Catheter-free (wireless) ambulatory oesophageal pH monitoring for gastro-oesophageal reflux disease (GORD)

February 2014

MSAC application no. 1203

Assessment report

Electronic copies of the report can be obtained from the Medical Service Advisory Committee's internet site at http://www.msac.gov.au/

Enquiries about the content of the report should be directed to the above address.

This report was commissioned for use by the Medical Services Advisory Committee (MSAC) to inform its deliberations. MSAC is an independent committee that has been established to provide advice to the Minister for Health on the strength of evidence available on new and existing medical technologies and procedures in terms of their safety, effectiveness and cost-effectiveness. This advice will help to inform government decisions about which medical services should attract funding under Medicare.

MSAC's advice does not necessarily reflect the views of all individuals who participated in the MSAC evaluation.

This report should be referenced as follows:

Kessels S, Schubert C, Newton S & Merlin T 2014, *Catheter-free (wireless) ambulatory oesophageal pH monitoring for gastro-oesophageal reflux disease (GORD)*, MSAC Application 1203, Assessment Report, Commonwealth of Australia, Canberra, ACT.

Contents

Tables

Boxes

Figures

Executive summary

Assessment of catheter-free (wireless) ambulatory oesophageal pH monitoring for gastro-oesophageal reflux disease

Rationale for assessment

An application requesting Medicare Benefits Schedule (MBS) listing of catheter-free (wireless) ambulatory oesophageal pH monitoring for gastro-oesophageal reflux disease (GORD) was received from Given Imaging Pty Ltd (the applicant) by the Department of Health in January 2012. Catheter-*free* monitoring is proposed for patients who have previously failed catheter-*based* monitoring or in whom the use of the latter is anatomically inappropriate.

Gastro-oesophageal reflux disease (GORD) and pH monitoring

Gastric acid reflux in the oesophagus is a normal physiological event and it usually occurs after the consumption of a meal. Acid reflux can lead to impairment in quality of life or physical complications such as oesophagitis, Barrett's oesophagus or oesophageal adenocarcinoma. Although there is no clear internationally applied definition for GORD, one of the accepted definitions is: 'a condition which develops when the reflux of stomach contents causes troublesome symptoms and/or complications'. Reflux is considered to impact on quality of life when symptoms are present on two or more days a week.

When a GORD patient's symptoms are resistant to acid suppression therapy, endoscopy is the primary test to exclude reflux disease and identify other causes. If the endoscopy result is negative, oesophageal pH monitoring is widely used to provide additional information. Oesophageal pH-monitoring measures the exposure of the oesophagus to acid reflux over a defined period of time, whereas endoscopy can only provide information at a single point in time. Currently, the 24-hour catheter-*based* pH-monitoring test is the most commonly used method of oesophageal pH monitoring. Catheter-based monitoring quantifies the time the oesophagus is exposed to acid and analyses the association between symptoms and reflux events. However, not every patient is able to tolerate catheter-based monitoring, e.g. due to gagging or anatomical challenges; as a consequence, they do not receive any pH monitoring.

Catheter-*free* (wireless) ambulatory oesophageal pH monitoring is a system that consists of a small capsule that is usually delivered transorally during an oesophagoscopy, and a

wireless receiver. The pH-monitoring capsule is temporarily attached to the oesophageal wall to measure and transmit pH information for at least 48 hours, sending it to the receiver and thereby circumventing the need for an oesophageal catheter.

The primary comparator for catheter-free monitoring is 'no pH monitoring', although the Protocol Advisory Sub-Committee (PASC) of the Medicare Services Advisory Committee (MSAC) requested that the accuracy of catheter-free monitoring be compared against 'catheter-based monitoring'.

Current arrangements for public reimbursement

There have been no previous MSAC considerations of catheter-free monitoring. The current test for GORD, 24-hour catheter-based monitoring, is listed on the MBS (item 11810) (see Table 2). There were 3,590 claims for catheter-based monitoring from July 2012 to June 2013.

The device used in catheter-free monitoring was listed on the Australian Register of Therapeutic Goods (ARTG) by the Therapeutic Goods Administration (TGA), and is currently classified as a Medical Device Class IIa (see Table 1). Its use in Australia is still limited: only 27, 22 and 46 catheter-free monitoring capsules were sold to Australian sites in 2011, 2012 and 2013, respectively.

Clinical need

Reflux disease is common. Of patients attending GPs, 10.4% have GORD (Knox et al. 2008). Studies indicate that between 14% and 25% of adults experience reflux symptoms at least once a week. Although the prevalence of GORD is high, there is a relatively low usage of MBS item 11810 (catheter-based pH monitoring), so it would appear that pH monitoring is not often used for the study of this condition. The new intervention (catheter-free or wireless monitoring) will be used in patients who either fail catheter-based monitoring or in whom its use is anatomically inappropriate. It may avoid lengthy periods of unnecessary pharmacologic therapy in this group, expected to be 5–10% of the patients eligible for pH monitoring. An algorithm outlining the clinical management of patients with GORD, including the use of pH monitoring, can be found on page 37, (Figure 3).

Results of assessment

Safety: Only one study was identified that assessed the safety of catheter-free monitoring compared with no monitoring. The findings indicated that chest pain was more likely to

occur in those undergoing catheter-free monitoring than in those who did not undergo monitoring (RR=2.33, 95%CI 0.81, 6.76). Twelve studies compared the safety of catheter-free monitoring with catheter-based monitoring. No deaths or life-threatening events caused by pH monitoring were reported. Chest pain was significantly more prevalent in patients undergoing catheter-free monitoring compared with catheter-based monitoring, possibly due to the attachment of the capsule to the oesophageal wall. In a high-quality randomised controlled trial by Andrews et al. (2012), the median chest pain score during the pH test was 29 ± 4 mm on a 100-point visual analogue scale (VAS) in the catheter-free monitoring group, compared with only 14 ± 3 mm in the catheter-based monitoring group (p=0.001). The difference is likely to be clinically meaningful. However, other adverse events such as nose and throat pain, dysphagia, eating and drinking difficulties, and headache were more prevalent with catheter-based monitoring. Furthermore, catheter-based monitoring caused more overall discomfort and had a greater negative impact on normal daily activities and work attendance.

Twenty-nine non-comparative studies of level IV interventional evidence were included to determine the rate of other complications and/or adverse events caused by catheter-free monitoring. Most adverse events reported were related to discomfort caused by the capsule attaching to the oesophageal mucosa, e.g. pharyngeal irritation, retrosternal discomfort and throat pain. More-severe complications from catheter-free monitoring, such as a rupture in the oesophageal mucosa, were rare.

Effectiveness: Studies assessing the direct health impact of catheter-free monitoring versus no monitoring in the selected study population were not available. Two studies that did not meet the pre-specified inclusion criteria, but reported on the health effects of catheter-free monitoring, were included in the absence of more relevant data. One comparative study reported an improvement or disappearance of the principal symptom in 73% of patients in the wireless group (n=51) and 69% in the catheter-based group (n=51). A case series reported 9 patients with a good or moderate improvement in symptoms, out of 26 patients who received medical therapy or conservative advice after a catheter-free monitoring test.

The sensitivity and specificity of catheter-free monitoring could be calculated from eight studies. In three of these studies the reference standard was catheter-based monitoring. The highest quality study, of level II diagnostic accuracy, indicated that a catheter-free monitoring test with a diagnostic cut-off point of 4.4% of the time that pH <4 had good sensitivity and specificity, with rates of 86.4% and 77.8%, respectively. The accuracy of catheter-free monitoring in the two other studies was similar, although test sensitivity was

slightly lower and specificity slightly higher. Five diagnostic case-control studies, of level III-3 diagnostic accuracy, used clinical diagnosis as a reference standard. The study with the largest patient population reported sensitivity ranging between 59% and 88% and specificity between 75% and 96%, depending on which cut-off value was used (between 1.9% and 4.4% of the time that pH <4).

There were nine studies that reported on oesophageal acid exposure times during catheter-free monitoring in adults. Most (five out of eight) of the studies reporting on adults found no difference between the two measurement methods. There was one study that reported concordance data on oesophageal acid exposure in children, finding significantly more reflux with wireless monitoring compared with catheter-based monitoring (p=0.01). Three of four available studies reported significantly more reflux events with the catheter-based monitoring system. Monitoring time was increased in the wireless-based compared with the catheter-based method. Six studies compared the yield of GORD diagnoses after 24 and 48 hours of catheter-free monitoring, and reported an additional yield on day 2 of monitoring, varying from 2.0% to 15.6% with a median of 7.8%.

Nine studies reported on the technical efficacy of the wireless system. A meta-analysis was conducted and showed that there was over three times the risk (RR=3.3, 95%CI 1.63, 6.81) of having technical problems with the catheter-free monitoring system compared with catheter-based monitoring in adults. The risk decreased slightly (RR=2.87, 95%CI 1.47, 5.62) when used in a combined population of adults and children. The most commonly reported technical failures associated with catheter-free monitoring were early capsule detachments, causing incomplete data capture, and errors in placement of the system, i.e. in the wrong place in the oesophagus or where the capsule fails to attach to the oesophageal mucosa. On average, problems occurred more often during insertion or placement of the catheter-free system (median 7.5% of procedures, range 0–12%) than the catheter-based system (median 0.9% of procedures, range 0–4%). In most studies (15/20) the oesophageal acid exposure remained the same across 2 days of catheter-free monitoring in adults.

Five studies reported on change in patient management after catheter-free monitoring, in patients who were suspected of GORD but who could have potentially tolerated a catheter-based test. The results of these studies were described due to a lack of more-relevant data in patients unable to tolerate a catheter-based test. One study reported that 42.2% (38/90) of patients continued to take proton pump inhibitors (PPIs) despite a negative pH test result. A second study reported that medical therapy or conservative advice was pursued in 68% of cases at a median of 2 years after the test, and two other studies showed that

monitoring led to a change in management in 63% of adults and 88% of children. Finally, one other study stated that concordance between the test and treatment was higher after catheter-free monitoring than catheter-based monitoring: 78% vs 58% (p<0.05).

To answer the question whether a change in management leads to better health outcomes, a second literature search was conducted. No evidence was found to determine the psychological or physical impact of a false negative test result. For people initially suspected of having GORD but eventually given an alternative diagnosis (i.e. true negatives), it was assumed that their management had been optimised as a result of obtaining the correct diagnosis. However, two retrospective case series reported poorer outcomes and worse treatment responses in infants with infantile spasms misdiagnosed as GORD (false positive test results) compared with infants with no treatment lag. Two systematic reviews comparing medical treatment with anti-reflux surgery reported that surgery was more effective than medical management at treating GORD symptoms, at least in the short to medium term.

Two randomised controlled trials (RCTs) have been published since the high-quality Wileman et al. (2010) review was published. After 5 years laparoscopic fundoplication continued to provide better relief of GORD symptoms than medical management.

Economic and financial considerations

Two economic models were constructed to determine: the cost-effectiveness of catheter-free pH monitoring vs no monitoring (empirical treatment) in an Australian population who cannot tolerate catheter-based monitoring; and the cost-effectiveness (or not) of catheter-free monitoring vs catheter-based monitoring should use of the proposed listing 'leak' to include patients who are not intolerant of catheter-based monitoring.

Characteristics of the cost–utility models, which assume a health system perspective from public and private healthcare providers and defined patient contributions are outlined in Table ES 1.

Table ES 1   Summary of the economic evaluation (applies to both models)

Time horizon	Base-case analysis of 15 years
Outcomes	Quality-adjusted life years
Methods used to generate results	Markov model with half-cycle correction
Cycle length	1 year
Discount rate	5% for both costs and outcomes
Software packages used	TreeAge Pro and Excel (hybrid)

Key structural assumptions were: that a positive pH test result provided patients the option of surgery to treat GORD symptoms; and that a negative result would prompt investigation for other diagnoses and reduce the use of ongoing high-dose PPIs. When modelling, it was considered that inclusion of the accuracy of wireless pH monitoring when measured against catheter-based monitoring—an imperfect reference standard—would distort the results; therefore, the base-case assumes wireless pH monitoring, catheter-based monitoring and a trial of empirical treatment (high-dose PPIs) to be equally accurate. Sensitivity analyses explore alternative test accuracy values.

The base-case analysis found that wireless pH monitoring vs no pH monitoring had an incremental cost-effectiveness ratio of $14,457 per quality-adjusted life year (QALY) gained. However, when compared against catheter-based monitoring, wireless monitoring was dominated—it had a higher cost and was less effective, with the lesser effectiveness due to an increased technical failure rate. Thus, if catheter-free monitoring were to be MBS listed, leakage of use into patients who could otherwise tolerate catheter-based monitoring may substantially reduce the assumed cost-effectiveness. The overall cost-effectiveness of an MBS listing is dependent on the predicted extent of leakage, with 30% leakage into cohorts of patients who are able to tolerate, or do not fail, catheter-based monitoring increasing the overall incremental cost-effectiveness ratio (ICER) to $58,429/QALY.

Other key areas of uncertainty in the cost–utility models relate to: the assumed accuracy of the test—where imperfect sensitivity and specificity values are incorporated, sensitivity analyses showed that wireless monitoring was dominated, with sensitivity values ≤90%; and the assumption that, in the absence of a pH test (which incorporates a follow-up assessment of results), some patients will be trialled on high-dose PPIs and inappropriately remain on this treatment indefinitely. That this occurs is supported by the literature and expert opinion, but quantification is highly uncertain. If the assumption is removed altogether, wireless pH monitoring is dominated by an empirical trial of high-dose medication. The base-case assumes that inappropriate ongoing high-dose PPI use occurs in 1 in 10 patients

diagnosed with non-erosive reflux disease (NERD)-like symptoms but who actually have a non-acid related condition.

It is estimated that the 'true' catheter-intolerant patient population who would benefit from wireless monitoring would be approximately 400 patients per year (2015–16 estimate = 404), increasing by 4%–8% per year, based on the current extent and growth of catheter-based pH monitoring MBS claims. This estimate would result in MBS costs of approximately $369,000 directly associated with the new item number, and a further $67,000 in associated MBS-funded anaesthesia items, per year (with annual growth). There are no MBS cost offsets in this patient population. Of this expenditure of $436,000, patients are expected to pay approximately $100,000.

The overall financial impact of the proposed listing is also highly dependent on the extent of predicted use beyond the intended catheter-intolerant population, and it is noted that the proposed fee, in including an 'equipment depreciation' component, provides an incentive that encourages increased servicing.

If 90% (an extreme upper plausible limit) of patients who are tolerant to catheter-based monitoring 'switched' to wireless monitoring, the net MBS impacts, including cost offsets of 'unused' catheter-based monitoring, would exceed $3.5 million per year, of which approximately $880,000 would be paid through patient contributions.

Other non-MBS costs incurred are private hospital costs, which may range from $275,000 to $2.5 million across the population, depending on leakage.

Other relevant considerations / discussion

The catheter-free monitoring test is seen as more convenient and has less impact on daily activities than catheter-based monitoring. It can be difficult to determine who is 'unable' to tolerate a catheter. Physicians might therefore be inclined to give more patients catheter-free monitoring than catheter-based monitoring to avoid discomfort and embarrassment. Thus, it is possible that the population using catheter-free monitoring may expand beyond the population currently being considered for eligibility for the device, and this 'leakage' could lead to extra costs (see 'Economic and financial considerations').

The validity of catheter-based monitoring as a reference standard to determine the diagnostic accuracy of catheter-free monitoring is questionable. First, catheter-free monitoring is able to monitor acid reflux for 48 hours instead of the 24 hours used with catheter-based monitoring; this leads to increased diagnostic yield. Second, many of the

short reflux events that are measured by catheter-based monitoring and not by catheter-free monitoring could be artefacts that are not related to GORD. It is possible that catheter-free monitoring is more accurate than catheter-based monitoring.

There are other limitations with the reported test accuracy data. There was considerable variation and a lack of consensus across the studies in the evidence-base with regard to diagnostic thresholds (cut-off points), defining a positive test result (worst day analysis, positive value on both days, reflux symptom correlation) and determining the optimal position of the capsule in the oesophagus. There is therefore a considerable lack of certainty with regard to the reported test accuracy results.

Glossary and abbreviations

AHTA	Adelaide Health Technology Assessment
ARTG	Australian Register of Therapeutic Goods
catheter-based monitoring	catheter-based pH monitoring
catheter-free monitoring	catheter-free (wireless) pH monitoring
CBPM	catheter-based pH monitoring
CFPM	catheter-free (wireless) pH monitoring
CI	confidence interval
DAP	decision analytic protocol
FH	functional heartburn
GOJ	gastro-oesophageal junction
GORD	gastro-oesophageal reflux disease
H. pylori	*Helicobacter pylori*
H2	histamine 2 (receptor)
HESP	Health Expert Standing Panel
HO	hypersensitive oesophagus
ICER	incremental cost-effectiveness ratio
LOS	lower oesophageal sphincter
LR+	likelihood ratio positive
LR–	likelihood ratio negative
MBS	Medicare Benefits Schedule
MII	multichannel intraluminal impedance
MSAC	Medical Services Advisory Committee
NERD	non-erosive reflux disease
NHMRC	National Health and Medical Research Council

NPV	negative predictive value
NS	not significant
OeO	oesinophilic oeosphagitis
PASC	Protocol Advisory Sub-Committee (of MSAC)
PBS	Pharmaceutical Benefits Scheme
PPIs	proton pump inhibitors
PPV	positive predictive value
QALY	quality-adjusted life year
QoL	quality of life
RCT	randomised controlled trial
SCJ	squamo-columnar junction
VAS	visual analogue scale

Introduction

This assessment report is intended for the Medical Services Advisory Committee (MSAC). MSAC evaluates new and existing health technologies and procedures for which funding is sought under the Medicare Benefits Schedule (MBS) in terms of their safety, effectiveness and cost-effectiveness, while taking into account other issues such as access and equity. MSAC adopts an evidence-based approach to its assessments, based on reviews of the scientific literature and other information sources, including clinical expertise.

Adelaide Health Technology Assessment (AHTA), School of Population Health, University of Adelaide, has been commissioned by the Australian Government Department of Health to conduct a systematic literature review and economic evaluation of the use of catheter-free (wireless) oesophageal pH monitoring in the diagnosis of gastro-oesophageal reflux disease (GORD). This evaluation has been undertaken in order to inform MSAC's decision-making regarding public funding of the intervention.

The proposed use of catheter-free (wireless) oesophageal pH monitoring in Australian clinical practice was outlined in a decision analytical protocol (DAP) that guided the evaluation undertaken by AHTA. The DAP was released for public comment on 28 May 2012 and closed for comments on 5 July 2012. No public consultation responses were received. The DAP was finalised as a result of PASC deliberation on 16–17 August 2012.

Rationale for assessment

Given Imaging Pty Ltd (the applicant) submitted an application to the Department of Health in January 2012, requesting an MBS listing for catheter-*free* (wireless) ambulatory oesophageal pH monitoring for the diagnosis of GORD. Its use is proposed in patients who have previously failed catheter-*based* monitoring or in whom the use of the latter is anatomically inappropriate.

