from the overall prevalence of Barrett's oesophagus. Data from Cancer Research UK (2024) reports the 2019-2021 annual average for the proportion of oesophageal cancer cases by known stage at diagnosis for Wales. The proportion of stage I oesophageal cancer, presented in Table A6, is used to estimate prevalence and sub-distribution with OAC (late-stage) removed, where stage I is assumed to represent early-stage OAC. Additionally, the proportion of dysplasia cases which were high grade was reported as 19.3% (95% CI 8.2% to 33.7%) by Eusebi et al. (2021), and used in sub-distribution estimates. The derived values for prevalence and sub-distribution for Barrett's oesophagus, used as baseline values for the base case model, are presented in Table A7.

HTW researchers asked experts if the estimates for prevalence and sub-distribution appeared appropriate for the chronic reflux population. Several experts considered our estimates appropriate, while others reinforced the need for these estimated to align with the modelled population.

The base case analysis did not consider prevalence from the Kadri et al. (2010) study, which reported a prevalence of 3%, as the outcomes reported in Saha et al. (2024) meta-analysis are expected to be more reflective of current estimates. The discrepancy between the prevalence reported between the two studies may be due to a couple of factors. The patient population in Kadri et al. (2010) includes those who have received reflux medication for more than three months in a five year period, therefore it is possible that this group is broader than the gastroesophageal reflux population and may include patients whose reflux has resolved. Additionally, prevalence of Barrett's oesophagus may have increased since the Kadri et al. (2010) study was conducted. Notably, a study by Alexandropoulou et al. (2013) analysed UK data and concluded the incidence of Barrett's oesophagus doubled between 1996 to 2005 with the incidence of the gastroesophageal reflux remaining stable, suggesting an increase in Barrett's oesophagus prevalence over time. The reported prevalence of 3% form the Kadri et al. (2010) study has been explored in scenario analysis.

Baseline characteristic inputs are explored in sensitivity and scenario analyses.

Table A6 – Baseline characteristics

Characteristic	Mean[a]	SE, distribution	Source
Age (years)	62.0	6.2[b], normal	Kadri et al. (2010)
Male (%)	45.7	4.6[b], beta	Kadri et al. (2010)
BO prevalence (%)	8.6 (5.0 – 14.4)	2.4, beta	Saha et al. (2024)
Dysplastic BO prevalence (%)	0.6 (0.3 – 1.1)	0.2, beta	Saha et al. (2024)
OAC prevalence (%)	0.6 (0.4 – 1.0)	0.2, beta	Saha et al. (2024)
Dysplastic BO with HGD (%)	19.3 (8.2 – 33.7)	6.5, beta	Eusebi et al. (2021)
OAC with Stage I (%)	4.7	0.5[b], beta	Cancer Research UK (2024)

Abbreviations: BO, Barrett Oesophagus; HGD, high-grade dysplasia; OAC, oesophageal adenocarcinoma, SE, standard error

[a] 95% CIs displayed in brackets where reported.
[b] SE assumed 10% of the mean.

Table A7 – Derived Barrett's oesophagus prevalence and sub-distribution

Characteristic	Value (%)	Calculation
OAC (early-stage) prevalence	0.03	$OAC\ prevalence \times (OAC\ with\ Stage\ I)$
BO prevalence[a]	8.0	$BO\ prevalence - [OAC\ prevalence \times (1 - OAC\ (early\ stage)\ prevalence)]$
Dysplastic BO with LGD (%)	80.7	$1 -$ Dysplastic BO with HGD
Sub-distribution: NDBO	92.2	$\left[\begin{matrix}BO\ prevalence^a - Dysplastic\ BO\ prevalence - \\ OAC\ (early\ stage)\ prevalence\end{matrix}\right] / BO\ prevalence^a$
Sub-distribution: LGD	6.0	Dysplastic BO with LGD $\times [Dysplastic\ BO\ prevalence / BO\ prevalence^a]$
Sub-distribution: HGD	1.4	Dysplastic BO with HGD $\times [Dysplastic\ BO\ prevalence / BO\ prevalence^a]$
Sub-distribution: OAC (early-stage)	0.4	$OAC\ (early\ stage)\ prevalence / BO\ prevalence^a$

Abbreviations: BO, Barrett Oesophagus; HGD, high-grade dysplasia; LGD, low-grade dysplasia; OAC, oesophageal adenocarcinoma; SE, standard error

[a] OAC (late-stage) removed from overall BO prevalence.

2.4 Transition probabilities

The model estimates disease progression via annual transition probabilities.

Base case transition probabilities are presented in Table A10 and are aligned with values used in Swart et al. (2021), who utilises values used in previous economic models. In model calculations, transitions are applied to survivors of mortality from any cause.

Table A8 – Transition probabilities (annual)

Transition	Mean	SE[a]	Source
No BO to NDBO	0.005	0.002	Swart et al. (2021)[b]
NDBO to LGD	0.029	0.003	Swart et al. (2021)[b]
NDBO to HGD	0.005	0.001	Swart et al. (2021)[c]
NDBO to OAC (early)	0.003	0.0003	Swart et al. (2021)[c]
LGD to HGD	0.028	0.003	Swart et al. (2021)[c]
LGD to OAC (early)	0.014	0.001	Swart et al. (2021)[c]
HGD to OAC (early)	0.119	0.012	Swart et al. (2021)[b]
OAC (early) to OAC (late)	0.800	0.080	Swart et al. (2021)[d]

Abbreviations: BO, Barrett's oesophagus; HGD, high-grade dysplasia; LGD, low-grade dysplasia; NDBO, nondysplastic Barrett's oesophagus; OAC, oesophageal adenocarcinoma; RCT, randomised controlled trial; SE, standard error

[a] sampled from a beta distribution.
[b] study used value informed from Inadomi et al. (2009), adjusted by Benaglia et al. (2013).
[c] study used value informed from Inadomi et al. (2009), adjusted by Pollit et al. (2019).
[d] study used value based on expert opinion from the BEST3 RCT team.

2.5 Treatment outcomes

Treatment effects used in the base case model for endotherapy and oesophagectomy are presented in Table A11.