Background

Gastro-oesophageal reflux disease (GORD)

Gastric acid reflux in the oesophagus is a normal physiological event and it usually occurs after the consumption of a meal. GORD occurs when acid reflux impairs a person's quality of life or when it leads to a risk of physical complications such as oesophagitis, Barrett's oesophagus or oesophageal adenocarcinoma (Gastroenterological Society of Australia 2011). Although there is no clear internationally applied definition for GORD, it is defined by consensus (Dent et al. 2005; Katz, Gerson & Vela 2012). The current definition of the disorder being used by the International Foundation of Functional Gastrointestinal Disorders (IFFGD) and the American Gastroenterological Association is: 'a condition which develops when the reflux of stomach contents causes troublesome symptoms and/or complications' (IFFGD 2014; Kahrilas 2008; Kahrilas, Shaheen & Vaezi 2008). Reflux is considered to impact on quality of life when symptoms are present on two or more days a week. The definition used by the Gastroenterological Society of Australia (2013) is 'persistent reflux that occurs more than twice a week'. Oesophagitis, which is present in about one-third of GORD patients, occurs when excessive reflux of acid and pepsin results in necrosis of surface layers of the oesophageal mucosa, causing erosions and ulcers.

Although the absolute risk of oesophageal cancer in an individual with GORD is small (Gastroenterological Society of Australia 2011), oesophageal adenocarcinoma is epidemiologically linked to GORD and the incidence of both is increasing (Kahrilas 2008; Lagergren et al. 1999). Up to 15% of patients with chronic reflux disease undergoing endoscopy may be shown to develop Barrett's oesophagus. This condition involves the development of columnar metaplasia in place of the normal squamous epithelium, and is a risk factor for the development of oesophageal adenocarcinoma (Shaheen & Richter 2009).

Several factors can predispose patients to GORD, such as a hiatus hernia, lower oesophageal sphincter (LOS) hypotension, loss of oesophageal peristaltic function, obesity, increased compliance of the hiatal canal, gastric hypersecretory states, delayed gastric emptying, overeating and alcohol consumption. Often, multiple risk factors are present (Gastroenterological Society of Australia 2011; Kahrilas 2008). Also, the following specific patient groups are at higher risk of GORD and its complications (Gastroenterological Society of Australia 2011):

- patients with connective tissue diseases such as scleroderma

- patients with chronic respiratory diseases such as asthma and cystic fibrosis

- people who are institutionalised and/or intellectually handicapped

- patients nursed in a supine position for prolonged periods.

Most GORD patients experience a burning feeling rising up from the stomach or lower chest towards the neck, known as heartburn. However, the symptoms of GORD, ulcer disease, irritable bowel syndrome (IBS) and dyspepsia are often similar. When a patient exhibits typical GORD symptoms and responds to therapy, no additional diagnostic tests are necessary. Diagnostic testing is usually done to avert misdiagnosis, identify complications of GORD and evaluate treatment failures. Other diagnoses that could be considered when GORD treatment fails include coronary heart disease; gallbladder disease; gastric or oesophageal cancer; peptic ulcer disease; oesophageal motility disorders; and oesinophilic, infectious or pill oesophagitis (Kahrilas 2008). When the patient's symptoms are resistant to acid suppression therapy, endoscopy is the primary test to exclude reflux disease and to diagnose other causes. However, the yield of endoscopy is low because of the poor correlation between symptoms of GORD and oesophagitis, the likelihood that oesophagitis may already have been resolved with previous therapy, and its poor sensitivity at detecting motility disorders. If the endoscopy result is negative, other physiological testing can be considered, such as oesophageal manometry to detect motility disorders, ambulatory pH monitoring to demonstrate abnormal exposure to acid (see below), or impedance monitoring to identify reflux events regardless of acidic content (Gastroenterological Society of Australia 2011; Kahrilas 2008).

Treatments for GORD depend on symptom severity and include lifestyle modification, pharmacologic therapy [proton pump inhibitors (PPIs), H2 (histamine 2) antagonists] and surgical intervention. PPIs and H2 antagonists both reduce gastric acid production, although PPIs seem to be more effective than H2 antagonists (Eriksson et al. 1995). Reflux disease patients with severe oesophagitis may be treated with a standard dose of PPIs on an ongoing basis. A surgical intervention, usually laparoscopic fundoplication, is indicated in patients whose GORD symptoms cannot be controlled using medical therapy, or who wish to avoid being medicated over the long term and accept the risks associated with a surgical procedure.

pH monitoring

Oesophageal pH monitoring is widely used to monitor GORD after a negative endoscopy result. It provides additional information by monitoring acid reflux over a period of time

rather than, in the case of endoscopy, at one point in time. It is useful when the relationship between symptoms and reflux aetiology remains unclear.

A pH-monitoring study may reveal PPI failure and associated ongoing acid reflux, which may require dose escalation of therapy, adequate acid control or no reflux. When the pH study in refractory GORD patients is negative, those with heartburn may be classified as having 'functional heartburn'; and in those with other ongoing, atypical symptoms, other diagnoses may be considered (Katz, Gerson & Vela 2012).

The current MBS-approved pH-monitoring test for GORD: catheter-based ambulatory oesophageal pH monitoring (MBS item 11810)

Currently, the 24-hour catheter-*based* pH-monitoring system is commonly used to objectively diagnose GORD (Weber, Davis & Fisichella 2011). Before the measurement starts, a pH-measuring sensor with a catheter is passed transnasally and placed with manometric guidance. This is often a cause of discomfort for the patient. The catheter is then taped to the patient's nose, and the patient returns to the hospital or rooms of the specialist 24 hours later to have the catheter removed and the data analysed.

Although catheter-based monitoring quantifies the time the oesophagus is exposed to acid and measures the association between symptoms and reflux events, it does have some disadvantages. It is claimed that poor patient tolerance is the main drawback, as it leads to patient discomfort, social embarrassment, and alteration of daily activities and eating habits during the monitoring period (Figure 1). Other claimed disadvantages include variable test accuracy and a limited observation period that leads to less-reproducible results. It is very important that the sensors are correctly placed in order for the results to be valid (Lee 2012; Weber, Davis & Fisichella 2011). The information from the sensor is used to guide or reassess treatment strategies for patients.
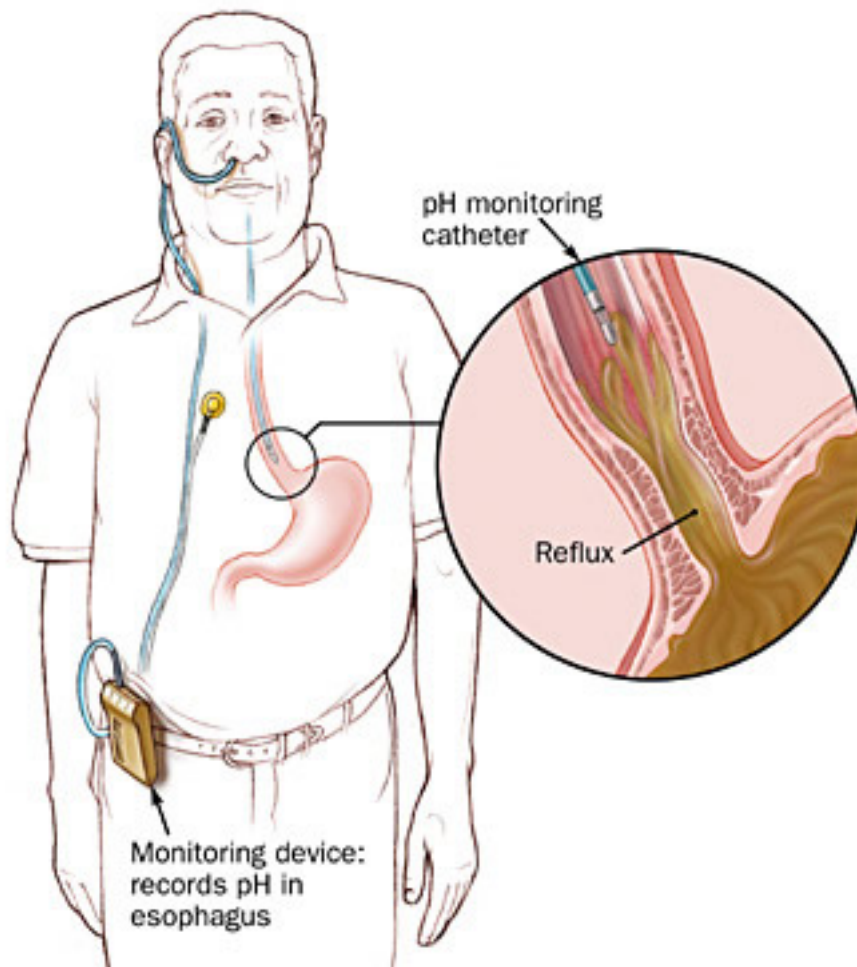
Figure 1   Catheter-based pH monitoring

Source: http://gastroarkansas.com/Web_Development/Web_Pages/24_HOUR_pH.html [accessed August 2013]

Catheter-free (wireless) ambulatory oesophageal pH monitoring

Catheter-*free* (wireless) ambulatory oesophageal pH monitoring for use in the confirmatory diagnosis of GORD is a system that consists of a small capsule (which is usually delivered transorally during an oesophagoscopy or sometimes transnasally during manometry) plus a wireless receiver that the patient wears on his or her belt. The radiotelemetry pH-monitoring capsule is temporarily attached to the oesophagus wall to measure and transmit pH information every 6 seconds for at least 48 hours, sending it to the receiver and thereby circumventing the need for an oesophageal catheter (Maerten et al. 2007; Weber, Davis & Fisichella 2011). The wireless capsule is invisible, so the patient should be better able to carry out his/her normal daily activities, which should lead to more reliable and accurate results.

The procedure

First, the capsule is calibrated by submersion in pH buffer solutions, and is then placed through the mouth, and sometimes the nose, into the oesophagus via endoscopic guidance. The distal end of the capsule is placed approximately 6 cm above the squamo-columnar junction (SCJ) or 5 cm above the proximal border of the LOS via manometric guidance. The LOS is usually localised before capsule placement (Nusrat, Roy & Bielefeldt 2012). The capsule itself measures 6 x 5.5 x 25 mm and contains a 3.5 mm-deep well that is connected to an external vacuum unit. When the capsule is being placed, suction is applied through the suction channel, causing the adjacent mucosa to be drawn into the well of the capsule. At this point a spring-loaded mechanism puts a little stainless steel pin through the well of the capsule, securing the mucosa in the well and attaching the capsule to the oesophageal wall. The delivery system is then removed (Maerten et al. 2007; Pandolfino et al. 2003).

The single-use capsule makes catheter-free monitoring more expensive than the traditional catheter-based monitoring, which usually uses a reusable catheter, although single-use catheters are also available. During the pH recording days, patients keep a diary where they record food intake, symptoms and activities, including changes in their position. After 48 hours the patient returns to the hospital or specialist's rooms, and the information on the receiver is uploaded to a computer and interpreted by a physician. AccuView is the software used to review and analyse the test results. Recent versions of the software are compatible with both the Bravo pH (catheter-free) and Digitrapper pH-Z (catheter-based) monitoring systems, eliminating the need to learn and maintain separate systems. The information in the diary is used to interpret the pH data. The capsule usually spontaneously dislodges from the oesophageal mucosa within 15 days. However, in some cases it does not spontaneously detach and needs to be endoscopically removed (Pandolfino et al. 2003).

Intended purpose

Between 5% and 10% of patients eligible for pH monitoring are intolerant of the intubation required for the transnasal catheter or they fail to complete 24-hour catheter-based monitoring (Sweis et al. 2009). Early referral for pH monitoring, and pH monitoring of patients who fail catheter-based monitoring, may avoid lengthy periods of unnecessary medical therapy (Kleiman et al. 2012). The expectation is that pH monitoring will change management in patients suspected of GORD in approximately 50% of cases, although such changes are not always maintained, and this may lead to an increased quality of life for this patient group (Eckardt, Dilling & Bernhard 1999; Hirano, Richter & Practice Parameters Committee of the American College of 2007; Netzer et al. 1999).

Indications

The American College of Gastroenterology documented the usefulness of pH monitoring in GORD patients and its impact on patient management (Hirano, Richter & Practice Parameters Committee of the American College of 2007; Katz, Gerson & Vela 2012). They determined that pH monitoring is useful to:

- quantify reflux and assess the relationship between reflux episodes and the patient's symptoms

- document abnormal oesophageal acid exposure in endoscopy-negative individuals being considered for a surgical anti-reflux procedure

- evaluate endoscopy-negative patients with typical or atypical reflux symptoms that are refractory to PPI therapy

- document the adequacy of PPI therapy in oesophageal acid control in patients with complications of GORD that include Barrett's oesophagus.

Of these patients, catheter-*free* monitoring would be indicated for those in whom catheter-*based* monitoring has failed or is anatomically inappropriate.

Contraindications for catheter-free (wireless) ambulatory oesophageal pH monitoring include the following (Maerten et al. 2007):

- haemorrhagic diathesis

- oesophageal varices

- severe oesophagitis

- oesophageal strictures or diverticula

- having a pacemaker or defibrillator

- pregnancy.

Clinical need

Reflux disease is common. Knox et al. (2008) estimated the prevalence of GP-diagnosed GORD in Australia to be around 10.4% of patients. The prevalence in the Australian population was estimated at 9.2%, which equates to approximately 2 million Australians with GORD. Knox et al. reported on the percentage of patients who were currently under clinical management, which means that this study probably underestimated the total GORD prevalence. Other studies report higher prevalence rates, ranging between 14% and 25% of adults experiencing symptoms of reflux at least once a week (Gastroenterological Society of Australia 2011; Kahrilas 2008; Miller & Pan 2009). The lower prevalence recorded from GP-

encounter data compared with community-based studies may indicate that there is a significant level of unmet need for management of this disease in the Australian community.

There is a lack of epidemiological data on GORD in infants and children. However, one American study reported that at least one episode of regurgitation a day occurs in approximately half of infants aged 0–3 months, and heartburn and regurgitation were reported on a weekly basis in about 2% of 3–9 year old children and 5%–8% of 10–17 year old children (Nelson et al. 2000).

GORD occurs more often in developed countries and may be associated with reduction of *Helicobacter pylori* (*H. pylori*) colonisation of the gastric mucosa due to better hygiene, a greater use of antibiotics, and the increasing frequency of obesity in the community (Knox et al. 2008). However, the relationship between *H. pylori* infection and GORD is still controversial (Katz, Gerson & Vela 2012). It is known that GORD patients with daily or weekly symptoms are more likely to have time off work, a decrease in work productivity, lower quality sleep and a decrease in physical functioning. Nocturnal GORD has an even greater impact on quality of life compared with daytime symptoms (Katz, Gerson & Vela 2012).

Other existing tests for GORD

Oesophageal impedance testing

In 1991 multichannel intraluminal impedance (MII) was first described. It is a technology that is able to detect refluxate moving from the stomach into the oesophagus, independent of the content. With this method it is possible to determine the composition, distribution and clearance of gastro-oesophageal reflux, differentiating between the refluxate of liquid, gas, and combined liquid and gas. Impedance depends on the conductivity of material through which current travels; this means that MII measures the electrical conductivity across a pair of electrodes in the oesophageal lumen, and changes in conductivity as a result of intraoesophageal materials moving will be recorded. It can detect movement in both the retrograde (reflux) and the anterograde (swallow) direction. This test uses a catheter and can be combined with catheter-based monitoring. Therefore, it could be an important tool to identify reflux of all types, independent of their acid content (Hirano 2006; Tutuian 2008; Weber, Davis & Fisichella 2011; Wilson & Vela 2008).

Marketing status of device/technology

All therapeutic products marketed in Australia require listing on the Australian Register of Therapeutic Goods (ARTG). The device for catheter-free monitoring is listed on the ARTG under the item shown in Table 1. It is stated that 'the Bravo pH-monitoring system is intended to be used for gastro-oesophageal pH measurement and monitoring of gastric reflux'.

Table 1   Catheter-free monitoring device listed on the ARTG

ARTG no.	Product no.	Product description	Product category	Sponsor
205795	58357	Gastrointestinal telemetric monitoring system	Medical Device Class IIa	Given Imaging Pty Limited

Source: Therapeutic Goods Administration, https://www.ebs.tga.gov.au/ [accessed 14 August 2013]

Current reimbursement arrangements

Currently, there is no listing on the MBS for catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD. As a result, its use in Australia is limited: only 27, 22 and 46 capsules were sold to Australian sites in 2011, 2012 and 2013, respectively[1].

The diagnostic test for GORD currently listed on the MBS—24-hour ambulatory catheter-based pH monitoring—is shown in Table 2.

Table 2   Relevant MBS item for the currently reimbursed test for GORD

Category 2 – Diagnostic procedures and interventions
MBS 11810 CLINICAL ASSESSMENT of GASTRO-OESOPHAGEAL REFLUX DISEASE involving 24 hour pH monitoring, including analysis, interpretation and report and including any associated consultation Fee: $174.45 Benefit: 75% = $130.85 85% = $148.30

Although GORD is relatively common, there is fairly low usage of the current MBS item for pH monitoring; there were only 3,590 claims for catheter-based monitoring from July 2012 until June 2013 (Figure 2). It is claimed that part of the low usage for the current test could reflect the discomfort and inconvenience of having a catheter in place for 24 hours, so that

---

[1] Information provided by Given Imaging Pty Ltd, North Ryde, Australia on 4 November 2013; sales data for 2013 are from January until October

use of pH-monitoring devices might increase if catheter-free monitoring is publicly funded on the MBS.
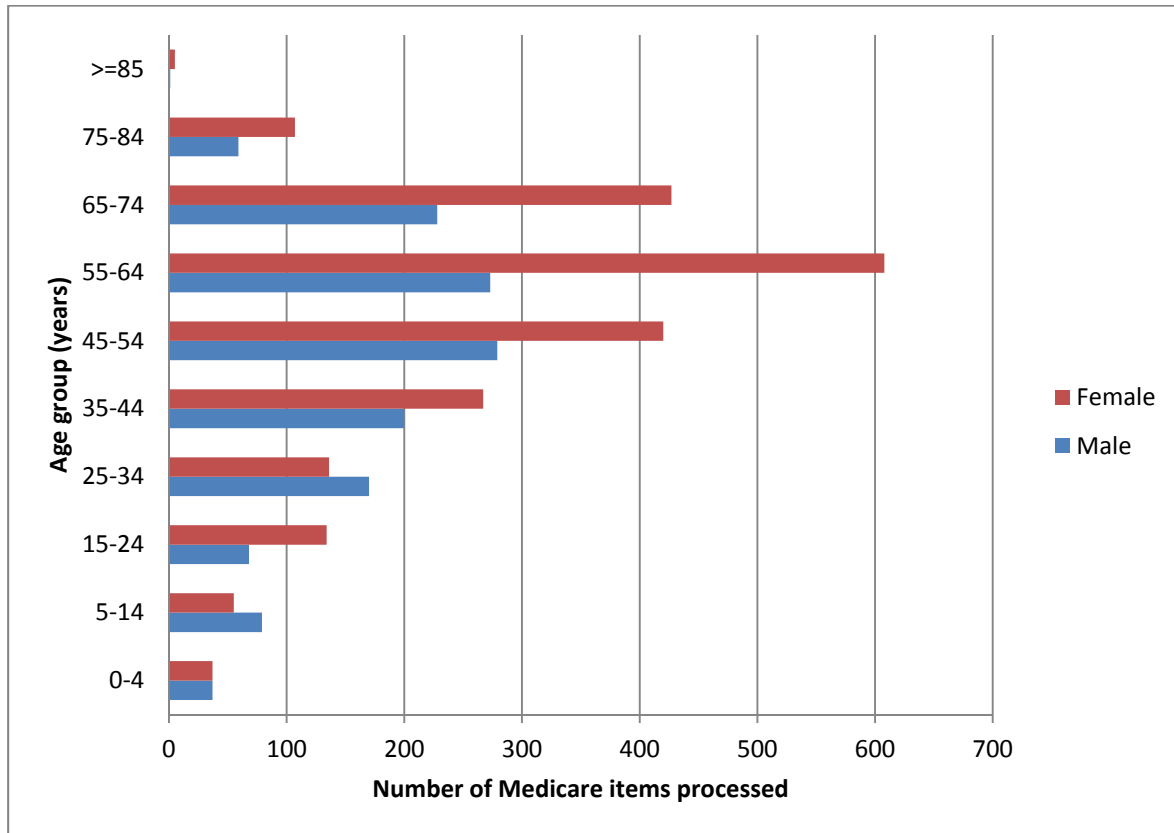
Figure 2   Usage of catheter-based pH monitoring, July 2012 – June 2013

Source: Medicare (2013)

Surgical treatments for GORD are listed under five different MBS item numbers: 43951, 43954, 31464, 30527 and 30530. A total of 1,834 anti-reflux procedures were claimed from July 2012 to June 2013 (Medicare 2013). For the description and usage of each individual item number, see Appendix I.