For patients receiving endotherapy, a proportion of patients will experience a treatment effect whereby dysplasia is completely eradicated. Additionally, patients may also experience complete eradication of any concurrent intestinal metaplasia. Previous economic models (Benaglia et al. 2013, Sami et al. 2021, Swart et al. 2021) have used two clinical studies (Phoa et al. 2014, Shaheen et al. 2011) to inform this treatment effect. Shaheen et al. (2011) reports eradication outcomes of an RCT exploring RFA in Barrett's oesophagus with dysplasia whereby for those with HGD, 89% experienced complete intestinal metaplasia eradication and 93% experienced complete dysplasia eradication after two years. Phoa et al. (2014) reports eradication outcomes of an RCT exploring RFA in Barrett's oesophagus patients with LGD whereby 88% experienced complete intestinal metaplasia eradication and 93% experienced complete dysplasia eradication at end of endotherapy treatment. It is assumed these outcomes are appropriate for use in the model at the end of the model cycle in which patients receive endotherapy treatment. Following a positive treatment effect, patients move to the no Barrett's oesophagus or NDBO health states. Patients experiencing no treatment effect are subject to natural disease progression. Aligned to the modelling approach Sami et al. (2021), it is assumed that intestinal metaplasia and dysplasia eradication are equivalent for HGD and OAC (early-stage) patients receiving endotherapy.

For OAC (late-stage) patients receiving oesophagectomy, the impact on mortality is considered. In the model by Sami et al. (2021), a study investigating the long-term survival from OAC after oesophagectomy (Ovrebo et al. 2012) was identified and used to inform mortality following oesophagectomy. Sami et al. (2021) used a five year survival following surgery of 15% based on patients with tumour stage 2. As it is unclear how Sami et al. (2021) arrived at this value, our model considers the five year survival probability of 25% based on all OAC patients receiving oesophagectomy, as reported in Ovrebo et al. (2012). Assuming exponential decay (i.e. a constant hazard rate over time), this has been translated to an annual transition probability from OAC (late-stage) to death, following oesophagectomy, for use in the economic model.

Table A9 – Treatment effect

Treatment effect	Mean (%)	SE[a] (%)	Source
Endotherapy			
LGD to No BO	0.88	0.088	Phoa et al. (2014)
LGD to NDBO	0.05	0.005	Phoa et al. (2014)
HGD/OAC (early-stage) to No BO	0.89	0.089	Shaheen et al. (2011)
HGD/OAC (early-stage) to NDBO	0.04	0.004	Shaheen et al. (2011)
Oesophagectomy			
OAC (late-stage) to death	0.24[b]	0.024	Ovrebo et al. (2012)

Abbreviations: BO, Barrett's oesophagus; HGD, high-grade dysplasia; LGD, low-grade dysplasia; NDBO, nondysplastic Barrett's oesophagus; OAC, oesophageal adenocarcinoma; SE, standard error

[a] SE assumed 10% of the mean and sampled from a beta distribution.
[b] based on a five year survival probability of 25% following oesophagectomy in OAC patients. This has been translated to an annual transition assuming exponential decay using the following formula: $p = 1 - e^{-r}$, where $r = -\frac{Ln(S)}{t}$, S is the survival probability at time t, and p is the annual transition probability.

2.6 Life expectancy and mortality

Population mortality rates for 2017-2019, by age and sex, were sourced from life tables for Wales, published by the Office for National Statistics (2024). These were used to calculate the annual probability of mortality from any cause and applied to each modelled cycle.

Those with OAC (late-stage) who are not suitable for oesophagectomy are assumed to transition to death within one cycle (i.e. one year). No increase in mortality is assumed for patients receiving endotherapy.

2.7 Resource use and costs

To reflect the perspective of the analysis, only costs that are relevant to the UK NHS were included. All costs in this analysis reflect 2023/24 prices. Unit costs and resource use inputs are presented in Table A8.

The cost of diagnostic testing with upper gastrointestinal tract endoscopy with biopsy has been informed by 2023/24 National Cost Collection data (NHS England 2024). The national average unit cost for a day case patient was reported as £745 (FE21Z). This cost is applied to all patients undergoing an endoscopy in the comparator and intervention arms.

The cost of diagnostic testing with Cytosponge has been informed by the Medtech innovation briefing (MIB240) for Cytosponge for detecting abnormal cells in the oesophagus (NICE 2020). The cost is reported as £280 which includes the cost of the device itself, the immunohistochemical assay test (TFF3), and haematoxylin and eosin stain. The base case analysis assumes the test is administered in primary care by a general practice (GP) nurse. According to the 2024 PSS Research Unit (PSSRU) report (Jones et al. 2025), the hourly cost of a qualified GP nurse is estimated as £53. In line with the economic model by Swart et al. (2021), which utilises the average Cytosponge-TFF3 test time indicated by the BEST3 RCT, 20 minutes of a GP nurse's time is considered for test administration. These costs are applied to all patients undergoing Cytosponge diagnostic testing in the intervention arm. Training costs for administering the capsule sponge device has not been incorporated into this economic evaluation as the cost is expected to be negligible on a per-patient basis.

Safety and adverse events related to Cytosponge testing have been reported in Section 5.5 of the main report and comparative outcomes were not identified. However, as the ability to swallow the capsule sponge and sponge detachment would not apply to the comparator, considerations for these events have been made in the intervention arm. In a retrospective study performing a patient-level review of five prospective trials assessing Cytosponge (Januszewicz et al. 2019), they found that 3.5% of patients failed to swallow the capsule sponge. Therefore, it is assumed patients failing to swallow the capsule sponge would receive an endoscopic biopsy. Then in the BEST3 RCT (Fitzgerald et al. 2020), the capsule sponge detached in 0.06% of patients, leading to endoscopic retrieval. A cost of £1,941 is applied to these patients based on non-elective therapeutic upper gastrointestinal tract endoscopy (FE20Z) from 2023/24 National Cost Collection data (NHS England 2024). As these patients all receive one capsule sponge device without the immunohistochemical assay test (TFF3), and haematoxylin and eosin stain, a cost of [redacted] for only the capsule sponge device itself is applied. This cost was provided by Medtronic, the manufacture of the Cytosponge device.

Patients diagnosed with LGD, HGD or early-stage OAC were treated with a regimen of endotherapy aligned to NICE guidelines (NG231) (NICE 2023a). The treatment regimen is assumed to occur within one annual cycle. For those diagnosed with LGD, patients receive an additional endoscopic biopsy followed by radiofrequency ablation (RFA). For those diagnosed with HGD or early-stage

OAC, patients receive endoscopic mucosal resection (EMR) followed by RFA, where it is assumed all patients have residual Barrett's oesophagus following EMR. It is assumed early-stage OAC patients do not require oesophagectomy. A cost of £1,220 is considered for an EMR procedure, based on the national average day case unit cost for major therapeutic endoscopic procedures of the upper or lower gastrointestinal tract (FE02A – FE02C) (NHS England 2024). For an RFA procedure, a cost of £1,493 is considered based on the national average day case unit cost for complex therapeutic endoscopic procedures of the upper or lower gastrointestinal tract (FE01Z) (NHS England 2024). It was deemed appropriate to classify RFA as a complex procedure, as one expert contacted by HTW researchers noted that the cost of a BARRX RFA catheter alone is £1,200, suggesting an overall cost higher than the national average for major procedures. In a RCT exploring RFA in Barrett's oesophagus with dysplasia (Shaheen et al. 2009), a mean of 3.5 treatments were performed per patient. Therefore, patients treated with endotherapy are assumed of receive 3.5 RFA sessions within their regimen. For patients treated with EMR, it is assumed only one EMR session is received within their regimen.