Proposal for public funding

The proposed MBS item is summarised in Table 3. The fee was suggested by the applicant. It includes a $350 fee for professional time for performing the test, $430.40 for the capsule itself and $133.25 for the reader system. The proposed item descriptor makes it clear that it is a requirement that all individuals undergoing catheter-*free* monitoring have previously failed catheter-*based* monitoring or are anatomically inappropriate for a catheter-based system. Furthermore, the test should only be performed by a specialist or consultant

physician with endoscopic training that is recognised by the Conjoint Committee for the Recognition of Training in Gastrointestinal Endoscopy.

Table 3   Proposed MBS item descriptor

Category 2 – Diagnostic procedures and interventions

CLINICAL ASSESSMENT of GASTRO-OESOPHAGEAL REFLUX DISEASE that involves 48 hour catheter-free wireless ambulatory oesophageal pH monitoring including administration of the device and any endoscopy associated with this, analysis and interpretation of the data and all attendances providing the service, if
 	
    (a)   The service is performed by a specialist or consultant physician with endoscopic training that is recognised by the Conjoint Committee for the Recognition of Training in Gastrointestinal Endoscopy; and
 	
    (b)   The patient has previously failed (rather than intolerant of) a catheter-based ambulatory oesophageal pH-monitoring or is anatomically inappropriate for a catheter-based-system

Fee: $913.64

Consumer impact statement

PASC received one response from the public when the DAP was put out for public consultation—from the professional body, the Gastroenterological Society of Australia (GESA).

Approach to assessment

The purpose of a DAP is to describe in detail a limited set of decision option(s) associated with the possible public funding of a proposed medical service. A DAP also accurately captures current clinical practice, reflects likely future practice with the proposed medical service, and describes all potentially affected healthcare resources. A DAP was developed prior to the commencement of this assessment, and was approved by the PASC of MSAC. The guiding framework of the DAP was used throughout this assessment.

Objective

To determine whether there is sufficient evidence, in relation to safety, effectiveness and cost-effectiveness, to have catheter-free monitoring for GORD listed on the MBS for patients who have previously failed a catheter-based test or for whom catheter-based monitoring is anatomically inappropriate.

Clinical pathway

A flowchart can help define the place of a new intervention in the clinical management of a patient. The solid lines in Figure 3 indicate the current clinical management pathway, whereas the dashed lines show the proposed change in management if catheter-free monitoring becomes available. The tests prior to pH monitoring are also shown in the pathway, and the relevant population being targeted by catheter-free monitoring can be found in the shaded box.

The proposed pathway indicates that there would be more diagnostic options for patients who have either failed a catheter-based pH test or in whom it is anatomically inappropriate to use the currently available test (i.e. as an alternative to no pH monitoring). Instead of no further testing and continuation of potentially ineffective PPI treatment, patients are able to undergo a pH-monitoring test that would lead to other treatment options such as surgery or different medication.
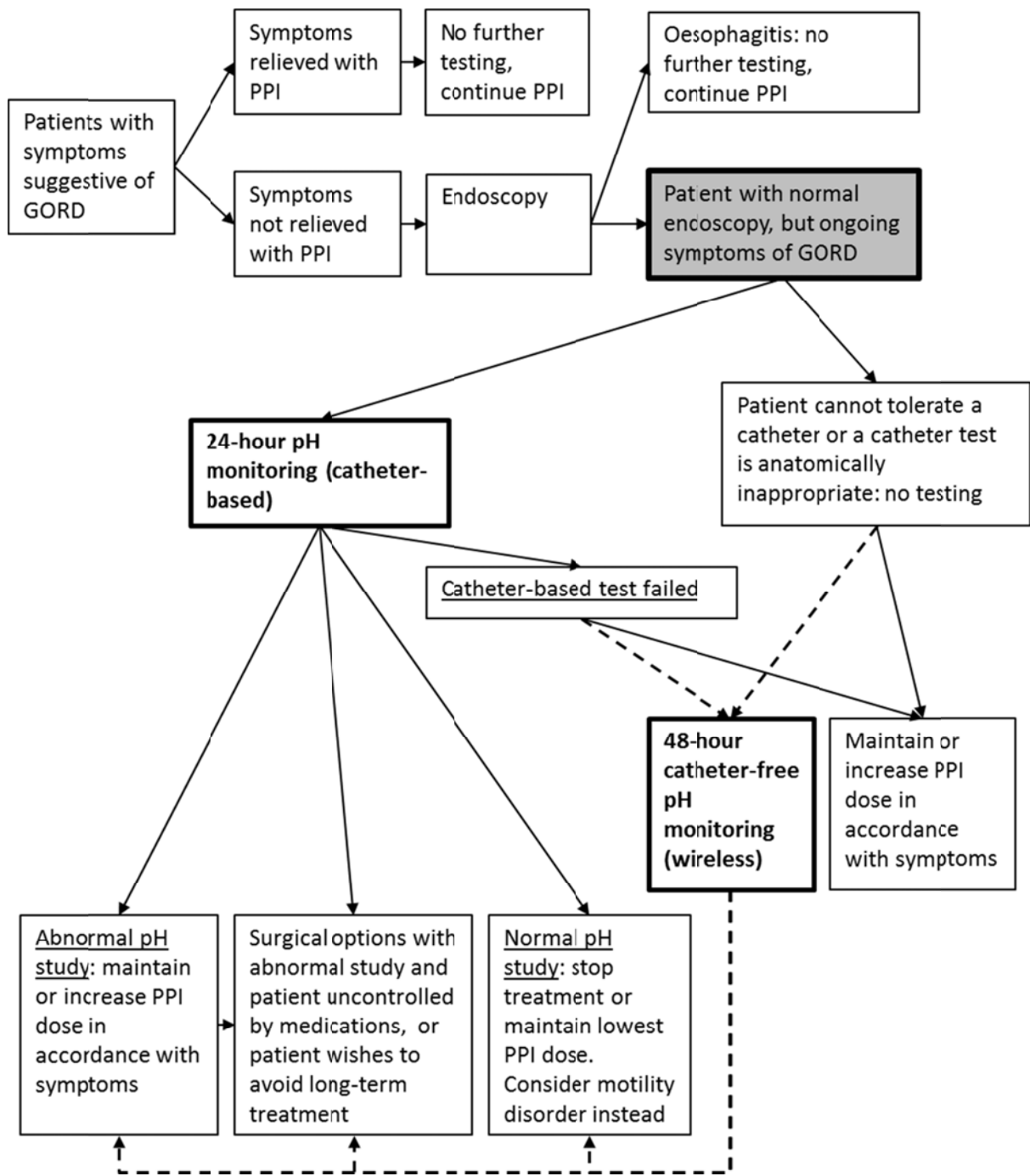
Figure 3   Clinical management algorithm for the proposed new intervention

Comparator

Comparators are usually selected by determining the diagnostic test that is most likely to be replaced, or added to, by the technology submitted for a new MBS item number. As catheter-free monitoring (the index test) is indicated for people that currently have no monitoring due to a failed catheter-based test or people for whom a catheter-based test is anatomically inappropriate, the appropriate comparator is 'no pH monitoring' and empirical therapy. However, for catheter-free monitoring to be implemented in this group, PASC determined that it should be demonstrated that the diagnostic sensitivity, specificity and safety of the proposed new test is at least equivalent to catheter-based monitoring. Thus, to establish diagnostic accuracy and safety, the selected comparator was catheter-based monitoring.

The reference standard

Oesophageal pH monitoring is used to diagnose GORD. In this assessment the accuracy of catheter-free monitoring in diagnosing GORD was benchmarked against the evidentiary standard as agreed to by PASC in August 2012—catheter-based pH monitoring. Patients usually have a GORD diagnosis if they respond to anti-reflux medication therapy, i.e. symptoms disappear completely or decrease by at least 50%. Clinical diagnosis based on response to treatment (which is the 'gold standard' in diagnosing GORD) has therefore also been used as a secondary reference standard.

Research questions

In the event that direct evidence was available to assess the safety, effectiveness and cost-effectiveness of catheter-free monitoring, the following research question was to be addressed by this evaluation:

- What is the safety, effectiveness and cost-effectiveness of catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD, compared with no pH monitoring, in patients who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate?

In the event that linked evidence (see 'Diagnostic assessment framework' on page 40) was the only evidence available to assess the safety, effectiveness and cost-effectiveness of catheter-free monitoring, the following research questions were also to be addressed:

Safety:

- Is catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD as safe as, or safer than, no pH monitoring?
- Is catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD as safe as, or safer than, catheter-based oesophageal pH monitoring?

Accuracy:

- What is the diagnostic accuracy of catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD compared with catheter-based oesophageal pH monitoring?

Effectiveness / patient management:

- Compared with no pH monitoring, does catheter-free (wireless) ambulatory oesophageal pH monitoring change the clinical management or treatment options available for patients with suspected GORD who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate?

Effectiveness / health outcomes, in case of a change in management:

- Do alterations in clinical management and treatment options have an impact on the health outcomes of patients with suspected GORD who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate?
  - Do alterations in clinical management and treatment options have an impact on the health outcomes of patients with GORD (true positives)?
  - Do alterations in clinical management and treatment options have an impact on the health outcomes of patients who were initially suspected of GORD but who received an incorrect diagnosis after catheter-free monitoring (false positives and false negatives)?

Cost-effectiveness:

- What is the cost-effectiveness of catheter-free (wireless) ambulatory oesophageal pH monitoring compared with no pH monitoring for patients with suspected GORD who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate?
- What is the cost-effectiveness of catheter-free (wireless) ambulatory oesophageal pH monitoring compared with catheter-based oesophageal pH monitoring in patients with symptoms of GORD?

Diagnostic assessment framework

This assessment uses the theoretical framework outlined in the MSAC *Guidelines for the assessment of diagnostic technologies* (MSAC 2005b).

This means that evidence of the clinical effectiveness of catheter-free monitoring requires either the first or, if not available, the second of the following:

- evidence of the effectiveness of catheter-free monitoring from high-quality comparative studies evaluating its use and subsequent treatment compared with no monitoring and treatment (direct evidence); randomised controlled trials (RCTs) provide the highest quality evidence for this comparison.

- evidence of treatment effectiveness from high-quality comparative studies evaluating the treatment for GORD, linked with applicable and high-quality evidence of the accuracy of catheter-free monitoring to diagnose GORD compared with catheter-based monitoring or no monitoring; this is called 'linked evidence'.

There was limited, low-quality direct evidence available that assessed the impact of catheter-free monitoring, so in this assessment the available direct evidence was supplemented with a linked evidence approach. This means that evidence from studies that report on diagnostic test performance (diagnostic accuracy), the impact on clinical decision-making and the impact of the treatment of diagnosed patients on health outcomes was narratively linked in order to infer the effect of the diagnostic test on patient health outcomes. For the last step of the linked analysis, to answer the questions, two separate searches were conducted (Table 6):

1. Evidence-based guidelines and/or systematic reviews providing evidence on treatment effectiveness in patients with GORD was collated, as this provides data on the health outcomes of those who are correctly diagnosed (i.e. true positives). For people initially suspected of GORD but who are eventually given an alternative diagnosis, it was assumed that their management/treatment would be optimised as a consequence of obtaining the correct diagnosis.

2. To address the implications associated with either inappropriately treating people who are incorrectly diagnosed with GORD (i.e. false positives) or not properly treating people who are incorrectly given an alternative diagnosis to GORD (i.e. false negatives), a second separate literature search was conducted.

Review of literature

Literature sources and search strategies

The medical literature was searched to identify relevant studies and reviews for the period between 2001 and September 2013, as catheter-free monitoring became available in 2001. Searches were conducted via the databases described in Table 4. Search terms are described in Table 5 and Table 6. Table 6 contains the search terms for the last step of the linked analysis.

Table 4　　　Electronic databases searched

Electronic database	Period covered
Cochrane Library – including, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, the Cochrane Central Register of Controlled Trials (CENTRAL), the Health Technology Assessment Database, the NHS Economic Evaluation Database	2001 – 9/2013
Current Contents	2001 – 9/2013
Embase	2001 – 9/2013
PubMed	2001 – 9/2013
Web of Science – Science Citation Index Expanded	2001 – 9/2013
Cinahl	2001 – 9/2013
Econlit	2001 – 9/2013
Scopus	2001 – 9/2013

Table 5　　　Search terms used

Element of clinical question	Search terms
Population	Gastroesophageal reflux [MeSH] OR Gastric Acid [MeSH] OR heartburn [MeSH] OR (GORD OR GERD OR Heartburn OR gastro oesophageal reflux disease OR gastro esophageal reflux disease OR gastroesophageal reflux disease
Intervention	((hydrogen-ion concentration [MeSH]) OR hydrogen-ion concentration OR pH) AND (Esophageal pH monitoring [MeSH] OR Gastric Acidity Determination /instrumentation [MeSH] OR "catheter free" OR catheterless OR wireless OR tubeless OR telemetry [MeSH] OR telemetry OR (radio transmit*) OR (radio transmis*) OR radiotransmit* OR radiotransmis* OR bravo)
Comparator (if applicable)	N/A

Element of clinical question	Search terms
Outcomes (if applicable)	N/A
Limits	2001–2013

MeSH = Medical Subject Heading, based on a Medline/PubMed platform

Table 6   Search terms used to identify studies for the last step of the linked analysis (health outcomes)

Element of clinical question	True positives	False negatives and false positives
Population	Gastroesophageal reflux [MeSH] OR Gastric Acid [MeSH] OR heartburn [MeSH] OR (GORD OR GERD OR heartburn OR "gastro oesophageal reflux disease" OR "gastro esophageal reflux disease" OR "gastroesophageal reflux disease")	(Gastroesophageal reflux [MeSH] OR Gastric Acid [MeSH] OR heartburn [MeSH] OR (GORD OR GERD OR heartburn OR "gastro oesophageal reflux disease" OR "gastro esophageal reflux disease" OR "gastroesophageal reflux disease")
Intervention	(Omeprazole OR lansoprazole OR dexlansoprazole OR esomeprazole OR pantoprazole OR rabeprazole OR ilaprazole OR "proton pump inhibitor" OR PPI OR "proton pump inhibitors") OR (cimetidine OR rantidine OR famotidine OR nizatidine OR (H2 receptor antagonist*) OR (H2 antagonist*))	(untreated OR "not treated" OR "no treatment" OR "delayed treatment" OR "non-optimal treatment" OR "inappropriate treatment") OR ("wrong diagnosis" OR "missed diagnosis" OR misdiagnosis) OR "false negative" OR "false negatives" OR "false reassurance" OR "false positive" OR "false positives" OR "false alarm" OR "Diagnostic Errors"[Mesh]
Comparator (if applicable)	-	-
Outcomes (if applicable)	-	-
	"Systematic review" OR guideline OR guidelines	-

MeSH = Medical Subject Heading, based on a Medline/PubMed platform

Selection criteria

In general, studies were excluded (Figures 4, 5 and 6) if they:

- did not provide information on the pre-specified target population

- did not address one of the pre-specified outcomes and/or provided inadequate data on these outcomes

- were in a language other than English and were of a lower level of evidence than the studies in English

- did not have an eligible study design

- used impedance monitoring as a comparator

- included a non-Bravo wireless pH-monitoring device as the index test.

If the same data were duplicated in multiple articles, only results from the most comprehensive or most recent article were included.

Search results

PRISMA flowchart
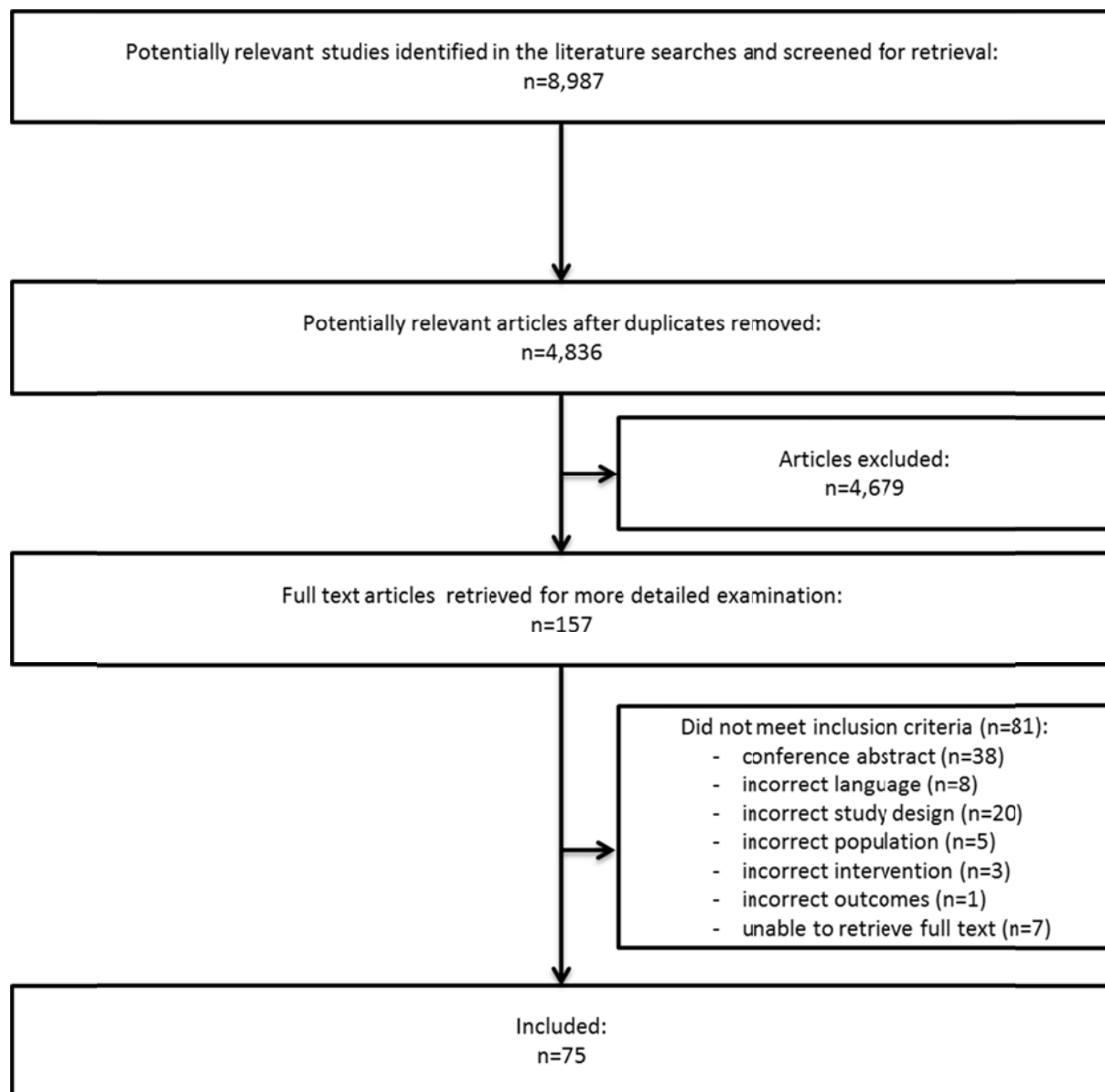
Figure 4   Summary of the process used to identify and select studies for the review (in first search)
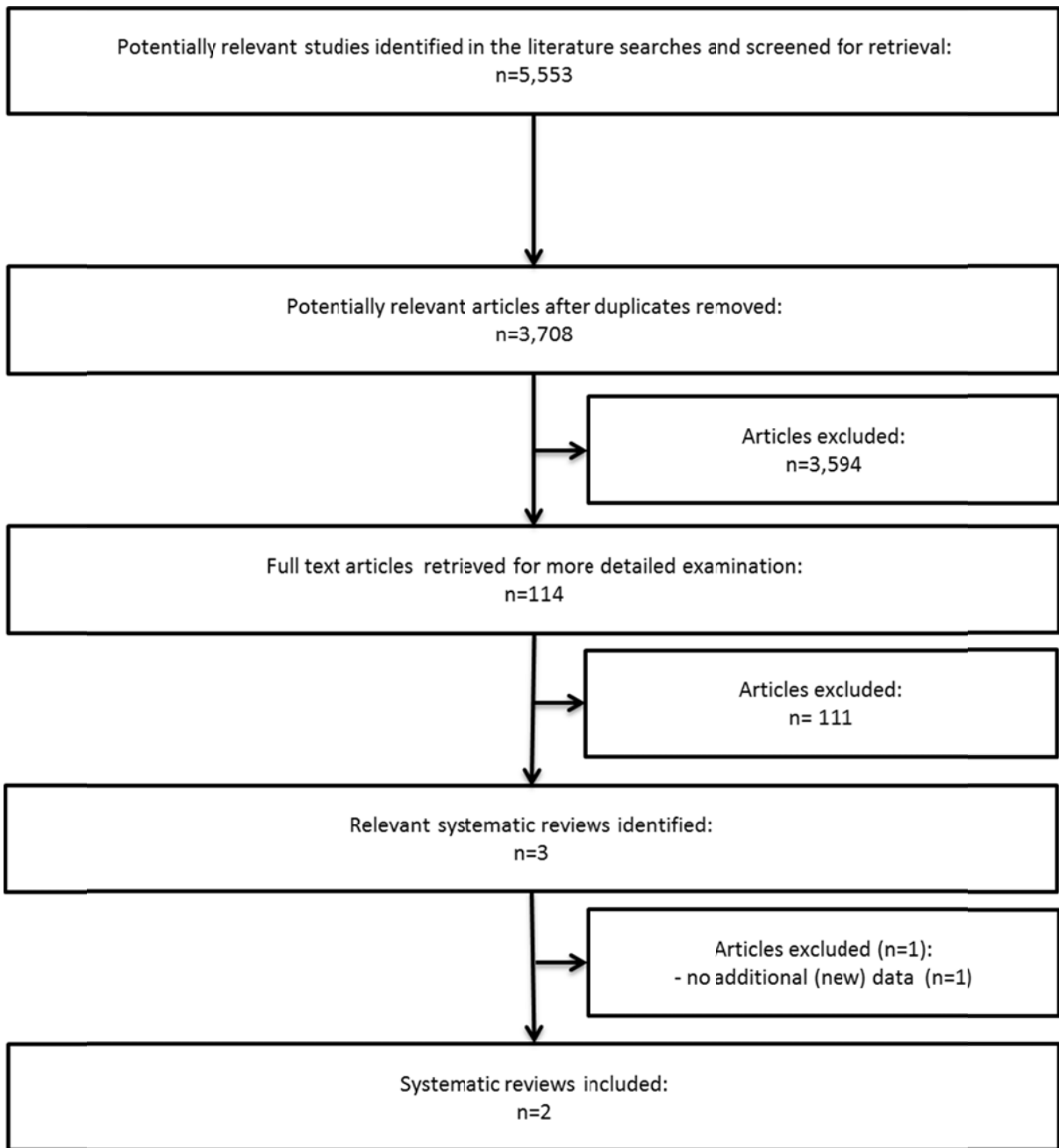
Figure 5   Summary of the process used to identify and select studies for the review (in search 'true positives')
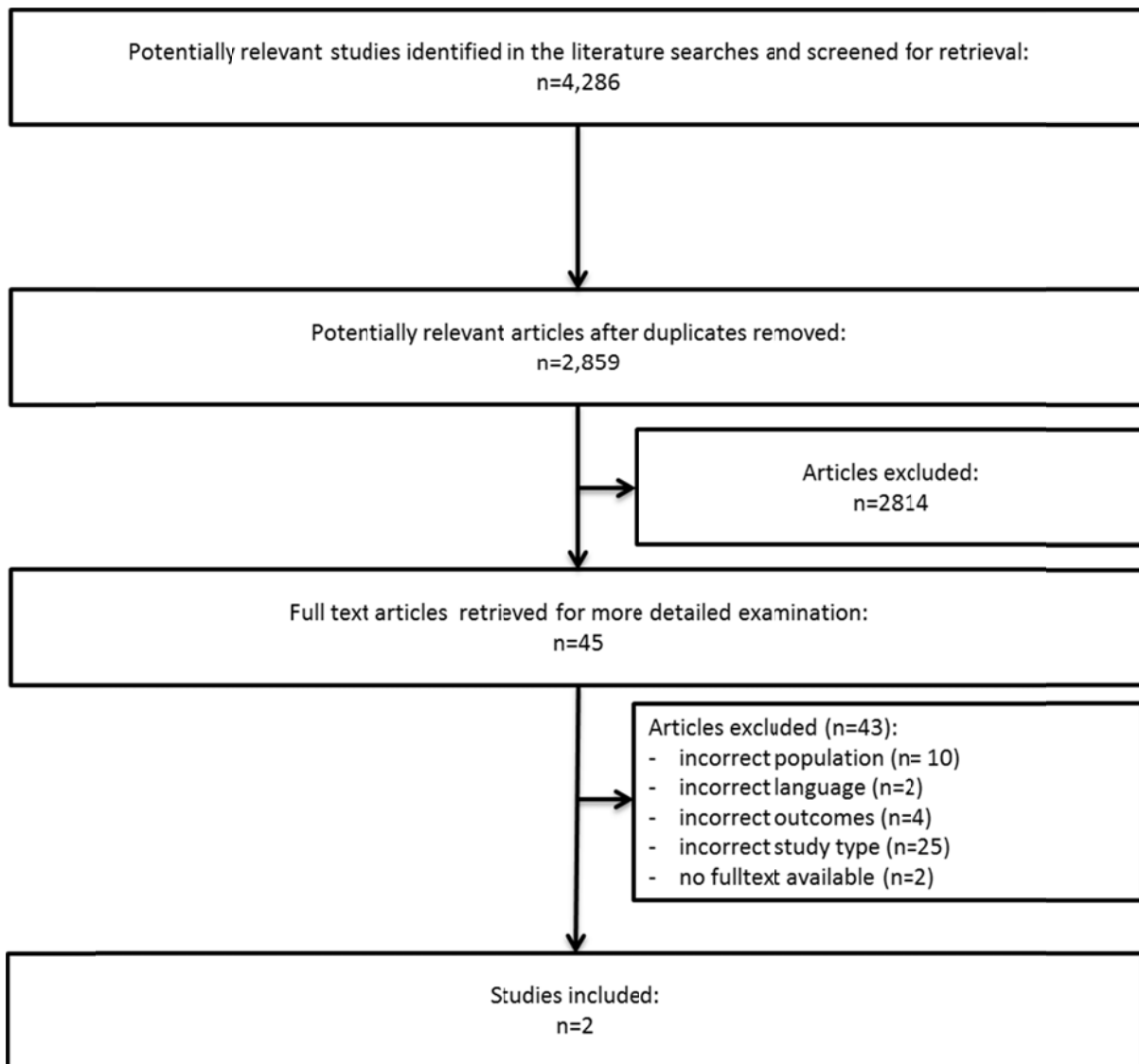
Figure 6   Summary of the process used to identify and select studies for the review (in search 'false negatives and false positives')

Data extraction and analysis

A profile of key characteristics was developed for each included study (see Appendix C). Each study profile described the level of evidence, design and quality of the study, authors, publication year, location, criteria for including/excluding patients, study population characteristics, type of intervention, comparator intervention and/or reference standard (where relevant), and outcomes assessed. Studies that could not be retrieved or that met the inclusion criteria but contained insufficient or inadequate data for inclusion are listed in Appendix D. Definitions of all technical terms and abbreviations are provided in the Glossary (page 23). Descriptive statistics were extracted or calculated for all safety and effectiveness outcomes in the individual studies.

Assessing diagnostic accuracy

To assess the diagnostic accuracy of catheter-free monitoring, where possible the sensitivity, specificity, negative and positive predictive values (NPV, PPV) and likelihood ratios of the tests were calculated with corresponding 95% confidence intervals (CIs). Data were extracted into a classic 2×2 table, in which the results of the index diagnostic test were cross-classified (Table 7) against the results of the reference standard (Armitage, Berry & Matthews 2002; Deeks 2001), and Bayes' Theorem was applied:

Table 7　　　Diagnostic accuracy data extraction for GORD

		Reference standard (catheter-*based* oesophageal pH monitoring or clinical diagnosis)		
		Disease +	*Disease –*	
Index test (catheter-*free* [wireless] ambulatory oesophageal pH monitoring)	*Test +*	true positive	false positive	Total test positive
	Test –	false negative	true negative	Total test negative
		Total with GORD	Total without GORD	

Primary measures

Test sensitivity was calculated as the proportion of people with GORD, as determined by catheter-based monitoring or clinical diagnosis, who had a positive test result on catheter-free monitoring:

Sensitivity (true positive rate) = number with true positive result / total with GORD

Test specificity was calculated as the proportion of people without GORD, as determined by catheter-based monitoring or clinical diagnosis, who had a normal test result on catheter-free monitoring:

Specificity (true negative rate) = number with true negative result / total without GORD

Where possible, the PPV and NPV were also provided. A PPV is the proportion of positive test results that are true positives, and the NPV is the proportion of negative test results that are true negatives. They reflect the probability that a positive (or negative) test result corresponds with having (or not having) GORD. However, the values depend on the prevalence of GORD.

PPV = number true positive / total test positive

NPV = number true negative / total test negative

When a 95%CI was not provided, it was calculated by exact binomial methods.

Summary measures

Positive and negative likelihood ratios (LRs) were reported if available. These ratios measure the probability of the test result in patients with GORD compared with those without GORD.

LR+ = sensitivity / 1–specificity

LR– = 1–sensitivity / specificity

A likelihood ratio of 1 means that the test does not provide any useful diagnostic information, whereas LR+ >5 and LR– <0.2 can suggest strong diagnostic ability (MSAC 2005a).

The summary receiver–operator characteristic curve plots the estimated sensitivity versus 1–specificity from different studies to produce a global measure of test accuracy.

All statistical calculations were undertaken using the biostatistical computer package Stata version 12.0.

Meta-analysis of diagnostic accuracy results was not conducted owing to the heterogeneous nature of the available evidence-base assessing the test performance of GORD; specifically, different thresholds for diagnosis were used and there were not enough studies reporting data within each threshold to stratify and meta-analyse the results by cut-point.

Meta-analysis of technical failure data was conducted using the metan command in Stata version 12.0 and specifying a random-effects model. Subgroup analyses were not performed.

A narrative meta-synthesis of the data was undertaken where meta-analysis could not be performed.

Appraisal of the evidence

Appraisal of the evidence was conducted in three stages:

Stage 1: Appraisal of the applicability and quality of individual studies included in the review (strength of the evidence).

Stage 2: Appraisal of the precision, size of effect and clinical importance of the results for primary outcomes in individual studies—used to determine the safety and effectiveness of the intervention.

Stage 3: Integration of this evidence for conclusions about the net clinical benefit of the intervention in the context of Australian clinical practice.

The evidence presented in the selected studies was assessed and classified using the dimensions of evidence defined by the National Health and Medical Research Council (NHMRC 2000).

These dimensions (Table 8) consider important aspects of the evidence supporting a particular intervention and include three main domains: strength of the evidence, size of the effect and relevance of the evidence. The first domain is derived directly from the literature identified as informing a particular intervention; the last two each requires expert clinical input as part of its determination.

Table 8   Evidence dimensions

Type of evidence	Definition
Strength of the evidence:	
Level	The study design used, as an indicator of the degree to which bias has been eliminated by design.[a]
Quality	The methods used by investigators to minimise bias within a study design.
Statistical precision	The p-value or, alternatively, the precision of the estimate of the effect. It reflects the degree of certainty about the existence of a true effect.
Size of effect	The distance of the study estimate from the 'null' value and the inclusion of only clinically important effects in the confidence interval.
Relevance of evidence	The usefulness of the evidence in clinical practice, particularly the appropriateness of the outcome measures used.

[a] See Table 9

Stage 1: strength of the evidence

The three sub-domains—level, quality and statistical precision—are collectively a measure of the strength of the evidence.

The 'level' of evidence reflects the effectiveness of a study design to answer a particular research question. Effectiveness is based on the probability that the design of the study has reduced or eliminated the impact of bias on the results. The NHMRC evidence hierarchy provides a ranking of various study designs ('levels of evidence') by the type of research question being addressed (Table 9).

Table 9   Designations of levels of evidence according to type of research question (including table notes)

Level	Intervention [a]	Diagnostic accuracy [b]
I [c]	A systematic review of level II studies	A systematic review of level II studies
II	A randomised controlled trial	A study of test accuracy with: an independent, blinded comparison with a valid reference standard,[d] among consecutive persons with a defined clinical presentation [e]
III-1	A pseudo-randomised controlled trial (i.e. alternate allocation or some other method)	A study of test accuracy with: an independent, blinded comparison with a valid reference standard,[d] among non-consecutive persons with a defined clinical presentation [e]
III-2	A comparative study with concurrent controls: ▪ non-randomised, experimental trial [f] ▪ cohort study ▪ case-control study ▪ interrupted time series with a control group	A comparison with reference standard that does not meet the criteria required for level II and III-1 evidence
III-3	A comparative study without concurrent controls: ▪ historical control study ▪ two or more single arm studies [g] ▪ interrupted time series without a parallel control group	Diagnostic case-control study [e]
IV	Case series with either post-test or pre-test/post-test outcomes	Study of diagnostic yield (no reference standard) [h]

Sources: Merlin et al. (2009); hierarchies adapted and modified from: NHMRC 1999; Bandolier 1999; Lijmer et al. 1999; Phillips et al. 2001

Explanatory notes:

[a] Definitions of these study designs are provided in NHMRC (2000; pp. 7–8) and in the Glossary accompanying Merlin et al. (2009).

[b] These levels of evidence apply only to studies assessing the accuracy of diagnostic or screening tests. To assess the overall effectiveness of a diagnostic test there also needs to be a consideration of the impact of the test on patient management and health outcomes (MSAC 2005; Sackett & Haynes 2002). The evidence hierarchy given in the 'Intervention' column should be used when assessing the impact of a diagnostic test on health outcomes relative to an existing method of diagnosis or comparator test(s). The evidence hierarchy given in the 'Screening' column should be used when assessing the impact of a screening test on health outcomes relative to no screening or alternative screening methods.

[c] A systematic review will only be assigned a level of evidence as high as the studies it contains, excepting where those studies are of level II evidence. Systematic reviews of level II evidence provide more data than the individual studies and any meta-analyses will increase the precision of the overall results, reducing the likelihood that the results are affected by chance. Systematic reviews of lower level evidence present results of likely poor internal validity and thus are rated on the likelihood that the results have been affected by bias, rather than whether the systematic review itself is of good quality. Systematic review quality should be assessed separately. A systematic review should consist of at least two studies. In systematic reviews that include different study designs, the overall level of evidence should relate to each individual outcome/result, as different studies and study designs might contribute to each different outcome.

[d] The validity of the reference standard should be determined in the context of the disease under review. Criteria for determining the validity of the reference standard should be pre-specified. This can include the choice of the reference standard(s) and its timing in relation to the index test. The validity of the reference standard can be determined through quality appraisal of the study (Whiting et al. 2003).

^e Well-designed population-based case-control studies (e.g. screening studies where test accuracy is assessed on all cases, with a random sample of controls) do capture a population with a representative spectrum of disease and thus fulfil the requirements for a valid assembly of patients. However, in some cases the population assembled is not representative of the use of the test in practice. In diagnostic case-control studies a selected sample of patients already known to have the disease is compared with a separate group of normal/healthy people known to be free of the disease. In this situation patients with borderline or mild expressions of the disease, and conditions mimicking the disease, are excluded, which can lead to exaggeration of both sensitivity and specificity. This is called spectrum bias or spectrum effect because the spectrum of study participants will not be representative of patients seen in practice (Mulherin & Miller 2002).

^f This also includes controlled before-and-after (pre-test/post-test) studies as well as adjusted indirect comparisons, i.e. utilising A vs B and B vs C to determine A vs C, with statistical adjustment for B.

^g Comparing single arm studies, i.e. case series from two studies. This would also include unadjusted indirect comparisons, i.e. utilising A vs B and B vs C to determine A vs C, but where there is no statistical adjustment for B.

^h Studies of diagnostic yield provide the yield of diagnosed patients, as determined by an index test, without confirmation of the accuracy of this diagnosis by a reference standard. These may be the only alternative when there is no reliable reference standard.

Note A: Assessment of comparative harms/safety should occur according to the hierarchy presented for each of the research questions, with the proviso that this assessment occurs within the context of the topic being assessed. Some harms (and other outcomes) are rare and cannot feasibly be captured within randomised controlled trials, in which case lower levels of evidence may be the only type of evidence that is practically achievable; both physical and psychological harms may need to be addressed by different study designs; harms from diagnostic testing include the likelihood of false positive and false negative results; harms from screening include the likelihood of false alarms and false reassurance results.