Those diagnosed with Barrett's oesophagus or progressed up to early-stage OAC receive endoscopic surveillance. In NICE guidelines (NG231) (NICE 2023a), it is recommended for patients diagnosed with Barrett's oesophagus to undergo an endoscopic surveillance every two to five years, dependant on segment length and intestinal metaplasia. For the base case model, it is assumed endoscopic surveillance is undertaken every three years. The frequency of endoscopic surveillance is explored in scenario and sensitivity analyses.

For patients receiving endotherapy, additional short-term surveillance is incorporated based on expert feedback received during the review process. One expert indicated that endoscopic eradication therapy is labour intensive and will generally involve endoscopies every three months for the first year, and then every six months for the second year. Therefore, it is assumed that patients receive four endoscopic biopsies in the year of treatment and two endoscopic biopsies in the subsequent year, applied in place of the standard rate of surveillance. Following the additional short-term surveillance, patients revert to the standard surveillance schedule.

Patients progressing to late-stage OAC are treated with oesophagectomy or palliative cancer treatments depending on if they are suitable for surgery. The base case model considers 50% of late-stage OAC patients are unsuitable for surgery. This is aligned to previous modelling work by Sami et al. (2021) and historic modelling by NICE (CG106, now obsolete (NICE 2010, cited in Sami et al. 2021)) using estimates from a study conducting a systematic review, expert workshop and economic modelling for the surveillance of Barrett's oesophagus (Garside et al. 2006) based on those with symptomatic cancer. For oesophagectomy, a cost of £11,224 is considered based on the national average unit cost for elective inpatient procedures involving complex, oesophageal, stomach or duodenum treatments (FF02A - FF02C) (NHS England 2024). Patients in the OAC (late-stage) health state suitable for surgery are also assumed to receive two outpatient visits per year, aligned to considerations made in Benaglia et al. (2013). A cost of £181 is considered for an outpatient visit, based on the national average unit cost for a consultant led, non-admitted face-to-face follow-up attendance in the upper gastrointestinal surgery service (WF01A) (NHS England 2024).

For palliative cancer treatments in those unsuitable for surgery, it is assumed 25% receive chemotherapy and 75% receive palliative RFA and stent, aligned to treatments received by stage III/IV patients in the model by Swart et al. (2021) based on BEST3 RCT data. The cost of chemotherapy and palliative RFA has been lifted from Swart et al. (2021) and inflated from 2018/19 (assumed) to 2023/24 values. A cost of £4,808 is considered for palliative stent, based on the national average unit cost for elective inpatients receiving endoscopic insertion of luminal stent into the gastrointestinal tract (FE10A-FE10D) (NHS England 2024).

Patients with late-stage OAC who are not suitable for oesophagectomy are assumed to transition to death within one year. Palliative care costs are applied to late-stage OAC patients transitioning to death, excluding those who reach 100 years of age expected to die of natural causes. The cost of palliative care has been lifted from the Swart et al. (2021) and inflated from 2018/19 (assumed) to 2023/24 values.

Table A10 – Unit costs and resource use inputs

Input	Mean	SE[a], distribution	Source
Unit costs (£)			
Endoscopic biopsy	745.00	74.50, gamma	Day case (FE21Z) NHS England (2024)
Cytosponge testing (overall)	280.00	28.00, gamma	NICE (2020)
Cytosponge device only	[redacted]	[redacted], gamma	Medtronic
Qualified GP nurse (per hour)	53.00	5.30, gamma	Jones et al. (2025)
Endoscopic retrieval	1,941.00	194.10, gamma	Non-elective (FE20Z) NHS England (2024)
Endotherapy – EMR	1,219.31	121.93, gamma	Day case (FE02A – FE02C) NHS England (2024)
Endotherapy - RFA	1492.51	149.2, gamma	Day case (FE01Z) NHS England (2024)
Oesophagectomy	6,362.86	636.29, gamma	Elective (FF02A - FF02C) NHS England (2024)
Oesophagectomy follow up outpatient visit	181.00	18.10, gamma	Consultant led (WF01A) NHS England (2024)
Chemotherapy (1-year)	3,965.78	396.58, gamma	Swart et al. (2021) using NICE NG83 (NICE 2018)
Palliative RFA (1-year)	4,505.53	450.55, gamma	Swart et al. (2021) using NICE NG83 (NICE 2018)
Palliative stent (1-year)	4,807.78	480.78, gamma	Elective (FE10A – FE10D) NHS England (2024)
Palliative care (final year of life)	8,759.82	875.98, gamma	Swart et al. (2021) using NICE NG83 (NICE 2018)
Resource use			
Cytosponge administration time (minutes)	20.0	2.0, gamma	Swart et al. (2021) using BEST3 RCT data
Failed to swallow Cytosponge device (%)	3.5	0.35, beta	Januszewicz et al. (2019)
Cytosponge detachment (%)	0.06	0.006, beta	BEST3 RCT (Fitzgerald et al. 2020)
Endotherapy for LGD: Additional endoscopy (n)	1.0	Fixed	NICE (2023a)
Endotherapy for LGD: RFA sessions (n)	3.5	Fixed	Shaheen et al. (2009)

Input	Mean	SE[a], distribution	Source
Endotherapy for HDG/OAC: EMR sessions (n)	1.0	Fixed	Assumption
Endotherapy for HDG/OAC: RFA sessions (n)	3.5	Fixed	Shaheen et al. (2009)
Endoscopies following endotherapy year 1 (n)	4.0	Fixed	Expert opinion
Endoscopies following endotherapy year 2 (n)	2.0	Fixed	Expert opinion
Endoscopic surveillance frequency (years)	3.0	Fixed	Assumption based on NICE NG231 NICE (2023a)
OAC (late-stage) treatment: Suitable for surgery (%)	50.0	5.0, beta	Assumption aligned to Sami et al. (2021)
Post surgery follow up appointments per year (n)	2.0	Fixed	Assumption aligned to (Benaglia et al. 2013)
OAC (late-stage) treatment: Unsuitable for surgery and receiving palliative RFA and stent (%)	75.0	7.5, beta	Swart et al. (2021) using BEST3 RCT data
OAC (late-stage) treatment: Unsuitable for surgery and receiving chemotherapy (%)	25.0	-	Calculated[b]

Abbreviations: EMR, endoscopic mucosal resection; GP, general practice; HGD, high-grade dysplasia; LGD, low-grade dysplasia; OAC, oesophageal adenocarcinoma; RFA, radiofrequency ablation; SE, standard error

[a] SE assumed 10% of the mean.
[b] Calculated as: 1 − OAC (late − stage) treatment: Unsuitable for surgery and receiving palliative RFA and stent

2.8 Quality of life

The model estimates effectiveness in terms of quality adjusted life years (QALYs). These are estimated by combining life year estimates with QoL utility values associated with being in a particular health state. QoL utility values used in the model are presented in Table A9 which are closely aligned with values used in previous economic models (Benaglia et al. 2013, Sami et al. 2021, Swart et al. 2021) and in historic modelling by NICE (NICE CG106, now obsolete (NICE 2010, cited in Sami et al. 2021)).

To capture the impact of the aging population, the model incorporates general population age-adjusted QoL utility values, sourced from the NICE Decision Support Unit (Hernández Alava et al. 2022). These are applied to health state utilities multiplicatively throughout the modelled time horizon.

Disutility associated with endoscopy and capsule sponge was not captured due to the lack of comparable data.

Table A11 – Quality of life health state utility values

Health state	Mean	SE[a]	Source
No BO	1.000	Fixed	Benaglia et al. (2013), Sami et al. (2021)
NDBO	0.910	0.130	Benaglia et al. (2013), Sami et al. (2021)
LGD	0.850	0.120	Benaglia et al. (2013), Sami et al. (2021)
LGD (during endotherapy)	0.770	0.140	Benaglia et al. (2013)
HGD	0.770	0.140	Benaglia et al. (2013), Sami et al. (2021)
OAC (early-stage)	0.770	0.140	Benaglia et al. (2013), Sami et al. (2021)
OAC (late-stage)	0.675	0.032	Benaglia et al. (2013), Sami et al. (2021)
OAC (late-stage): due to surgery	0.414	0.19	Benaglia et al. (2013)
OAC (late-stage): during the year of surgery	0.731	-	Calculated[b]
OAC (late-stage): years following surgery	0.836	0.016	Benaglia et al. (2013), Sami et al. (2021)

Abbreviations: BO, Barrett's oesophagus; EMR, endoscopic mucosal resection; HGD, high-grade dysplasia; LGD, low-grade dysplasia; NDBO, nondysplastic Barrett's oesophagus; OAC, oesophageal adenocarcinoma; RFA, radiofrequency ablation; SE, standard error

[a] sampled from a beta distribution unless fixed.
[b] the utility value for OAC (late-stage) during the year of surgery is calculated in the model where the utility due to surgery (0.414) is assumed to last for a duration of 3 months until reaching the utility following surgery (0.836). Therefore, the adjusted value applied to the year during surgery is 0.731 based on a weighted average. The assumption for the utility due to surgery (0.414) lasting a duration of 3 months is aligned a modelling assumption based on expert opinion in Swart et al. (2021).

3. Results

3.1 Base case results

Base case health economic results are provided in Table A12. Over a lifetime horizon, the results show that use of Cytosponge in primary care, followed by endoscopic biopsy in those with a positive result, is expected to reduce costs by [redacted] per patient with a loss of 0.02 QALYs, compared to endoscopic biopsy in all patients. This results in a corresponding incremental cost-effectiveness ratio (ICER) of [redacted], representing the cost savings per QALY lost. Therefore, Cytosponge is estimated to be cost effective at the commonly accepted cost effectiveness threshold of £20,000 per QALY.

In this context, where the intervention is less costly and less effective than the comparator, an ICER above the commonly accepted cost effectiveness threshold of £20,000 per QALY is considered cost effective as cost savings outweigh the reduction in health outcomes. All ICERs in this evaluation should be interpreted using this framework.

Table A12 – Base case health economic results (per patient)

	Intervention	Comparator	Incremental
Total Costs	[redacted]	£1,403	[redacted]
Total QALYs	11.74	11.76	-0.02
ICER (cost savings per QALY lost)			[redacted]
Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year			

A breakdown of initial diagnostic outcomes, per 1,000 patients, are presented in Table A13. In overall diagnostic outcomes, the intervention arm is estimated to result in 73.8% of the Barrett's oesophagus population being detected and appropriately managed. In the comparator arm, 100% of the population is detected and appropriately managed due to the assumption of perfect diagnostic accuracy.

When considering outcomes specific to the intervention arm, 3.6% of the population are estimated to receive an endoscopy due to sponge detachment or failing to swallow the capsule sponge. In those with a positive capsule sponge result, 49.3% of patients receive an unnecessary endoscopy.

Table A13 – Initial diagnostic outcomes (base case, per 1,000 patients)

	Intervention	Comparator
Capsule sponge device outcomes		
Failed capsule sponge swallow	35 (No BO = 32, BO = 3)	NA
Capsule sponge detachment	1 (No BO = 1, BO = 0)	NA
Positive capsule sponge result	112 (TN = 55, TP = 57)	NA
Overall diagnostic outcomes		
No BO detected	920	920
BO detected and confirmed	59	80
BO undetected	21	0
Abbreviations: BO, Barrett's oesophagus; NA, not applicable; TN, true negative; TP, true positive		

The per patient QALY and cost breakdown of the base case analysis are summarised in Table A14 and Table A15, respectively.

The overall QALYs in the intervention arm is estimated to be less than the comparator arm. When QALYs are disaggregated by health state, QALYs are higher in progressed health states (LGD, HGD and OAC) for the intervention arm, reflecting the additional patients progressing to these states. Conversely, early health states (No Barrett's oesophagus and NDBO) show fewer QALYs for the intervention arm as fewer people remain healthy.

A breakdown of costs shows estimated savings of [redacted] in diagnostic and surveillance costs, primarily driven by fewer patients requiring an endoscopic biopsy in the intervention arm. An additional estimated saving of [redacted] in endotherapy costs is attributed to fewer Barrett's oesophagus patients detected, and therefore subsequently treated. This reduction in detections results in an increase in progression to late-stage OAC, and therefore an increase of [redacted] in related treatment costs.