Note B: When a level of evidence is attributed in the text of a document, it should also be framed according to its corresponding research question, e.g. level II intervention evidence; level IV diagnostic evidence; level III-2 prognostic evidence.

Note C: Each individual study that is attributed a 'level of evidence' should be rigorously appraised using validated or commonly used checklists or appraisal tools to ensure that factors other than study design have not affected the validity of the results.

Individual studies assessing diagnostic effectiveness were graded according to pre-specified quality and applicability criteria (MSAC 2005), as shown in Table 10.

Table 10   Grading system used to rank included studies

Validity criteria	Description	Grading system
Appropriate comparison	Did the study evaluate a direct comparison of the index test strategy versus the comparator test strategy?	C1 direct comparison CX other comparison
Applicable population	Did the study evaluate the index test in a population that is representative of the subject characteristics (age and sex) and clinical setting (disease prevalence, disease severity, referral filter and sequence of tests) for the clinical indication of interest?	P1 applicable P2 limited P3 different population

Quality of study	Was the study designed to avoid bias?
	
	High quality = no potential for bias based on pre-defined key quality criteria
	
	Medium quality = some potential for bias in areas other than those pre-specified as key criteria
	
	Poor quality = poor reference standard and/or potential for bias based on key pre-specified criteria
		Q1 high quality
		
		Q2 medium
		
		Q3 poor reference standard
		
		    poor quality
		
		    or insufficient information

The appraisal of intervention studies pertaining to treatment safety and effectiveness was undertaken using the Downs and Black (1998) checklist, which was used for trials and cohort studies. Uncontrolled before-and-after case series are a poorer level of evidence with which to assess effectiveness. The quality of this type of study design was assessed according to a checklist developed by the UK National Health Service (NHS) Centre for Reviews and Dissemination (Khan et al. 2001). Studies of diagnostic accuracy were assessed using the QUADAS-2 quality assessment tool (Whiting et al. 2011), whereas systematic reviews included in the last step of the linked analysis were assessed with the PRISMA checklist (Liberati et al. 2009).

Stage 2: precision, size of effect and clinical importance

Statistical precision was determined using statistical principles. Small CIs and p-values give an indication as to the probability that the reported effect is real and not attributable to chance (NHMRC 2000). Studies need to be appropriately powered to ensure that a real difference between groups will be detected in the statistical analysis.

For intervention studies it was important to assess whether statistically significant differences between patients receiving catheter-free monitoring and those receiving no pH monitoring or catheter-based monitoring were also clinically important. The size of the effect needed to be determined, as well as whether the 95%CI included only clinically important effects.

The outcomes being measured in this report were assessed as to whether they were appropriate and clinically relevant (NHMRC 2000).

Stage 3: Assessment of the body of evidence

Appraisal of the body of evidence was conducted along the lines suggested by the NHMRC in their guidance on clinical practice guideline development (NHMRC 2008). Five components are considered essential by the NHMRC when judging the body of evidence:

1. the evidence-base—which includes the number of studies sorted by their methodological quality and relevance to patients;

2. the consistency of the study results—whether the better quality studies had results of a similar magnitude and in the same direction, i.e. homogeneous or heterogeneous findings;

3. the potential clinical impact—appraisal of the precision, size and clinical importance or relevance of the primary outcomes used to determine the safety and effectiveness of the test;

4. the generalisability of the evidence to the target population; and

5. the applicability of the evidence—integration of this evidence for conclusions about the net clinical benefit of the intervention in the context of Australian clinical practice.

A matrix for assessing the body of evidence for each research question, according to the components above, was used for this assessment (Table 11).

Table 11   Body of evidence matrix

Component	A	B	C	D
	Excellent	Good	Satisfactory	Poor
Evidence-base [a]	One or more level I studies with a low risk of bias or several level II studies with a low risk of bias	One or two level II studies with a low risk of bias or an SR or several level III studies with a low risk of bias	One or two level III studies with a low risk of bias, or level I or II studies with a moderate risk of bias	Level IV studies, or level I to III studies/SRs with a high risk of bias
Consistency [b]	All studies consistent	Most studies consistent and inconsistency may be explained	Some inconsistency reflecting genuine uncertainty around clinical question	Evidence is inconsistent
Clinical impact	Very large	Substantial	Moderate	Slight or restricted
Generalisability	Population(s) studied in body of evidence are the same as target population	Population(s) studied in body of evidence are similar to target population	Population(s) studied in body of evidence differ from target population for guideline but it is clinically sensible to apply this evidence to target population [c]	Population(s) studied in body of evidence differ from target population and it is hard to judge whether it is sensible to generalise to target population
Applicability	Directly applicable to Australian healthcare context	Applicable to Australian healthcare context with few caveats	Probably applicable to Australian healthcare context with some caveats	Not applicable to Australian healthcare context

Source: adapted from NHMRC (2008)

[a] Level of evidence determined from the NHMRC evidence hierarchy (Table 9)

[b] If there is only one study, rank this component as 'not applicable'.

[c] For example, results in adults that are clinically sensible to apply to children OR psychosocial outcomes for one cancer that may be applicable to patients with another cancer

SR = systematic review; several = more than two studies

Expert advice: Health Expert Standing Panel (HESP)

The Health Expert Standing Panel (HESP) has been established as a panel of the MSAC and is a pool of experts collated from various medical fields who are nominated by their associated professional body or by applicants. HESP members are engaged to provide practical, professional advice to evaluators that directly relates to each application and the service being proposed for the MBS. HESP members are not members of either MSAC or its subcommittees. Their role is limited to providing input and guidance to the assessment groups to ensure that the pathway is clinically relevant and takes into account consumer interests. HESP members' advice is used to inform the deliberations that MSAC presents to the Federal Minister for Health.

Results of assessment

Is it safe?

Summary—Is catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD as safe as, or safer than, no pH monitoring?

Only one study was identified that assessed the safety of catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD compared with no pH monitoring. This RCT found that chest pain was more likely to occur in those undergoing catheter-free monitoring (66%) than those undergoing no monitoring (28%) (RR=2.33, 95%CI 0.81, 6.76). Chest pain is likely to be attributed to the placement of the capsule.

Summary—Is catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD as safe as, or safer than, catheter-based oesophageal pH monitoring?

Twelve comparative studies were identified that reported on complications of catheter-free monitoring compared with catheter-based monitoring. Included were four RCTs, one pseudo-RCT, three cohort studies, two non-randomised controlled trials and two case-control studies.

No deaths or life-threatening events caused by pH monitoring were reported in the comparative studies. Chest pain was significantly more prevalent in patients undergoing catheter-free monitoring compared with catheter-based monitoring, as reported in four out of seven studies. In a high-quality RCT by Andrews et al. (2012), the median chest pain score during the pH test was 29 ± 4 mm on a 100-point VAS in the catheter-free monitoring group, compared with only 14 ± 3 mm in the catheter-based monitoring group (p=0.001). The difference is likely to be clinically meaningful. On the other hand, other adverse events such as nose and throat pain, dysphagia, eating and drinking difficulties, and headache were significantly more prevalent in catheter-based monitoring compared with catheter-free monitoring. Catheter-based monitoring tends to produce more overall discomfort than catheter-free monitoring and is more likely to have a negative impact on normal daily activities and work attendance.

In addition to the comparative studies, 29 non-comparative level IV studies were included to determine other complications and/or adverse events caused by catheter-free monitoring. The most common reported adverse events were chest pain and foreign-body sensation. Other complications included diminished appetite, extreme gagging, nausea, epistaxis, pharyngeal irritation, retrosternal discomfort on swallowing, throat pain, back pain, rash, mucosal abrasion with (minor) haemorrhage, capsule dislodgement, capsule detachment failure, laryngospasm, vasovagal reaction, poor capsule

tolerance with vomiting, and a dizzy spell during insertion.

In children, two oesophageal tears were reported, at least one due to a capsule release failure. Less-severe adverse events in children were overall discomfort, mild chest discomfort, coughing and dysphagia.

Finally, seven case reports were identified that reported some additional (rare) complications resulting from catheter-free monitoring, such as capsule dislodgment and/or aspiration, retention of the capsule in a colonic diverticulum and oesophageal perforation.

In conclusion, most complications reported were mild and did not require medical therapy. However, some (rare) complications can become severe when left untreated.

Studies were included to assess the safety of catheter-free (wireless) oesophageal pH monitoring according to criteria outlined *a priori* in Box 1.

Box 1   Inclusion criteria for studies assessing the safety of catheter-free (wireless) oesophageal pH monitoring

Selection criteria	Inclusion criteria
Population	1. Patients with symptoms of GORD for whom pH monitoring is indicated 2. Patients with symptoms of GORD in whom catheter-based pH monitoring has failed or in whom it is anatomically inappropriate Subgroups: adults versus children
Intervention	Catheter-free (wireless) ambulatory oesophageal pH monitoring
Comparators	1. No pH monitoring 2. Catheter-based oesophageal pH monitoring
Outcomes	Any adverse events arising from monitoring including: • oesophageal injury • chest pain • late detachment • stricture • naso-pharyngeal discomfort • nausea Any adverse events arising from (incorrect) treatment (as a result of no monitoring), such as: • worsening of or no decrease in reflux complaints • Barrett's oesophagus / oesophageal adenocarcinoma
Publication type	Randomised or non-randomised controlled trials, cohort studies, registers, case series, case reports or systematic reviews of these study designs

Search period	2001 – 5/2013
Language	Non-English language articles were excluded unless they provided a higher level of evidence than the English language articles identified

There were 48 studies that met the selection criteria that reported on the safety of catheter-free monitoring. There was 1 comparative study on the safety of (proximal) catheter-free monitoring compared with no monitoring, and 12 studies compared catheter-based monitoring with catheter-free monitoring in terms of safety. The remaining studies were case series or case reports. There were no deaths or life-threatening events reported as a direct result of either catheter-free or catheter-based monitoring.

Safety of catheter-free (wireless) pH monitoring

Safety results have been divided into the commonly reported outcomes of chest pain, foreign-body sensation, nose and throat pain, dysphagia, eating and drinking difficulties, impact on daily life, overall discomfort and satisfaction, and 'other'. Within each table of results per outcome, studies are listed in order of methodological quality and sample size. For the complete overview of safety data extracted per study, including non-significant results, see Appendix E (page 231).

Chest pain

One randomised blinded controlled trial (Francis et al. 2012) reported on chest discomfort during proximal (16 cm above the SCJ) catheter-free pH monitoring compared with no proximal monitoring (Table 12). All patients had a wireless pH-monitoring capsule positioned 6 cm proximal to the gastro-oesophageal junction (GOJ), and 48 hours later they were randomised into proximal capsule placement (i.e. another capsule placed 16 cm above the SCJ) or sham capsule placement. Patients were blinded regarding whether the second capsule was deployed and there were 11 patients in each group. In the intervention group 66% reported chest pain scores of 3 or greater on a 10-point VAS where higher scores indicate more pain, compared with only 28% in the control group (RR=2.33, 95%CI 0.81, 6.76). For higher levels of chest discomfort, 32% reported a chest pain score of ≥7 in the intervention group, compared with 8% in the sham group. Multivariate analysis showed an increased odds of severe chest pain in the intervention group when comparing subjects with the same chest pain score during the first monitoring period (adjusted OR=8.44; 95%CI 1.35, 52.6).

Four (three level II and one level III-2) intervention studies reported significantly more chest pain or oesophageal discomfort with catheter-free monitoring than with catheter-based monitoring in adults. In a high-quality RCT by Andrews et al. (2012), the median chest pain score during the pH test was 29 ± 4 mm on a 100-point VAS in the catheter-free monitoring group, compared with only 14 ± 3 mm in the catheter-based monitoring group (p=0.001) This difference is likely to be clinical meaningful. An RCT by Wong et al. (2005) reported 9/25 patients (36%) with chest discomfort with catheter-free monitoring, compared with only 2/25 patients (8%) in the catheter-based monitoring group (RR=4.5, 95%CI 1.08, 18.77, p=0.037, p=0.037). More chest discomfort (RR=1.95, 95%CI 1.47, 2.60) and chest pain (RR=1.76, 95%CI 1.2, 2.57) in the catheter-free monitoring group, relative to the catheter-based monitoring group, was also reported in a cohort study by Bradley et al. (2011). In one average-quality case-control study (Pandolfino et al. 2003) 34% of the patients undergoing catheter-free monitoring had oesophageal discomfort, compared with only 17% of those undergoing catheter-based monitoring. Three studies (Schneider et al. 2007; Grigolon et al. 2007; Wenner, Johnsson et al. 2007) did not report significant differences in chest pain or oesophageal discomfort.

Table 12   Chest discomfort or chest pain resulting from catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Wenner, Johnsson et al. (2007)	Level II Quality: 20/26 Randomised crossover trial	31 participants, catheter group 31 participants, wireless group	<u>Patients' experience, median 10 cm VAS (IQR)</u> Chest symptoms: 2.4 (0.3–5.9)	<u>Patients' experience, median 10 cm VAS (IQR)</u> Chest symptoms: 1.1 (0.3 – 2.9)	(0.084)[a]
Andrews et al. (2012)	Level II Quality: 19/26 Dual centre, randomised, non-blinded trial	43 participants, catheter group 43 participants, wireless group	<u>pH placement discomfort (100 mm VAS ± SE)</u> Chest: 14 ± 3 <u>pH-test discomfort (mm VAS ± SE)</u> Chest: 29 ± 4	<u>pH placement discomfort (100 mm VAS ± SE)</u> Chest: 13 ± 3 <u>pH-test discomfort (mm VAS ± SE)</u> Chest: 14 ± 3	(0.968)[b] (0.001)[b]
Wong et al. (2005)	Level II Quality: 16.5/26 Randomised controlled trial	25 participants, catheter group 25 participants, wireless group	Chest discomfort: 9/25 (36%)	Chest discomfort: 2/25 (8%)	4.5 [1.08, 18.77] (0.037)
Grigolon et al. (2007)	Level III-1 Quality: 13.5/26 Pseudo-randomised controlled trial	78 participants, catheter group 55 participants, wireless group	Chest discomfort: 17/55 (30.9%)	Not stated	-

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Bradley et al. (2011)	Level III-2 Quality: 15/26 Cohort study	106 participants, catheter group 234 participants, wireless group	Chest discomfort: 54/106 (51%) Chest pain: 35/106 (33%)	Chest discomfort: 61/234 (26%) Chest pain: 44/234 (19%)	1.95 [1.47, 2.60] (<0.001) 1.76 [1.2, 2.57] (0.009)
Pandolfino et al. (2003)	Level III-2 Quality: 11.5/26 Case-control study	30 participants, catheter group 29 participants, wireless group	Oesophageal discomfort: 10/29 (34.5%) Severe discomfort leading to capsule removal: 2/29 (6.9%) Moderate chest pain: 4/29 (13.8%)	Oesophageal discomfort: 5/30 (16.7%)	2.07 [0.80, 5.32] (<0.05)
Schneider et al. (2007)	Level III-2 Quality: 9/26 Non-randomised controlled trial	78 participants, wireless group 55 participants, catheter group	Oesophagus discomfort: 14/78 (17.9%)	Oesophagus discomfort: 18/55 (32.7%)	0.55, [0.30, 1.01] (NS)

NS = not significant

a Analysis was done using the Mann-Whitney U-test.

b Analysis was done using the Student t-test.

Chest pain was reported in 21/31 non-comparative studies and was the most reported adverse event resulting from catheter-free monitoring. In the studies that reported chest pain or discomfort as an adverse event, rates varied between 0.27% and 45%, with a median of 17.19%. Severe or intolerable chest pain was reported in 6 of the 21 studies and occurred in 0.52% to 5.13% of cases, with a median of 1.44% (see Table 75 in Appendix E).

Children had six times more chest pain undergoing catheter-free monitoring (39%) compared with catheter-based monitoring (6%; p=0.02) (Croffie et al. 2007). Of the four non-comparative studies on complications of catheter-free monitoring in children, one study (Gunnarsdottir, Stenstrom & Arnbjornsson 2007, 2008) reported a case of mild chest discomfort (Table 13 and Table 14).

Table 13   Chest pain or chest discomfort resulting from catheter-free monitoring compared with catheter-based monitoring in children

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Croffie et al. (2007)	Level II Quality: 14.5/25 Randomised controlled trial	16 participants, catheter group 18 participants, wireless group	Chest pain 7/18 (39%)	Chest pain 1/16 (6%)	6.22 [0.86, 45.25] (0.02)

Table 14   Complications and adverse events resulting from catheter-free (wireless) oesophageal pH monitoring in children

Study	Level and quality	Population	Complications per procedure
Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008)	Level IV Q1 Case series	58 children	1/58 procedures 1 mild chest discomfort

Foreign-body sensation

A foreign-body sensation during monitoring was reported in one comparative study and nine non-comparative studies. Bradley et al. (2011) reported significantly more chest foreign-body sensation in patients with catheter-free monitoring compared with catheter-based monitoring (Table 15).

In a large case series by Crowell et al. (2009) the foreign-body sensation rate was as high as 50% (147/294) (Table 16). Other smaller studies reported a lower rate of globus sensation (i.e. feeling a lump or foreign body in their throat): the proportion of procedures associated with this sensation ranged from 0.22% to 22.97%, with a median of 5.9%.

Table 15   A foreign-body sensation resulting from catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Bradley et al. (2011)	Level III-2 Quality: 15/26 Cohort study	106 participants, catheter group 234 participants, wireless group	Foreign body chest: 66/106 (62)% Foreign body throat: 22/106 (21%)	Foreign body chest: 80/234 (34%) Foreign body throat: 51/234 (22%)	1.82 [1.44, 2.30] (<0.001) 0.95 [0.61, 1.48] (0.887)

Table 16   A foreign-body sensation resulting from catheter-free (wireless) oesophageal pH monitoring in adults

Study	Level and quality	Population	Chest discomfort / complications per procedure
Crowell et al. (2008)	Level IV Q1 Case series	294 patients	147/294 procedures 147 foreign-body sensation
Tseng et al. (2005)	Level IV Q1 Retrospective case series	190 procedures, 186 patients	Most patients noted a foreign-body sensation

Remes-Troche et al. (2005)	Level IV Q1 Prospective case series	77 patients	11/77 procedures 11 foreign-body sensation
Prakash et al. (2006)	Level IV Q2 Retrospective case series	452 patients	1/452 procedures (but this only gives the major side effects of the patients who requested removal of the capsule) 1 severe foreign-body sensation
Domingues, Moraes-Filho & Domingues (2011)	Level IV Q2 Retrospective case series	74 patients	17/74 procedures 17 mild foreign-body sensation
Garrean et al. (2008)	Level IV Q2 Post-test case series	60 patients	1/60 procedures 1 globus sensation
Belafsky et al. (2004)	Level IV Q2 Prospective case series	46 patients	4/46 procedures 4 foreign-body sensation
Pandolfino et al. (2006)[a]	Level IV Q2 Case-control study	10 control subjects and 10 patients with GORD	7/20 7 mild chest discomfort or foreign-body sensation

[a] This study is not a case series but is still included as level IV safety evidence because it does not report a(n) (appropriate) comparison regarding complications or adverse events.