Table A14 – QALY breakdown (base case, per patient)

	Intervention	Comparator	Incremental
Health states			
No BO	10.613	10.670	-0.057
NDBO	0.950	0.954	-0.004
LGD	0.131	0.102	0.029
HGD	0.028	0.021	0.007
Early-stage OAC	0.008	0.006	0.002
Late-stage OAC	0.012	0.009	0.003
Total			
Total	11.742	11.762	-0.020

Abbreviations: BO, Barrett's oesophagus; EMR, endoscopic mucosal resection; HGD, high-grade dysplasia; LGD, low-grade dysplasia; NDBO, nondysplastic Barrett's oesophagus; OAC, oesophageal adenocarcinoma

Table A15 – Cost breakdown (base case, per patient)

	Intervention	Comparator	Incremental
Diagnostics and surveillance			
Capsule sponge testing	[redacted]	NA	[redacted]
Capsule sponge detachment	[redacted]	NA	[redacted]
Endoscopic biopsy	[redacted]	[redacted]	[redacted]
Endoscopic surveillance	[redacted]	[redacted]	[redacted]
Endotherapy[a]			
LGD	[redacted]	[redacted]	[redacted]
HGD	[redacted]	[redacted]	[redacted]
Early-stage OAC	[redacted]	[redacted]	[redacted]
Late-stage OAC			
Surgery	[redacted]	[redacted]	[redacted]
Post surgery follow ups	[redacted]	[redacted]	[redacted]
No surgery	[redacted]	[redacted]	[redacted]
OAC death	[redacted]	[redacted]	[redacted]
Total			
Total	[redacted]	£1,403.16	[redacted]

Abbreviations: HGD, high-grade dysplasia; LGD, low-grade dysplasia; NA, not applicable; OAC, oesophageal adenocarcinoma

[a] Includes short-term surveillance associated with endotherapy

3.2 Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was conducted to assess the combined parameter uncertainty in the model. In this analysis, inputs used in the base case were replaced with values drawn from distributions around the mean. The health economic outcomes of this analysis were summarised as the mean average of PSA runs. To ensure robustness, the PSA was run 10,000 times.

Table A16 presents the health economic results summarised from the PSA. Under this analysis, the intervention is cost effective with an ICER of [redacted] per QALY. At a cost-effectiveness threshold of £20,000, 65.8% of PSA estimates were cost effective.

The cost-effectiveness plane, presented in Figure 5, provides a detailed visualisation of the incremental costs and QALYs associated with each PSA estimate, along with the mean result. The £20,000 threshold considered for cost effectiveness is also presented (dotted line), where PSA estimates below this line are cost effective.

Figure 6 presents the probability of the intervention being considered cost effective at various cost-effectiveness thresholds. At a threshold of £3,000, the probability of cost effectiveness reaches 100.0%.

Table A16 – Probabilistic sensitivity analysis results (per patient)

Incremental costs	Incremental QALYs	ICER	% cost effective[a]
[redacted]	-0.02	[redacted]	65.8%

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

[a] The percentage of PSA estimates which are cost effective at a threshold of £20,000 per QALY

Figure 5 – Cost-effectiveness plane

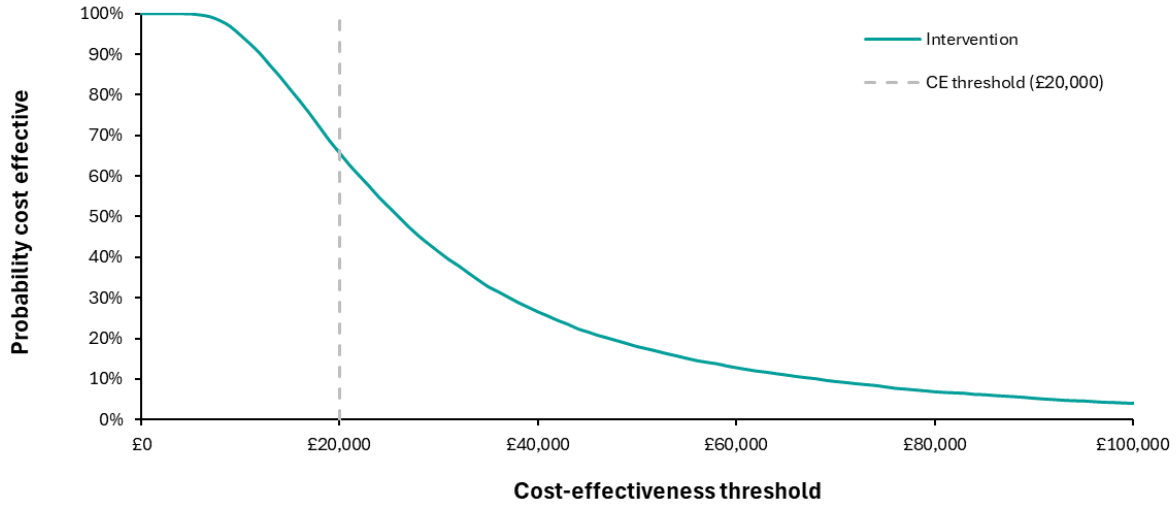


Figure 6 – Cost-effectiveness acceptability curve

3.3 Deterministic sensitivity analysis

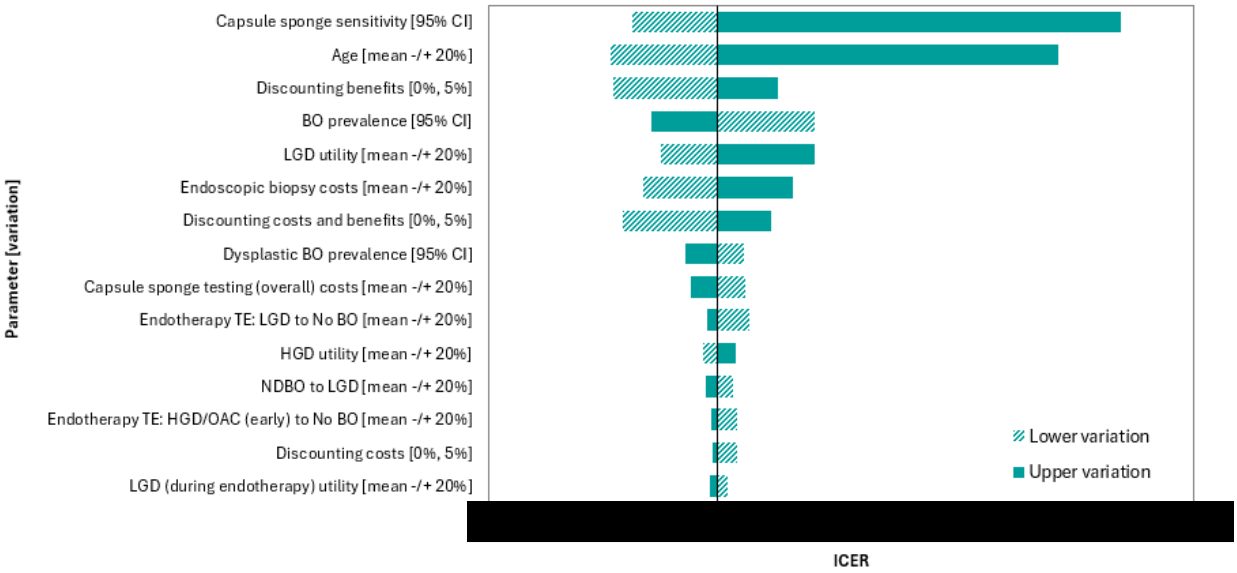
Deterministic sensitivity analyses were conducted by varying one input parameter at a time, running the model, and recording the health economic outcomes. This approach helps estimate uncertainty and identify the key drivers of the model results. Where available, inputs were varied within 95% confidence intervals (CIs); otherwise, input parameters were adjusted by 20% above and below the mean value. The only exception was discounting costs and benefits (both together and separately) which was varied between 0% and 5%.

The 15 most influential parameters on the ICER in deterministic sensitivity analysis are presented in Figure 6.

This analysis revealed Cytosponge sensitivity, age and Barrett's oesophagus prevalence to be key drivers on health economic outcomes, with the ICER ranging from [redacted]. Additionally, varying the discount rate applied to benefits was shown to be impactful on health economic outcomes, with the ICER ranging from [redacted].

The most influential parameter was capsule sponge sensitivity, with higher sensitivity levels resulting in higher ICERs, due to fewer Barrett's oesophagus patients being undetected. An older population was also associated with higher ICERs, as the shorter remaining lifetime reduces the opportunity for patients with undetected Barrett's oesophagus to progress to more severe health states. Furthermore, lower prevalence of Barrett's oesophagus increases the ICER due to fewer patients being at risk of progressing to advanced disease.

At a cost effectiveness threshold of £20,000, the following input parameters resulted in health economic results which were not cost effective when their lower variation was applied: capsule sponge sensitivity, age, discounting benefits, LGD utility, endoscopic biopsy costs, and discounting costs and benefits together. When their upper variation was applied, only Barrett's oesophagus prevalence and dysplastic Barrett's oesophagus prevalence resulted in health economic results which were not cost effective.



Abbreviations: BO, Barrett's oesophagus; CE, cost-effectiveness; HGD, high-grade dysplasia; ICER, incremental cost-effectiveness ratio; LGD, low-grade dysplasia; NDBO, nondysplastic Barrett's oesophagus; OAC, oesophageal adenocarcinoma; TE, treatment effect

Figure 7 – Deterministic sensitivity analysis, 15 most influential parameters on the ICER

3.4 Threshold analysis

As demonstrated in the deterministic sensitivity analysis (Section 3.3), Cytosponge sensitivity, age and Barrett's oesophagus prevalence were identified as key influential parameters on health economic outcomes. To explore this, a threshold analysis was conducted to determine the value of these parameters at which the intervention becomes cost effective, at a cost effectiveness threshold of £20,000. The health economic impact of varying these parameters is presented in Figure 8, Figure 9 and Figure 10, respectively.

For capsule sponge sensitivity, the analysis revealed that a minimum sensitivity of [redacted] is required for the intervention to achieve cost effectiveness. Below this level, the reduction health outcomes outweigh the associated cost savings.

For age, the analysis indicated that the intervention is cost effective for patients aged [redacted] years and over. At younger ages, the longer time horizon increases the opportunity for those with undetected Barrett's oesophagus to progress to more severe health states, leading to greater losses in QALYs. Therefore, older populations are more favourable for cost effectiveness.

In contrast, for Barrett's oesophagus prevalence, the analysis demonstrated that the intervention remains cost effective if the prevalence is below [redacted]. At higher prevalence levels, a greater proportion of patients may benefit from early detection and treatment, increasing the value of the comparator. This suggests that the cost effectiveness of the intervention is more favourable in lower prevalence populations.

Figure 8 – Threshold analysis, capsule sponge sensitivity

Figure 9 – Threshold analysis, age

Figure 10 – Threshold analysis, Barrett's oesophagus prevalence

3.5 Scenario analysis

Scenario analyses were conducted to test the robustness of the base case results to alternative modelling assumptions. A total of 15 scenarios were explored, each representing a plausible variation in key model parameters, structural assumptions or care setting where the capsule sponge device is administered. Full details of each scenario are provided in Table A17, and corresponding results are presented in Table A18. All results determine cost effectiveness at the commonly accepted threshold of £20,000 per QALY.

A scenario considering insourcing costs for endoscopic biopsy was not included due to the lack of a robust cost estimate. However, if insourced procedures are expected to be more costly than in-house procedures, this would increase the cost of the comparator. As a result, cost effectiveness conclusions would remain unchanged from the base case analysis.

Scenario analyses explored the impact of alternative care settings from the base case, which considered the capsule sponge device administered by a qualified GP nurse in primary care. A secondary care scenario applied the cost of a hospital-based nurse specialist from the 2024 PSSRU report (Jones et al. 2025), including qualifications and overheads. Two community-based settings were also considered using similar unit costs: one assuming administration by a Band 6 qualified nurse, and another by a community pharmacist, reflecting expert feedback that future delivery may move to community settings, including pharmacies. Across these scenarios, changes in incremental costs were minimal, with no impact on QALYs, suggesting that the care setting has limited influence of cost effectiveness results.

Two scenarios explored the use Endosign as an alternative to Cytosponge. As experts indicted that evidence between these two devices and associated biomarkers can be generalised, the diagnostic accuracy and adverse events of Endosign were assumed equivalent to Cytosponge. The first scenario replaced Cytosponge costs with the standard cost of Endosign. The second scenario accounted for a national discount structure for Endosign [redacted] The cost associated with Endosign have been provided by the manufacture. Results showed changes in incremental costs were minimal, with no impact on QALYs, suggesting the use Endosign is unlikely to affect the overall cost effectiveness conclusions.