Nose and throat pain

In contrast to the findings for chest pain and discomfort, other complications and discomforts from pH monitoring were more common with catheter-*based* monitoring compared with catheter-*free* monitoring. Five studies reported on nose and/or throat pain or discomfort, and in all studies this was significantly more severe in patients receiving catheter-based monitoring (Table 17). Wenner, Johnsson et al. (2007) reported a median score of 0.2 cm on a 10 cm VAS scale for nose/throat symptoms in patients undergoing catheter-free monitoring, compared with a score of 6.5 cm for patients in the catheter-based monitoring group (p<0.0001)—a clinically significant difference. Andrews et al. (2012) noted a greater difference in nasal discomfort than throat discomfort between the two pH-monitoring technologies, with results favouring catheter-free monitoring. Nasal pain was halved in patients receiving catheter-free monitoring compared with catheter-based monitoring in Wong et al. (2005). The differences in throat pain and discomfort between the two pH tests in Wong et al. (2005), Pandolfino et al. (2003) and Schneider et al. (2007) was much greater than in Andrews et al. (2012), with relative risk reductions favouring catheter-free monitoring that ranged from 67% to 96%. Croffie et al. (2007) reported that throat pain

in children was also significantly more common during catheter-based monitoring (Table 18).

Table 17   Nose and throat pain or discomfort resulting from catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Wenner, Johnsson et al. (2007)	Level II Quality: 20/26 Randomised crossover trial	31 participants, catheter group 31 participants, wireless group	Patients' experience, median cm VAS (IQR) Nose/throat symptoms: 0.2 (0.0–1.9)	Patients' experience, median cm VAS (IQR) Nose/throat symptoms: 6.5 (1.5–8.0)	(<0.0001) [a]
Andrews et al. (2012)	Level II Quality: 19/26 Dual centre, randomised, non-blinded trial	43 participants, catheter group 43 participants, wireless group	pH placement discomfort (mm VAS ± SE) Nasal: 6 ± 2 Throat: 32 ± 4 pH-test discomfort (mm VAS ± SE) Nasal: 10 ± 3 Throat: 19 ± 4	pH placement discomfort (mm VAS ± SE) Nasal: 36 ± 4 Throat: 37 ± 3 pH-test discomfort (mm VAS ± SE) Nasal: 39 ± 3 Throat: 43 ± 4	(<0.001) [b] (0.317) [b] (<0.001) [b] (<0.001) [b]
Wong et al. (2005)	Level II Quality: 16.5/26 Randomised controlled trial	25 participants, catheter group 25 participants, wireless group	Nose pain: 8/25 (32%) Runny nose: 13/25 (52%) Nose bleeding: 1/25 (4%) Throat pain: 4/25 (16%) Throat discomfort: 12/25 (48%)	Nose pain: 15/25 (60%) Runny nose: 24/25 (96%) Nose bleeding: 0/25 (0%) Throat pain: 12/25 (48%) Throat discomfort: 23/25 (92%)	0.53 [0.28, 1.03] (0.047) 0.54 [0.37, 0.8] (0.001) (1.0) 0.33 [0.12, 0.89] (0.032) 0.52 [0.34, 0.8] (0.001)
Pandolfino et al. (2003)	Level III-2 Quality: 11.5/26 Case-control study	30 participants, catheter group 29 participants, wireless group	Throat discomfort: 4/29 (13.8%)	Throat discomfort: 22/30 (73.3%)	0.19 [0.07, 0.48] (<0.001)
Schneider et al. (2007)	Level III-2 Quality: 9/26 Non-randomised controlled trial	55 participants, catheter group 78 participants, wireless group	Throat discomfort: 3/78 (3.8%)	Throat discomfort: 52/55 (94.5%)	0.04 [0.01, 0.12] (0.001)

[a] Analysis was done using the Mann-Whitney U-test.
[b] Analysis was done using the Student t-test.

Table 18   Nose and throat pain or discomfort resulting from catheter-free monitoring compared with catheter-based monitoring in children

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Croffie et al. (2007)	Level II Quality: 14.5/25 Randomised controlled trial	16 participants, catheter group 18 participants, wireless group	Throat pain 8/18 (44%)	Throat pain 15/16 (94%)	0.47 [0.28, 0.81] (0.001)

Three case series reported throat pain during catheter-free monitoring (Table 19). Throat pain was reported in 9.3% of cases in the study by Karamanolis et al. (2012), 5.0% of cases by Garrean et al. (2008) and 32.8% of patients by Sofi et al. (2011). No case series mentioned nose pain or discomfort as a side effect from catheter-free monitoring; however, some cases of epistaxis were reported (Table 19).

Table 19   Throat pain and epistaxis resulting from catheter-free (wireless) oesophageal pH monitoring in adults

Study	Level and quality	Population	Throat pain per procedure
Karamanolis et al. (2012)	Level IV Q1 Post-test case series	32 patients	3/32 procedures 3 throat pain
Garrean et al. (2008)	Level IV Q2 Post-test case series	60 patients	3/60 procedures 3 throat pain
Sofi et al. (2011)	Level IV Q3 Retrospective case series	58 patients	19/58 procedures 19 sore throat
des Varannes et al. (2005)[a]	Level IV Q1 In-subject simultaneous recording study	40 patients	1/40 procedures 1 epistaxis
Marchese et al. (2006)	Level IV Q1 Prospective case series	39 patients	2/39 procedures 2 mild epistaxis
Belafsky et al. (2004)	Level IV Q2 Prospective case series	46 patients	2/46 procedures 2 epistaxis

[a] This study is not a case series but is still included as level IV safety evidence because it does not report a(n) (appropriate) comparison regarding complications or adverse events.

Dysphagia

Dysphagia (difficulty swallowing) was also mentioned as a common adverse effect from pH monitoring. Francis et al. (2012) compared proximal wireless pH monitoring with no proximal pH monitoring (48 hours after the start of distal wireless pH monitoring) during a blinded RCT. They reported increased odds of dysphagia (adjusted OR=14.3; 95%CI 2.12, 96.6; p=0.006) and odynophagia (painful swallowing) (adjusted OR=49.5; 95%CI 4.70, 520; p=0.001) in the intervention group, i.e. those receiving catheter-free monitoring.

Wenner, Johnsson et al. (2007) compared swallow-induced symptoms in patients receiving catheter-free and catheter-based monitoring, and reported a median score of 5.2 on a 10 cm VAS scale (IQR=2.1–6.7) in the catheter-based monitoring group, compared with a score of only 2.1 (IQR=0.5–5.0) in the catheter-free monitoring group (p=0.033). Swallowing difficulties were a lot more common in patients undergoing catheter-based, as opposed to catheter-free, monitoring and were approximately halved in a cohort study by Bradley et al. (2011). This study also reported a significant difference in painful swallowing (Table 20).

In children the results were similar, with one high-quality case series by Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008) reporting that 3 out of 58 (5.2%) children undergoing catheter-free monitoring had dysphagia (see Appendix E).

Table 20   Dysphagia resulting from catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Dysphagia, wireless pH monitoring	Dysphagia, catheter-based pH monitoring	RR [95%CI] (p-value)
Wenner, Johnsson et al. (2007)	Level II Quality: 20/26 Randomised crossover trial	31 participants, catheter group 31 participants, wireless group	Patients' experience, median cm VAS (IQR) Swallow-induced symptoms: 2.1 (0.5–5.0)	Patients' experience, median cm VAS (IQR) Swallow-induced symptoms: 5.2 (2.1–6.7)	0.033 [a]
Bradley et al. (2011)	Level III-2 Quality: 15/26 Cohort study	106 participants, catheter group 234 participants, wireless group	Difficulty swallowing: 98/234 (42%) Painful swallowing: 103/234 (44%)	Difficulty swallowing: 96/106 (91%) Painful swallowing: 66/106 (62%)	0.46 [0.39, 0.54] (<0.001) 0.71 [0.57, 0.87] (0.002)

[a] Analysis was done using the Mann-Whitney U-test.

Eating and drinking difficulties

Six studies reported that pH monitoring affects diet or appetite (Table 21 and

Table 22). Andrews et al. (2012) reported that patients in the catheter-free monitoring group were more likely to be able to eat and drink without difficulties than those undergoing catheter-based monitoring, with scores of 75 ± 5 and 51 ± 4 on a VAS-scale[2], respectively (p<0.001). Catheter-based monitoring affected the diet of 47% of patients in the study by Pandolfino et al. (2003). In comparison, only 3% of the patients receiving catheter-free monitoring in the same study had their diet affected (p<0.001). Wenner, Johnsson et al. (2007) reported a median food intake interference of 1.4 with the catheter-free (wireless) test and 4.0 with the catheter-based test (on a 10-point VAS, where 0 was defined as no interference at all and 10 the worst possible interference). Although a difference of this size would likely be considered clinically significant, the sample was too small for this to be so. Bradley et al. (2011) also did not report any significant differences regarding early fullness or diminished appetite comparing the two pH tests. In children the results were again consistent, with patients undergoing catheter-free monitoring shown to have more appetite than those undergoing catheter-based monitoring (Croffie et al. 2007;

Table 22).

One non-comparative study, a large case series, addressed eating and drinking during catheter-free monitoring, and reported that 82/294 (27.9%) of the patients had a diminished appetite (Crowell et al. 2008; see Appendix E).

Table 21   Eating and drinking difficulties resulting from catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Wenner, Johnsson et al. (2007)	Level II Quality: 20/26 Randomised crossover trial	31 participants, catheter group 31 participants, wireless group	Patients' experience, median cm VAS (IQR) Interference food intake: 1.4 (0.2–4.9)	Patients' experience, median cm VAS (IQR) Interference food intake: 4.0 (1.8–7.0)	(0.056) [a]
Andrews et al. (2012)	Level II Quality: 19/26 Dual centre, randomised, non-blinded trial	43 participants, catheter group 43 participants, wireless group	pH-test discomfort (mm VAS ± SE) Eating/drinking: 75 ± 5	pH-test discomfort (mm VAS ± SE) Eating/drinking: 51 ± 4	(<0.001) [b]
Grigolon et al. (2007)	Level III-1 Quality: 13.5/26 Pseudo-randomised	78 participants, catheter group 55 participants, wireless group	Limitation of food intake (score 0–3) 0.4 ± 0.1	Limitation of food intake (score range 0–3) 0.9 ± 0.1	(<0.05)[b]

[2] A score of 100 means that their eating and drinking behaviour was completely normal.

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
	controlled trial				
Bradley et al. (2011)	Level III-2 Quality: 15/26 Cohort study	106 participants, catheter group 234 participants, wireless group	Early fullness: 68/234 (29%) Diminished appetite: 56/234 (24%)	Early fullness: 23/106 (22%) Diminished appetite: 34/106 (32%)	1.33 [0.89, 2.02] (0.188) 0.75 [0.52, 1.07] (0.144)
Pandolfino et al. (2003)	Level III-2 Quality: 11.5/26 Case-control study	30 participants, catheter group 29 participants, wireless group	Affected diet: 1/29 (3%)	Affected diet: 14/30 (47%)	0.07 [0.01, 0.53] (<0.001)

[a] Analysis was done using the Mann-Whitney U-test, 10 cm VAS.
[b] Analysis was done using the Student t-test, 100 mm VAS.

Table 22   Eating and drinking difficulties resulting from catheter-free monitoring compared with catheter-based monitoring in children

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	p-value
Croffie et al. (2007)	Level II Quality: 14.5/25 Randomised controlled trial	16 participants, catheter group 18 participants, wireless group	Appetite on scale of 1–5: mean 3.54	Appetite on scale of 1–5: mean 2.72	0.029[a]

[a] Analysis was done using the Student t-test.

Impact on daily life

Oesophageal pH monitoring can have an impact on daily activities. All seven studies reported a significant difference in interference in daily activities when comparing the two different tests, favouring catheter-free monitoring (Table 23 and Table 24). On a scale of 0–10—where 0 means no interference, and 10 means maximum interference, with physical activities—Wenner, Johnsson et al. (2007) reported a median score of 0.6 (IQR=0.2–2.7) with catheter-free monitoring, compared with a score of 5.0 (IQR=2.6–8.5) in the catheter-based monitoring group (Mann-Whitney U test p<0.0001). Interference with normal daily life was also assessed in this study and similar results were obtained. Andrews et al. (2012) assessed whether patients with wireless monitoring were better able to maintain their usual activities than patients with a catheter inserted, and found this was the case. Grigolon et al. (2007), Pandolfino et al. (2003) and Schneider et al. (2007) reported similar results (Table 23).

Gillies et al. (2005) reported that catheter-based monitoring seemed to affect work attendance significantly more than wireless monitoring: 82% of patients in the catheter-based group did not attend work, compared with only 14% in the catheter-free group (p<0.0001); and interference with work activities was also significantly less with catheter-free monitoring than with catheter-based monitoring in the study by Wenner, Johnsson et al. (2007).

Table 23   Impact on daily life from catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] (p-value)
Wenner, Johnsson et al. (2007)	Level II Quality: 20/26 Randomised crossover trial	31 participants, catheter group 31 participants, wireless group	Patients' experience, median cm VAS (IQR) Interference work activities: 0.3 (0.0–1.4) Interference physical activities: 0.6 (0.2–2.7) Interference normal daily life: 0.7 (0.2–3.4)	Patients' experience, median cm VAS (IQR) Interference work activities: 6.8 (1.7–9.0) Interference physical activities: 5.0 (2.6–8.5) Interference normal daily life: 5.7 (2.3–8.0)	(0.005) [a] (<0.0001) [a] (<0.0001) [a]
Andrews et al. (2012)	Level II Quality: 19/26 Dual centre, randomised, non-blinded trial	43 participants, catheter group 43 participants, wireless group	pH-test discomfort (mm VAS ± SE) Ability usual activities: 92 ± 2	pH-test discomfort (mm VAS ± SE) Ability usual activities: 75 ± 5	(<0.001) [b]
Grigolon et al. (2007)	Level III-1 Quality: 13.5/26 Pseudo-randomised controlled trial	78 participants, catheter group 55 participants, wireless group	Limitation of daily activities (score 0–3, mean ± SEM) 0.2 ± 0.1	Limitation of daily activities (score 0–3, mean ± SEM) 1.2 ± 0.1	(<0.001) [b]
Gillies et al. (2007)	Level III-2 Quality: 13.5/26 Non-randomised controlled trial	100 participants, catheter group 85 participants, wireless group (but 100 studies)	Impact on work attendance: 14/100 (14%)	Impact on work attendance: 82/100 (82%)	0.17 [0.10, 0.28] (<0.0001)
Pandolfino et al. (2003)	Level III-2 Quality: 11.5/26 Case-control study	30 participants catheter group, 29 participants, wireless group	Change daily routine: 0/29 (0%) Affected activity: 0%	Change daily routine: 11/30 (37%) Affected activity: 60%	(<0.001)
Schneider et al. (2007)	Level III-2 Quality: 9/26 Non-randomised controlled trial	55 participants, catheter group 78 participants, wireless group	Reduced activities: 5/78 (6.4%)	Reduced activities: 55/55 (100%)	0.12 [0.05, 0.29] (0.0001)

[a] Analysis was done using the Mann-Whitney U-test.

b Analysis was done using the Student t-test.

In children the results were similar to those described in adults (Table 24).

Table 24   Impact on daily life from catheter-free monitoring compared with catheter-based monitoring in children

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	p-value
Croffie et al. (2007)	Level II Quality: 14.5/25 Randomised controlled trial	16 participants catheter group 18 participants wireless group	Mean activity on scale of 1–5: 3.66	Mean activity on scale of 1–5: 2.33	0.001 [a]

[a] Analysis was done using the Student t-test.

Overall discomfort and satisfaction

Overall placement discomfort was not significantly different between the two pH monitoring tests, when Andrews et al. (2012) and Grigolon et al. (2007) reported on this outcome.

Overall discomfort during the period of monitoring was higher when a catheter was used. Overall discomfort scores on a VAS—where 0 means no discomfort and 100 means worst possible discomfort—ranged from 21 to 29 in patients undergoing catheter-free monitoring and from 33 to 51 in those undergoing catheter-based monitoring (Wenner, Johnsson et al. 2007; Andrews et al. 2012; Grigolon et al. 2007; p-values ≤0.012). Gilles et al. (2007) reported that 27% of the patients undergoing catheter-free monitoring experienced no overall discomfort, compared with only 3% undergoing catheter-based monitoring (p<0.0001). In the same patient population 23% in the catheter-free group and 45% in the catheter-based group reported discomfort scores >5 on a scale of 0–10. This difference was not considered statistically significant. Overall satisfaction scores were higher in patients who underwent catheter-free monitoring than catheter-based monitoring (Pandolfino et al. 2003; Sweis et al. 2009; Table 25).

Table 25   Overall discomfort and satisfaction from catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] p-value
Wenner, Johnsson et al. (2007)	Level II Quality: 20/26 Randomised crossover trial	31 participants, catheter group 31 participants, wireless group	Patients' experience, median cm VAS (IQR) All adverse symptoms: 2.1 (0.5–4.6)	Patients' experience, median cm VAS (IQR) All adverse symptoms: 5.1 (2.0–6.6)	(<0.001) [a]

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	RR [95%CI] p-value
Andrews et al. (2012)	Level II Quality: 19/26 Dual centre, randomised, non-blinded trial	43 participants, catheter group 43 participants, wireless group	pH placement discomfort (mm VAS ± SE) Overall: 29 ± 4 pH-test discomfort (mm VAS ± SE) Overall: 26 ± 4	pH placement discomfort (mm VAS ± SE) Overall: 33 ± 4 pH-test discomfort (mm VAS ± SE) Overall: 39 ± 4	(0.406) [b] (0.012) [b]
Grigolon et al. (2007)	Level III-1 Quality: 13.5/26 Pseudo-randomised controlled trial	78 participants, catheter group 55 participants, wireless group	Mean placement discomfort (mm VAS ± SE) 29 ± 4 Discomfort during whole test (mm VAS ± SE) 22 ± 3 None to mild discomfort: 41/55 (74.5%)	Mean placement discomfort (mm VAS ± SE) 32 ± 3 Discomfort during whole test (mm VAS ± SE) 37 ± 3 None to mild discomfort: 40/78 (51.3%)	NS (<0.001) [b] (<0.05) [b]
Sweis et al. (2009)	Level III-2 Quality: 15/26 Cohort study	110 participants, catheter group 134 participants, wireless group	Overall satisfaction: 4.4/5	Overall satisfaction: 3.5/5	(<0.001) [b]
Gillies et al. (2007)	Level III-2 Quality: 13.5/26 Non-randomised controlled trial	100 participants, catheter group 85 participants, wireless group (but 100 studies)	73/100 (73%) overall discomfort Median discomfort: 3 Discomfort score >5: 23/100 (23%)	97/100 (97%) overall discomfort Median discomfort: 5 Discomfort score >5: 45/100 (45%)	0.85 [0.68, 1.07] (<0.0001) [a] 0.60 [0.39, 0.94]
Pandolfino et al. (2003)	Level III-2 Quality: 11.5/26 Case-control study	30 participants, catheter group 29 participants, wireless group	Satisfaction (the lower the better): 0.8 ± 0.1	Satisfaction (the lower the better): 1.9 ± 0.2	(<0.001) [b]

NS = not significant

[a] Analysis was done using the Mann-Whitney U-test.