The cost applied to patients receiving an endoscopy was also explored. This scenario applied a weighted average of diagnostic testing with upper gastrointestinal tract endoscopy with and without biopsy, informed by 2023/24 National Cost Collection data (NHS England 2024). This cost (£713) replaces the cost of endoscopic biopsy (£745) used in the base case analysis. While endoscopic biopsy is the reference standard in the diagnostic accuracy studies used, this scenario aims to reflect that all patients undergoing endoscopy may not receive a biopsy in clinical practice. Results of this scenario showed the intervention remained cost effective, with reduced cost savings compared to the base case.

Several scenarios explored alternative population assumptions. The first scenario adjusted age and sex inputs to match the study used to inform Barrett's oesophagus prevalence in the base case (Saha et al. 2024), although these were not Europe specific. This led to a population younger than the base case, resulting in a lower ICER which was not cost effective. The second scenario applied the lower Barrett's oesophagus prevalence reported in the diagnostic study by Kadri et al. (2010) (3.0%) instead of the base case estimate (8.6%) drawn from the more recent meta-analysis (Saha et al. 2024). Here, the intervention remained cost effective with a stronger ICER of [redacted]. The third scenario modelled a male only population, consistent with previous economic evaluations (see Section 6.1 of the main report). This produced similar results to the base case, with a slightly improved ICER and no change in cost effectiveness conclusions.

A scenario considering an alternative proportion of capsule sponge detachments was explored based on a prospective cohort study (Chien et al. 2024a) reporting two sponge detachments out of 1,385 Cytosponge tests. This scenario resulted in minimal impact on health economic outcomes and did not change cost effectiveness conclusions.

Another scenario explores alternative diagnostic accuracy outcomes from Kadri et al. (2010) using a segment length of 2 cm or more. The higher sensitivity applied in this scenario resulted in more patients with Barrett's oesophagus being identified, which reduced QALY losses and produced a stronger cost effective ICER of [redacted]

Two scenarios explored variations in utility inputs. One removed age-adjusted utilities, applying fixed health state utility values across the lifetime horizon. This increased QALY losses and produced an ICER which was not cost effective. A second scenario assumed the utility for late-stage OAC following surgery to be equal to early-stage OAC, addressing the uncertainty around this input. This scenario resulted in minimal impact on health economic outcomes and did not change cost effectiveness conclusions.

Finally, two scenarios tested structural assumptions in the model. A shorter time horizon of 5-years estimated less QALY losses due to the reduced opportunity for patients to progress to advanced disease health states. This resulted in a very strong cost effective ICER of [redacted]. Another scenario removed standard endoscopic surveillance, assumed to be every three years in patients diagnosed with Barrett's oesophagus in the base case. Between each arm, this scenario led to fewer patients transitioning to post-treatment health states, resulting in reduced QALY losses. Whilst cost effectiveness conclusions did not change, this scenario produced a stronger cost effective ICER of [redacted].

Table A17 – Scenario analyses description and rationale

Scenario	Title	Description/rationale
Scenario 1	Capsule sponge in the secondary care setting	Considers the capsule sponge device administered in secondary care by a hospital-based nurse specialist (£62 per hour, including qualifications). Costs from Jones et al. (2025)
Scenario 2	Capsule sponge in community setting (qualified nurse)	Considers the capsule sponge device administered in the community setting by a qualified nurse (assumes band 6 - £64 per hour, including qualifications). Costs from Jones et al. (2025)
Scenario 3	Capsule sponge in community setting (pharmacist)	Considers the capsule sponge device administered in the community setting by a community pharmacist (£57 per hour). Costs from Jones et al. (2025)
Scenario 4	Endosign costs	Replaces Cytosponge costs with Endosign costs [redacted] Assumes equivalent diagnostic accuracy to Cytosponge. Costs from manufacture.
Scenario 5	Endosign costs (with discount)	As per scenario 4 with Endosign costs accounting for a national discount structure [redacted] Costs from manufacture.
Scenario 6	Weighted endoscopy costs (with and without biopsy)	Replaces endoscopic biopsy costs with a weighted average of daycase costs for diagnostic endoscopy with (FE21Z) and without (FE22Z) biopsy (£713). Costs from NHS England (2024).
Scenario 7	Age and sex based on Saha et al. (2024)	Aligns age and sex to the study (Saha et al. 2024) used to inform BO prevalence in the base case analysis (55.5 years of age and 52.6% male).
Scenario 8	Prevalence based on Kadri et al. (2010)	Aligns prevalence to the study (Kadri et al. 2010) used to inform diagnostic accuracy in the base case analysis (BO prevalence of 3%).
Scenario 9	Male only population	Explores the impact of setting the population to 100% male, aligning to previous economic studies identified in the economic review (Section 6.1 of the main report).
Scenario 10	Sponge detachment from Chien et al. (2024a)	Explores an alternative study (Chien et al. 2024a) reporting the proportion of sponge detachments (0.14% sponge detachments).
Scenario 11	2cm cut-off segment (diagnostic threshold)	Explores an alternative diagnostic threshold (a segment of 2cm or more) to define BO from Kadri et al. (2010) (90.0% sensitivity, 93.5% specificity).
Scenario 12	Utility for late-stage OAC following surgery equal to early-stage OAC	Explores the uncertainty of the utility assigned to late-stage OAC following surgery and assumes this is equal to early-stage OAC (0.77).
Scenario 13	No age-adjusted utilities included	Explores the impact of removing age-related utility decline.
Scenario 14	5-year time horizon	Replaces a lifetime horizon with a 5-year horizon to explore short-term cost effectiveness where long-term disease progression may be less influential.
Scenario 15	No standard endoscopic surveillance	Explores the impact of removing standard endoscopic surveillance.

Abbreviations: BO, Barrett's oesophagus; OAC, oesophageal adenocarcinoma

Table A18 – Scenario analyses results

Scenario	Title	Inc. costs	Inc. QALYs	ICER	CE outcome[a]
Scenario 1	Capsule sponge in the secondary care setting	[redacted]	-0.020	[redacted]	Cost effective
Scenario 2	Capsule sponge in community setting (qualified nurse)	[redacted]	-0.020	[redacted]	Cost effective
Scenario 3	Capsule sponge in community setting (pharmacist)	[redacted]	-0.020	[redacted]	Cost effective
Scenario 4	Endosign costs	[redacted]	-0.020	[redacted]	Cost effective
Scenario 5	Endosign costs (with discount)	[redacted]	-0.020	[redacted]	Cost effective
Scenario 6	Weighted endoscopy costs (with and without biopsy)	[redacted]	-0.020	[redacted]	Cost effective
Scenario 7	Age and sex based on Saha et al. (2024)	[redacted]	-0.029	[redacted]	Not cost effective
Scenario 8	Prevalence based on Kadri et al. (2010)	[redacted]	-0.009	[redacted]	Cost effective
Scenario 9	Male only population	[redacted]	-0.019	[redacted]	Cost effective
Scenario 10	Sponge detachment from Chien et al. (2024a)	[redacted]	-0.020	[redacted]	Cost effective
Scenario 11	2cm cut-off segment (diagnostic threshold)	[redacted]	-0.007	[redacted]	Cost effective
Scenario 12	Utility for late-stage OAC following surgery equal to early-stage OAC	[redacted]	-0.020	[redacted]	Cost effective
Scenario 13	No age-adjusted utilities included	[redacted]	-0.026	[redacted]	Not cost effective
Scenario 14	5-year time horizon	[redacted]	-0.001	[redacted]	Cost effective
Scenario 15	No standard endoscopic surveillance	[redacted]	-0.006	[redacted]	Cost effective