[b] Analysis was done using the Student t-test.

In children the results were similar to those found in adults (Table 26).

Table 26   Overall discomfort and satisfaction from catheter-free monitoring compared with catheter-based monitoring in children

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	p-value
Croffie et al. (2007)	Level II Quality: 14.5/25 Randomised controlled trial	16 participants, catheter group 18 participants, wireless group	Overall satisfaction on scale of 1–5: 4.31	Overall satisfaction on scale of 1–5: 3.11	0.003 [a]

[a] Analysis was done using the Student t-test.

Other reported side effects and complications

Wong et al. (2005) reported other fairly prevalent side effects such as coughing (20% in catheter-free and 28% in catheter-based monitoring) and headaches (20% in catheter-free and 56% in catheter-based monitoring, p=0.009). Furthermore, 19% experienced abdominal pain and 18% nausea, and 6% had to vomit during the wireless test, compared with 14%, 20% and 6%, respectively, in the catheter group in the study by Bradley et al. (2011); but these results were not significantly different. In the study by Pandolfino et al. (2003) 9% of patients undergoing catheter-free monitoring reported disrupted sleep, compared with 30% undergoing catheter-based monitoring; this difference was not considered statistically significant (see Appendix E).

In the non-comparative studies, other adverse events or complications resulting from wireless pH monitoring were extreme gagging (1/190); coughing (6/58); nausea (5/77); pharyngeal irritation (3/39); back pain (5.0% to 6.3%); rash (1/203); mucosal abrasion with minor haemorrhage (1/30); capsule dislodgement in the mouth, pyriform sinus or stomach (3/76); detachment failure of the capsule (1/66); laryngospasm (2/46); vasovagal reaction (1/46); poor tolerance with vomiting (2/40); and dizzy spell during insertion (1/40).

Four case series reported adverse events of catheter-free monitoring in children. There were two oesophageal tears reported, at least one due to a capsule release failure. A less severe adverse event reported was coughing (1/58).

Case reports were also included but provided less information than case series because it was impossible to determine the denominator, i.e. how many patients received the test and were at risk of harm but did not have any adverse effects. Case reports only provide descriptive information as to the possible types of adverse events and can be useful for describing rare complications. The case reports' results are shown in Appendix E. Serious

complications of catheter-free monitoring mentioned in these case reports and not anywhere else were: a retention of the capsule in a colonic diverticulum for over 2 years, a capsule aspiration followed by decreased oxygen saturation, an oesophageal perforation, and different capsule dislodgements—one in the left pyriform sinus and one in the mainstem bronchus. Capsule aspirations and dislodgements can lead to pain, breathing difficulties and choking, depending on where the capsules dislodge, and can therefore become a severe complication. Oesophageal perforations can lead to further complications such as permanent damage to the oesophagus, abscess formation, infection in and around the lungs, and can progress to shock and even death if remain untreated.

Is it effective?

Direct evidence of diagnostic effectiveness

Does catheter-free monitoring improve health outcomes?

Summary—What is the effectiveness of catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD in patients who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate? Does catheter-free monitoring improve health outcomes?

No studies were identified that assessed the direct health impact of catheter-free monitoring versus no monitoring in the selected study population. However, one matched-pairs study found that just slightly more patients who were monitored by the wireless system had an improvement or disappearance of the principal symptom, compared with those monitored by the catheter-based system (73% in the wireless group versus 69% in the catheter-based group (n=51 in each group)). A before-and-after study reported that out of 26 patients who received medical therapy or conservative advice after a catheter-free monitoring test, 9 patients had a good or moderate improvement in symptoms.

In a meta-analysis, catheter-free monitoring in adults was found to have over three times the risk (RR=3.3, 95%CI 1.63, 6.81) of having technical problems compared with catheter-based monitoring. A similar comparative risk of technical problems was seen in a trial of pH monitoring in adults and children combined (RR=2.87, 95%CI 1.47, 5.62).

Most studies (15/20) reported only minimal day-to-day variability in oesophageal acid exposure across the 2 days of monitoring, suggesting that catheter-free monitoring is a reliable means of monitoring pH levels.

Studies were included to assess the effectiveness of catheter-free (wireless) oesophageal pH monitoring according to the criteria outlined *a priori* in Box 2.

Box 2   Eligibility criteria for identification of studies relevant to an assessment of effectiveness of catheter-free (wireless) ambulatory oesophageal pH monitoring

Selection criteria	Inclusion criteria
Population	Patients with symptoms of GORD in whom catheter-based pH monitoring has failed or in whom it is anatomically inappropriate
Intervention	Catheter-free (wireless) ambulatory oesophageal pH monitoring
Comparators	No pH monitoring
Outcomes	Patient relevant outcomes: quality of life, reduction in progression to Barrett's oesophagus and/or oesophageal cancer, symptom resolution or symptom reduction

Publication type	Randomised or non-randomised controlled trials, cohort studies or systematic reviews of these study designs
Search period	2001 – 5/2013
Language	Non-English language articles were excluded unless they provided a higher level of evidence than the English language articles identified

No studies were identified that assessed the direct health impact of catheter-free monitoring versus no monitoring in the selected study population.

Two studies reported on the reduction in GORD symptoms after catheter-free monitoring. However, these studies did not include patients who had either failed a catheter-based test or in whom catheter-based monitoring was anatomically inappropriate. Furthermore, they did not have 'no pH monitoring' as a comparator, and therefore did not meet the inclusion criteria specified *a priori*. However, in the absence of more relevant data, results from these studies are presented.

A matched-pairs retrospective cohort study by Grigolon et al. (2011) evaluated the usefulness of a 96-hour[3] catheter-free monitoring test in patients' clinical management, comparing outcomes of 51 patients studied with the wireless technique and 51 patients studied with the traditional catheter-based 24-hour test. The percentage of patients with an improvement in or disappearance of the principal GORD symptom was 73% in the catheter-free group and 69% in the catheter-based group. The average satisfaction score was 7.0 in the catheter-free group compared with 6.5 in the catheter-based group, on a scale of 0 to 10; as expected, the average score for satisfaction increased progressively (p<0.001) in the two groups as their symptoms improved.

Sweis et al. (2011) assessed the clinical impact of prolonged, 96-hour catheter-free pH monitoring at clinical follow-up (6–36 months) in patients with negative 24-hour catheter-based monitoring referred for a second opinion due to ongoing symptoms suggestive of GORD. Follow-up data from at least 6 months after initiation of therapy were available for 33 out of 38 patients (87%). Twenty-six of these patients received medical therapy or conservative advice, with median follow-up at 24 months. Nine of these patients had a moderate or good improvement in symptoms and 17 had a poor outcome or were lost to follow-up. In these groups all 9 (100%) in the improvement group and 9 out of the 17 (53%)

---

[3] This is longer than the 'usual' 48-hour monitoring period

in the poor-outcome group had a GORD diagnosis on the basis of pathologic acid exposure or a positive association with GORD symptoms during catheter-free monitoring (p=0.361).

Technical efficacy

Studies were included to assess the technical efficacy of catheter-free (wireless) oesophageal pH monitoring according to the criteria outlined in Box 3.

Box 3   Eligibility criteria for identification of studies relevant to an assessment of the technical efficacy of catheter-free (wireless) ambulatory oesophageal pH monitoring

Selection criteria	Inclusion criteria
Population	1. Patients with symptoms of GORD in whom catheter-based pH monitoring has failed or in whom it is anatomically inappropriate 2. Patients with symptoms of GORD for whom pH monitoring is indicated
Intervention	Catheter-free (wireless) ambulatory oesophageal pH monitoring
Comparators	1. No pH monitoring 2. Catheter-based oesophageal pH monitoring
Outcomes	Technical efficacy: operative success and early detachment or hardware malfunctions
Publication type	Randomised or non-randomised controlled trials, cohort studies, registers, case series, case reports or systematic reviews of these study designs
Search period	2001 – 5/2013
Language	Non-English language articles were excluded unless they provided a higher level of evidence than the English language articles identified

Catheter-free monitoring versus catheter-based monitoring

Nine studies were included that reported not only the technical efficacy of catheter-free monitoring, but also catheter-based monitoring failure rates. There were no statistically significant differences in failure rates or recording efficacy when the two pH-measuring systems were compared. However, all eight studies conducted in adults reported a higher percentage of technical problems with the new catheter-free system compared with the traditional catheter-based system (Table 27). No differences were found between the two systems when used in children (Croffie et al. 2007).

A meta-analysis was conducted to determine the overall risk of technical failures of both systems when used in adults (Figure 7). The risk of having technical problems with catheter-free monitoring was over three times higher than with catheter-based monitoring (RR_p=3.3; 95%CI 1.63, 6.81; I^2=0%, p=0.906; k = 8 studies). Results of studies included in this meta-analysis were fairly consistent.

Other (non-comparative) technical efficacy results of catheter-free monitoring have been divided and described per outcome; for the complete overview of non-comparative technical efficacy data extracted by each study, see Appendix F . Studies are listed in order of level of evidence, quality and sample size.

Table 27　Failure rates / recording efficacy of catheter-free monitoring compared with catheter-based monitoring

Study	Level and quality	N	Catheter-free monitoring	Catheter-based monitoring	RR [95%CI]
des Varannes et al. (2005)	Level: II Quality: 20/26 In-subject simultaneous recording study	40 patients Adults	7/40 procedures (17.5%) 1 dysfunction of the capsule 2 insertion intolerances 1 detachment failure from the delivery system 1 recording failure 2 early detachments	1/40 procedures (2.5%) 1 recording failure	7.00 [0.90, 54.32]
Wenner, Johnsson et al. (2007)	Level: II Quality: 20/26 Randomised crossover trial	35 patients Adults	2/35 procedures (5.7%) 2 technical problems	0/35 procedures (0%)	5.00 [0.25, 100.53]
Andrews et al. (2012)	Level: II Quality: 19/26 Dual centre, randomised, non-blinded trial	43 patients in the wireless group 43 patients in the catheter group Adults	5/43 procedures (12%) 2 failed capsule calibrations 2 early detachments 1 insertion intolerance	3/43 procedures (7%) 1 equipment malfunction after the patient dropped the recorder 2 catheter intolerances with early probe removal	1.67 [0.42, 6.54]
Azzam et al. (2012)	Level: II Quality: 17/26 In-subject simultaneous recording study	25 patients Adults	1/25 procedures (4.0%) 1 early detachment	0/25 procedures (0%)	3.00 [0.13, 70.30]
Wong et al. (2005)	Level: II Quality: 16.5/26 Randomised controlled trial	25 patients in the wireless group 25 patients in the catheter group Adults	2/25 procedures (8.0%) 2 failures of transnasal insertion	1/25 procedures (4.0%) 1 catheter intolerance	2.00 [0.19, 20.67]

Grigolon et al. (2007)	Level: III-1 Quality: 13.5/26 Pseudo-randomised controlled trial	55 patients in the wireless group 78 patients in the catheter group Adults	1/55 procedures (1.8%) 1 early detachment 95.5% (91.5–97) recording time	0/78 procedures (0%) 98% (94–99) recording time	4.23 [0.18, 102.00]
Bradley et al. (2011)	Level: III-2 Quality: 15/26 Cohort study	234 patients in the wireless group 106 patients in the catheter group Adults	5/234 procedures (2.1%) 2 early detachments 3 endoscopic removals due to severe discomfort	1/106 procedures (0.9%) 1 catheter intolerance	2.26 [0.27, 19.15]
Hakanson et al. (2009)	Level: III-2 Quality: 14/26 Case-control study with in-subject simultaneous recording	53 volunteers and 55 patients Adults	13/108 procedures (12.0%) 5 attachment failures 4 early detachments 3 immediate detachments 1 recording failure	2/108 procedures (1.9%) 1 sensor displacement in the stomach 1 catheter intolerance	6.50 [1.5, 28.12]
Croffie et al. (2007)	Level: II Quality:14.5/26 Randomised controlled trial	30 patients in wireless + catheter group, 18 patients in wireless group and 18 in catheter group Children	2/48 procedures (4.2%) 1 recording failure 1 early detachment	2/48 procedures (4.2%) 1 recording failure 1 catheter intolerance	1.00 [0.15, 6.81]

Study Name	Year	Failures intervention	N intervention	Failures comparator	N comparator		RR (95%CI)	% Weight
Des Varannes et al.	2005	7	40	1	40		7.00 (0.90, 54.32)	12.20
Wenner et al.	2007	2	35	0	35		5.00 (0.25, 100.53)	5.69
Andrews et al.	2012	5	43	3	43		1.67 (0.42, 6.54)	27.39
Azzam et al.	2012	1	25	0	25		3.00 (0.13, 70.30)	5.15
Wong et al.	2005	2	25	1	25		2.00 (0.19, 20.67)	9.39
Grigolon et al.	2007	1	55	0	78		4.23 (0.18, 102.00)	5.06
Bradley et al.	2011	5	234	1	106		2.26 (0.27, 19.15)	11.24
Hakanson et al.	2009	13	108	2	108		6.50 (1.50, 28.12)	23.88
Overall (I^2 = 0.0%, p=0.906)							3.33 (1.63, 6.81)	100.00

NOTE: Weights are from random effects analysis

Figure 7   Meta-analysis of technical efficacy of catheter-free monitoring compared with catheter-based monitoring

<u>Insertion and placement of the catheter-free (wireless) system</u>

Six studies reported on the rate of problems associated with the insertion and placement of both the catheter-free and catheter-based monitoring systems (Table 27). On average, problems occurred more often during insertion or placement of the catheter-free (median 7.5% procedures, range 0%–12%) than the catheter-based monitoring system (median 0.9% procedures, range 0%–4%) (Table 27). Reported insertion and placement problems with the catheter-based system were catheter intolerance or sensor displacement. Appendix F outlines the errors during insertion and placement of the wireless system, which were reported in 24 studies for a total of 1,933 patients. The most common errors during insertion and placement of the wireless system were:

- an error in placement: either in the cardia or in the wrong place on the oesophagus wall (63–73 times).[4]

- attachment failure: the capsule does not attach (properly) to the oesophageal mucosa (28–38 times).[4]

- detachment failure: the capsule fails to deploy from the delivery system (6–10 times).

- insertion failure: the inability to insert the system, either due to intolerance or anatomical abnormalities (9 times).

- immediate detachment: detachment of the capsule from the oesophageal wall immediately after insertion, sometimes followed by aspiration of the capsule (4 times).

- failed capsule calibration or device malfunction: inability to correctly calibrate the capsule before the monitoring period, or malfunction of the device when inserting (2 and 2 times, respectively).

Most insertion and placement errors that occur with wireless monitoring are bothersome and inconvenient for the patient, since they have to return for another try or have the catheter inserted twice in the same sitting. However, insertion failures due to intolerance or anatomical abnormalities would occur in both catheter-free and catheter-based monitoring situations, because both systems are inserted using a catheter. The only difference would be

---

[4] Park et al. (2013) reported 10 cases of 'attachment failures or misplacements'. Since it is only known that each complication occurred 0–10 times, this was included as a range.

that the catheter could be inserted transnasally or transorally in catheter-free monitoring, whereas it has to be inserted transnasally in catheter-based monitoring.

Technical failures during data recording

The only complications reported in catheter-based monitoring in comparative studies during data recording were recording failures (recorded once in a child and once in an adult) (Croffie et al. 2007; des Varannes et al. 2005).

Technical failures during the (usually 48-hour) catheter-free monitoring period in adults were reported in 36 comparative and non-comparative studies with a total of 3,629 patients. The most commonly reported complication influencing technical efficacy was early capsule detachment (reported 133 times). When the capsule detaches from the oesophageal mucosa early, it passes through the stomach, which leads to unreliable data. The commonly reported errors leading to incomplete data were:

- early capsule detachment (133 times)

- incomplete data capture (58 times)

- recording failure or receiver malfunction (11 times)

- transmission failure (7 times)

- errors in data retrieval (4 times)

- poor data reception (3 times).

Furthermore, 'technical failures' were reported in 7 cases and the capsule malfunctioned in 2 patients. One patient lost his/her data recorder and therefore no analysis was possible (see Table 27 and Appendix F).

In children technical errors occurring during monitoring were similar to those with adults. Incomplete data capture was reported in 21/289 children by Cabrera et al. (2011). Furthermore, early detachment occurred in 5 children (Cabrera et al. 2011; Croffie et al. 2007), and there were 4 shorter monitoring periods of unknown cause, 2 recording failures and 1 transmission failure during the wireless monitoring period (see Table 27 and Appendix F).

Day-to-day variability in oesophageal acid exposure

With catheter-free monitoring it is possible to monitor pH for more than 24 hours: in most cases there is a 48-hour monitoring period. There is a concern that sedation used during insertion of the system may influence the results of the first 24 hours of monitoring, and oesophageal acid exposure has been reported to exhibit day-to-day variability (Wiener et al. 1988). If this is the case, monitoring for 48 hours instead of 24 hours, as with traditional catheter-based monitoring, would give truer and more-accurate results, and this may have clinical advantages. Twenty studies reported on the day-to-day variability in oesophageal acid exposure with catheter-free monitoring in adults (see Table 78 in Appendix G).

Two large case series of 124 and 77 patients reported significantly more reflux episodes on the second day of monitoring (Chander et al. 2012; Remes-Troche et al. 2005). Three other quality case series of 203, 148 and 26 patients reported more reflux on day 1 (Bechtold et al. 2007; Bhat, McGrath & Bielefeldt 2006; Turner et al. 2007). The remaining 14 studies did not report a significant difference in oesophageal acid exposure when comparing the first day with the second day of monitoring.

In children one out of three studies (n=44) reporting on day-to-day variability noted significantly more reflux on day 1 of monitoring (Hochman & Favaloro-Sabatier 2005). However, in the higher quality RCT and the case series with 27 and 58 included children, respectively, there were no significant differences in reflux reported when comparing the first 24 to the second 24 hours (Croffie et al. 2007; Gunnarsdottir, Stenstrom & Arnbjornsson 2007). See Table 79 in Appendix G for the full results.

Linked evidence

Is it accurate?

Summary—What is the diagnostic accuracy of catheter-free (wireless) ambulatory oesophageal pH monitoring for GORD compared with catheter-based oesophageal pH monitoring?

There was no consensus in the literature regarding what cut-offs should be used when pH monitoring is used to diagnose GORD. In the highest quality study available, a diagnosis of GORD was given if patients had a pH <4 for more than, or equal to, 4.4% of the time. Using these criteria, the catheter-free monitoring test had reasonable sensitivity (86.4%) and specificity (77.8%) when catheter-based monitoring was used as a reference standard. The two other studies that used this reference standard reported similar accuracy results, with slightly lower sensitivity and slightly higher specificity (variable cut-off points).