Abbreviations: CE: cost effectiveness; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; OAC, oesophageal adenocarcinoma

Key: green = cost effective, red = not cost effective

[a] Cost effectiveness is determined at a commonly accepted threshold of £20,000 per QALY.

4. Model assumptions and limitations

A summary of key assumptions and limitations of the base case analysis are presented in Table A19.

Table A19 – Key model assumptions and limitations

Model component	Description
Assumptions	
Capsule sponge delivery care setting	For the base case, it is assumed the capsule sponge device is administered by a qualified GP nurse in primary care.
Capsule sponge device	The Cytosponge device was considered for the base case to reflect the evidence base.
Capsule sponge diagnostic accuracy	Diagnostic accuracy is based on the use of TFF3 biomarker testing and a cut-off segment length of 1 cm or more. As described in Kadri et al. (2010).
Endoscopic biopsy accuracy	Endoscopic biopsy is assumed to be perfectly accurate with a sensitivity and specificity of 100%. Aligned to the reference standard in Kadri et al. (2010).
Standard endoscopic surveillance	Those diagnosed with Barrett's oesophagus receive endoscopic surveillance every three years. Endoscopic surveillance stops if patients progress to late-stage OAC.
PPI therapy costs	Treatment costs for PPI therapy are not considered as it is assumed all patients receive this due to their underlying chronic reflux.
Training costs	Training costs for administering the capsule sponge device has not been incorporated into this economic evaluation as the cost is expected to be negligible on a per-patient basis.
Capsule sponge fail	It is assumed those who are unable to swallow the capsule sponge device or experience detachment receive an endoscopic biopsy in secondary care.
Endotherapy	Patients in LGD, HGD and early-stage OAC health states are assumed to only have a single instance of endotherapy treatment regimens. Therefore, patients re-entering these health states following improvement from initial endotherapy would not undergo subsequent endotherapy treatment. This is aligned to previous economic evaluations.
Endotherapy	Patients identified with LGD, HGD, or early-stage OAC through surveillance, who have not previously undergone endotherapy treatment, proceed to receive endotherapy.
Endotherapy	Patients undergoing endotherapy receive four endoscopies in the year of treatment and two endoscopies in the subsequent year, applied in place of the standard rate of surveillance. Following this short-term surveillance, patients revert to the standard surveillance schedule.
Endotherapy	It is assumed that endotherapy treatment outcomes are equivalent for HGD and early-stage OAC patients.
Oesophagectomy	It is assumed early-stage OAC patients do not require oesophagectomy.
Early-stage OAC	Stage I oesophageal cancer is assumed to represent early-stage OAC.
Late-stage OAC	Progression to late-stage OAC is assumed to directly lead to clinical intervention due to the presence of symptoms.
Late-stage OAC	Patients progressing to late-stage OAC are treated with oesophagectomy or palliative cancer treatments, aligned to previous economic evaluations.
Late-stage OAC	Patients in the late-stage OAC health state suitable for surgery are assumed to receive two outpatient visits per year, aligned to previous economic evaluations.

Model component	Description
Late-stage OAC	For palliative cancer treatments in those unsuitable for surgery, it is assumed 25% receive chemotherapy and 75% receive palliative RFA and stent, aligned to treatments received by stage III/IV patients in the model by Swart et al. (2021) based on BEST3 RCT data.
Late-stage OAC mortality	Those who enter the late-stage OAC health state who are not suitable for oesophagectomy are assumed to transition to death in the subsequent model cycle, closely aligning to previous economic evaluations.
Late-stage OAC mortality	Deaths in the late-stage OAC health state are assumed to be OAC-related death. Excluding those who die of natural causes from reaching 100 years of age.
Mortality	Patients are assumed to die of natural causes once the cohort reaches age 100.
Mortality	In all health states up to OAC (early-stage), patients are at risk of mortality from any cause throughout the time horizon.
Model transitions	In model calculations, transitions are applied to survivors of mortality from any cause.
Limitations	
Modelled population	A separate evaluation for the surveillance population was not conducted due to the need for additional assumptions around surveillance intervals, disease progression risks, and repeat test performance. This led to a focus on the chronic reflux population for this evaluation, where available disease progression models are more established.
Disutility	A disutility associated with endoscopy and capsule sponge was not incorporated due to the absence of comparative evidence. The model may underestimate the full value of the intervention if endoscopy is associated with a temporary reduction in quality of life not experienced by capsule sponge.
Diagnostic accuracy population	In the study used to inform the accuracy of capsule sponge testing (Kadri et al. 2010), the patient population includes those who have received reflux medication for more than three months in a five year period, and it is possible that this group is broader than our target population and may include patients whose reflux has resolved.
Diagnostic accuracy	The study used to inform the accuracy of capsule sponge testing (Kadri et al. 2010) was undertaken in 2008 – 2009, and it's outcomes may not accurately reflect current practices.
Diagnostic accuracy	Diagnostic accuracy values are applied uniformly across all severities of Barrett's oesophagus. This may limit the applicability of result in clinical practice where detection may be more likely in patients with more advanced disease.
Prevalence	The base case did not consider prevalence from the Kadri et al. (2010) study, which reported a prevalence of 3%, as the outcomes reported in the Saha et al. (2024) meta-analysis are expected to be more reflective of current estimates.
Endoscopic biopsy accuracy	Experts contacted by HTW noted that endoscopic biopsy is not perfectly accurate in reality. However, this assumption was necessary to remain consistent with the evidence base and to allow a relative comparison to test performance.

Abbreviations: GP, general practice; HGD, high-grade dysplasia; LGD, low-grade dysplasia; OAC, oesophageal adenocarcinoma; PPI, proton pump inhibitor; RCT, randomised controlled trial; TFF3, trefoil factor 3