Five diagnostic case-control studies used clinical diagnosis as the reference standard. The study with the largest patient population reported test sensitivity between 59% and 88% and test specificity between 75% and 96%, depending on which cut-off value was used (between 1.9% and 4.4% of the time that patients had pH <4). The remaining studies used thresholds between 4.4% and 5.3% for 'proximal' oesophageal pH monitoring. The sensitivity and specificity in these studies varied from 67% to 100% and from 66% to 100%, respectively. No studies compared the accuracy of both catheter-free and catheter-based monitoring against the reference standard of clinical diagnosis.

Nine studies of adults reported on the time that the oesophagus was exposed to acid, rather than using a diagnostic cut-off, during pH monitoring. In four of the studies both tests were conducted in the same patient population, and therefore variations in results can be attributed with more certainty to the pH-monitoring method, rather than being associated with the use of two different samples of patients. However, the results were mixed, with two studies reporting that catheter-based monitoring detected more acid exposure time and one study reporting that catheter-free monitoring detected more acid exposure (Azzam et al. 2012). The remaining five studies reported no significant differences between the two measurement methods. One study reported on oesophageal acid exposure in children, finding significantly more reflux with wireless monitoring compared with catheter-based monitoring (p=0.01). Four studies reported not only on acid exposure time concordance, but also on variation in the number of reflux events. Three of the four studies—two level II and one level III-2—reported significantly more reflux events with the catheter-based monitoring system.

Two comparative diagnostic yield studies reported that significantly more patients were diagnosed with catheter-free monitoring compared with catheter-based monitoring (p<0.001). One reported advantage of catheter-free monitoring is that it allows a longer monitoring period than catheter-based monitoring,

i.e. 48 rather than 24 hours. Six studies reported that the additional day of monitoring increased diagnosed yield with a median of 7.8%.

Studies were included to assess the accuracy of catheter-free (wireless) oesophageal pH monitoring according to criteria outlined in Box 4.

Box 4   Inclusion criteria for identification of studies relevant to an assessment of the accuracy of catheter-free (wireless) ambulatory oesophageal pH monitoring

Selection criteria	Inclusion criteria
Population	Patients with symptoms of GORD for whom pH monitoring is indicated
Intervention	Catheter-free (wireless) ambulatory oesophageal pH monitoring
Reference standard (1)	Catheter-based oesophageal pH monitoring
Reference standard (2)	Clinical diagnosis
Outcomes	- Sensitivity - Specificity - Positive/negative predictive value - Level of agreement (concordance of data) - Diagnostic yield
Publication type	All study designs listed in the 'Diagnostic accuracy' column of Table 9
Search period	2001 – 5/2013
Language	Non-English language articles were excluded unless they provided a higher level of evidence than the English language articles identified

Only eight studies met the inclusion criteria for assessing the accuracy of catheter-free monitoring and provided sufficient data to determine its sensitivity and specificity. In three studies the reference standard used was catheter-based monitoring. Among them, two level II studies (a randomised crossover trial and an in-subject simultaneous recording study) and one level III-3 diagnostic accuracy study (a case-control study with in-subject simultaneous recording) were identified. All three studies with catheter-based monitoring as a reference standard scored high on flow and timing—the flow of patients through the study, and the timing of the index test(s) and reference standard—mostly because the index test and reference standard were performed almost simultaneously. Patient selection had an unknown to high risk of bias, depending on how patients enrolled and if there was the possibility of bias. The risk of bias regarding the index test and the reference standard varied

depending on how the results were interpreted. The results of these studies are presented in Table 28. All three studies included patients who were suspected of GORD and in whom pH monitoring was indicated, although one study (Hakanson et al. 2009) also included 45 volunteers—people without a history of GORD or any other pathology of the upper gastrointestinal tract.

The highest quality study available (level II diagnostic evidence) indicated that a catheter-free monitoring test with a cut-off point of 4.4% of the time that pH <4 had good sensitivity and specificity (86.4% and 77.8%, respectively). The accuracy of the two other studies was in the same range, although their sensitivity was slightly lower and their specificity slightly higher. There is no consensus on which cut-off value should be used, so the cut-off values for diagnosing GORD in both catheter-free and catheter-based monitoring are different in all studies, making it difficult to interpret the accuracy data.

Five studies used clinical diagnosis as a reference standard (Table 29). These were all diagnostic case-control studies (level III-3). This means that most of those included in the patient groups were GORD patients who had responded to therapy, because response to PPI therapy is a clear indication of having GORD and is widely used to diagnose the disease. Furthermore, the control groups consisted of patients who did not have GORD symptoms, so the applicability of the results based on these populations is questionable, since they do not match the target population, i.e. patients suspected of GORD who do not respond to PPIs. Most of these studies had a high risk of bias for patient selection when quality was assessed using the QUADAS-2 tool, and a medium to high risk of bias for flow and timing of the studies.

Different cut-off values were used for diagnosis, varying from 1.9% to 10.6% of the time that oesophageal pH <4. Wenner, Johansson et al. (2007) and Wenner et al. (2008) used multiple cut-off values (from 1.9% to 4.4%) to determine the optimal value for diagnosis, and Bansal et al. (2009) used two different cut-off values for catheter-free monitoring: 4.4% for pH monitoring more proximal in the oesophagus and 10.6% for more distal pH monitoring (1 cm above the GOJ). Since every study used different cut-off values, they were considered too heterogeneous in their methods to meta-analyse.

The study with the largest sample (Wenner, Johansson et al. 2007; Wenner et al. 2008) reported sensitivity values between 59% and 88% and specificity values between 75% and 96%, depending on which cut-off value was used. The remaining studies used cut-off values between 4.4% and 5.3% for 'proximal' oesophageal pH monitoring, i.e. with the capsule

placed approximately 5 cm above the LOS or 6 cm above the SCJ, respectively. The sensitivity and specificity in these studies varied from 67% to 100% and from 66% to 100%, respectively.

Whereas most studies use the 'total time that pH <4 in the oesophagus' to establish a positive or negative test result, some studies also use a 'worst day analysis'—they only take the data from the day with the most gastro-oesophageal reflux (Pandolfino et al. 2003; Scarpulla et al. 2007; Sweis et al. 2011). This can be either day 1 or day 2 of monitoring, and they then use similar cut-off values. Pandolfino et al. (2003) reported high sensitivity and specificity values with this type of analysis: 83.8% and 84.5%, respectively.

Table 28   Diagnostic accuracy of catheter-free monitoring compared with catheter-based monitoring for diagnosing GORD in adults

Study	Evidence level and quality	Population	Index test & cut-off point	Reference standard & cut-off point	Sensitivity % [95%CI]	Specificity % [95%CI]	PPV % [95%CI]	NPV % [95%CI]	LR+	LR–
des Varannes et al. (2005) In-subject simultaneous recording study	Level: II Quality: Patient selection: ? Index test: ☺ Reference standard: ? Flow and timing: ☺ C1 P2	36 subjects suspected of GORD	Bravo pH monitoring (48-hour) Time under pH 4 >2.9%	Catheter-based monitoring (24-hour) Time under pH 4 >4.2%	78.6% [0.49–0.94]	89.5% [0.65–0.98]	84.6% [0.54–0.97]	85.0% [0.61–0.96]	7.49	0.24
Wenner, Johnsson et al. (2007) Randomised crossover trial	Level: II Quality: Patient selection: ☺ Index test: ☺ Reference standard: ? Flow and timing: ☺ C1 P2	31 subjects suspected of GORD	Bravo pH monitoring (48-hour) Time under pH 4 >4.4%	Catheter-based monitoring (24-hour) Time under pH 4 >3.4%	86.4% [0.64–0.96]	77.8% [40.1–0.96]	90.5% [0.68–0.98]	70.0% [0.35–0.92]	3.89	0.17
Hakanson et al. (2009) Case-control study with in-subject simultaneous recording	Level: III-3 Quality: Patient selection: ? Index test: ☹ Reference standard: ? Flow and timing: ☺ C1 P2	92 subjects (45 healthy volunteers and 47 patients)	Bravo pH monitoring (48-hour) Time under pH 4 >1.9%	Catheter-based monitoring (24-hour) Time under pH 4 >4.0%	73.6% [0.59–0.84]	92.3% [0.78–0.98]	92.9% [0.79–0.98]	72.0% [0.57–0.83]	9.56	0.29

? = unable to determine / unclear risk; ☺ = low risk; ☹ = high risk;

Table 29 Diagnostic accuracy of catheter-free monitoring compared to clinical diagnosis for diagnosing GORD in adults

Study	Evidence level and quality	Population	Index test & cut-off point	Reference standard & cut-off point	Sensitivity % [95%CI]	Specificity % [95%CI]	PPV % [95%CI]	NPV % [95%CI]	LR+	LR–	Accuracy [95%CI]
Wenner, Johansson et al. (2007), Wenner et al. (2008) Diagnostic case-control study	Level: III-3 Quality: Patient selection: ☹ Index test: ? Reference standard: ☺ Flow and timing: ? CX P3	119 subjects (64 patients and 55 controls)	Bravo pH monitoring (48-hour) Time under pH 4 >1.9, 2.2, 2.3, 3.2, 3.6 and 4.4	Clinical diagnosis	1.9: 88% 2.2: 86% 2.3: 83% 3.2: 68% 3.6: 64% 4.4: 59%	1.9: 75% 2.2: 80% 2.3: 80% 3.2: 85% 3.6: 91% 4.4: 96%	-	-	1.9: 3.52 2.2: 4.30 2.3: 4.15 3.2: 4.53 3.6: 7.11 4.4: 14.75	1.9: 0.16 2.2: 0.18 2.3: 0.21 3.2: 0.38 3.6: 0.40 4.4: 0.43	-
Bansal et al. (2009) Diagnostic case-control study	Level: III-3 Quality: Patient selection: ☹ Index test: ? Reference standard: ☹ Flow and timing:? CX P2	56 subjects (40 patients and 16 controls)	Bravo pH monitoring (48-hour) Time under pH 4 >4.4% for the proximal location (6 cm above GOJ), and time under pH 4 >10.6% for the distal location (1 cm above GOJ)	Clinical diagnosis (no symptoms of reflux in control population)	Proximal: 67% Distal: 60%	Proximal: 66% Distal: 88%	-	-	-	-	-

Study	Evidence level and quality	Population	Index test & cut-off point	Reference standard & cut-off point	Sensitivity % [95%CI]	Specificity % [95%CI]	PPV % [95%CI]	NPV % [95%CI]	LR+	LR–	Accuracy [95%CI]
Ayazi et al. (2009) (2) Diagnostic case-control study	Level: III-3 Quality: Patient selection: ? Index test: ? Reference standard: ☺ Flow and timing: ? CX P3	44 subjects	Bravo pH monitoring (48-hour) Time under pH 4 >4.2%	Clinical diagnosis	100% [85–100]	95% [76–100]	-	-	-	-	98 [88–100]
Ayazi et al. (2009) (3) Diagnostic case-control study	Level: III-3 Quality: Patient selection: ☺ Index test: ? Reference standard: ☺ Flow and timing: ☺ CX P3	38 subjects (28 patients and 10 controls)	Bravo pH monitoring (48-hour) Time under pH 4 >4.2%	Clinical diagnosis (minimal clinical evidence of GORD in control population)	Abnormal combined 48 hours: 82% Abnormal first or second 24 hours: 93%	Abnormal combined 48 hours: 100% Abnormal first or second 24 hours: 100%	Abnormal combined 48 hours: 100% Abnormal first or second 24 hours: 67%	Abnormal combined 48 hours: 67% Abnormal first or second 24 hours: 83%	-	-	Abnormal combined 48 hours: 87% Abnormal first or second 24 hours: 95%
Pandolfino et al. (2003) Diagnostic case-control study	Level: III-3 Quality: Patient selection: ☹ Index test: ? Reference standard: ? Flow and timing: ☹ CX P3	37 subjects	Bravo pH monitoring (48-hour, day 1 (24-hour), worst day (W)) Time under pH 4 >5.3%	Clinical diagnosis	24-h: 67.5% 48-h: 64.9% W: 83.8%	24-h: 89.7% 48-h: 94.8% W: 84.5%	24-h: 86.2% 48-h: 92.3% W: 83.3%	24-h: 74.5% 48-h: 74.0% W: 84.6%	24-h: 6.55 48-h: 12.48 W: 5.4	24-h: 0.36 48-h: 0.37 W: 0.19	-

? = unable to determine / unclear risk; ☺ = low risk; ☹ = high risk

Concordance in acid exposure and reflux events in catheter-free monitoring versus catheter-based monitoring

There were nine studies that reported on the oesophageal acid exposure times during catheter-free monitoring compared with catheter-based monitoring in adults (Table 30). Four of these studies were in-subject simultaneous recording studies or randomised crossover trials, which means that both tests were conducted in the same patient population, allowing all confounding factors to be automatically controlled for (Azzam et al. 2012; des Varannes et al. 2005; Hakanson et al. 2009; Wenner, Johnsson et al. 2007). Results between the two styles of monitoring test are therefore expected to be more similar than when two different patient samples are compared. However, three of these studies reported a significant difference in acid exposure times between catheter-free and catheter-based monitoring. Two studies reported more acid exposure time with the catheter-based system (des Varannes et al. 2005; Hakanson et al. 2009), and one study reported more acid exposure with the wireless system (Azzam et al. 2012).

The remaining five studies also reported concordance data regarding acid exposure times, but in different patient groups (Ang et al. 2010; Grigolon et al. 2011; Martinez de Haro et al. 2008; Schneider et al. 2007; Wong et al. 2005). None of the studies reported significant differences in acid exposure times between the two measurement methods.

There was one study that reported concordance data on oesophageal acid exposure in children (Croffie et al. 2007). It reported significantly more reflux detected with wireless monitoring compared with catheter-based monitoring (p=0.0107). The tests were not performed in the same patient population; however, it was an RCT and therefore any baseline differences between the two groups are probably due to chance rather than a form of bias (Table 31).

Two level II studies and two level III-2 studies reported not only on acid exposure time concordance, but also on variation in the number of reflux events (Table 32). Three of the four studies—two level II and one level III-2—reported significantly more reflux events with the catheter-based monitoring system (Ang et al. 2010; des Varannes et al. 2005; Wenner et al. 2007b). Two of these studies used a within-subjects design, i.e. both tests were conducted in the same subjects, and in one they were compared simultaneously (des Varannes et al. 2005). Therefore, it can be concluded that the difference found in the number of reflux events recorded may be due to differences between the tests themselves and not other confounding factors such as the patient population variability or day-to-day variability in reflux.

Table 30   Concordance of acid exposure times using both monitoring tests in adults

Study	Study design	N	% of total time pH <4 wireless pH monitoring	% of total time pH <4 conventional pH catheter	p-value
Wenner, Johnsson et al. (2007)	Randomised crossover trial Level: II Quality: 20/26	31 patients	Median: 7.3	Median: 6.3	0.837 [a]
des Varannes et al. (2005)	In-subject simultaneous recording study Level: II Quality: 20/26	33 patients	Median: 2.4 (10th–90th percentile: 0.4–8.7)	Median: 3.6 (10th–90th percentile: 0.7–8.6)	0.0001 [a]
Azzam et al. (2012)	In-subject simultaneous recording study Level: II Quality: 17/26 C1 (for accuracy) P2	25 patients	Median: 6.1 (range: 0.1–21.4)	Median: 4.1 (range: 0.1–13.8)	0.001 [a]
Hakanson et al. (2009)	Case-control study with in-subject simultaneous recording Level: II Quality: 14/26	1st series: 26 patients, 27 controls 2nd series: 21 patients, 18 controls	Patients 1st: median 3.2 (day 1) Patients 2nd: median 2.4 (day 1) Controls 1st: median 0.6 (day 1) Controls 2nd: median 1.3 (day 1)	Patients 1st: median 6.8 Patients 2nd: median 7.1 Controls 1st: median 1.9 Controls 2nd: median 4.4	<0.001 [a] <0.001 [a] <0.05 [a] <0.05 [a]
Wong et al. (2007)	Randomised controlled trial Level: II Quality: 16.5/26	Catheter-free monitoring: 25 patients Catheter-based monitoring: 25 patients	Median: 1.9 (on PPI) Median: 10.2 (off PPI)	Median: 4.8 (on PPI) Median: 7.8 (off PPI)	0.7 0.78 [b]

Study	Study design	N	% of total time pH <4 wireless pH monitoring	% of total time pH <4 conventional pH catheter	p-value
Grigolon et al. (2011)	Matched-pairs retrospective cohort study Level: III-2 Quality: 16.5/26	Wireless monitoring: 51 patients Catheter-based monitoring: 51 patients	Median 24h: 3.5 (1st and 3rd quartile: 1.2–6.6) Median 48h: 4.0 (1st and 3rd quartile: 1.8–7.4) Median 96h: 4.7 (1st and 3rd quartile: 1.9–9.4)	Median: 3.5 (1st and 3rd quartile: 1.6–8.8)	-
Ang et al. (2010)	Comparative study of diagnostic yield Level: III-2 Quality: 14/26	Wireless monitoring: 66 patients Catheter-based monitoring: 55 patients	Day 1: 3.80 ± 6.09 Day 2: 3.91 ± 5.32 Overall: 3.86 ± 5.36	4.40 ± 5.34	Day 1: 0.57 [b] Day 2: 0.61 [b] Overall: 0.57 [b]
Martinez de Haro et al. (2008)	Non-randomised controlled trial Level: III-2 Quality: 11/26	Control group: 10 (simultaneous recording) No oesophagitis: 10 (5 catheter-based monitoring, 5 catheter-free monitoring) Non-Barrett's oesophagitis (5 catheter-based monitoring, 5 catheter-free monitoring) Barrett's oesophagitis (5 catheter-based monitoring, 5 catheter-free monitoring)	Controls: mean 1.7 No oesophagitis: mean 6.1 Non-Barrett's oesophagitis: mean 8.2 Barrett's oesophagitis: mean 27.3	Controls: mean 1.1 No oesophagitis: mean 5.8 Non-Barrett's oesophagitis: mean 7.9 Barrett's oesophagitis: mean 22.4	-
Schneider et al. (2007)	Non-randomised controlled trial Level: III-2 Quality: 9/26	Control group: 10 Oesophagitis: 68 catheter-free monitoring, 55 catheter-based monitoring After Nissen fundoplication: 43	Controls: median 0.8 Patients: median 14 Post-surgery: median 0.9	Controls: median 1 Patients: median 12 Post-surgery: median 1.2	0.1 [b] 0.8 [b] 0.9 [b]

[a] Analysis was done using a Wilcoxon signed rank test.
[b] Analysis was done using the Mann-Whitney U-test.

Table 31   Concordance of acid exposure times using both monitoring tests in children

Study	Study design	N	Reflux index wireless pH monitoring	Reflux index conventional pH catheter	p-value
Croffie et al. (2007)	Randomised controlled trial Level: II Quality:14.5/26	27 patients	Mean: 2.5 (range: 0.2–16.6)	Mean: 1.6 (range: 0.1–10.8)	0.0107[a]

[a] Analysis was done using the Mann-Whitney U-test.

Table 32   Number of reflux events in catheter-free monitoring compared with catheter-based monitoring

Study	Study design	N	Wireless pH monitoring	Conventional pH catheter	Test and p-value
Wenner, Johnsson et al. (2007)	Randomised crossover trial Level: II Quality: 20/26	31 patients	Median: 53	Median: 129	<0.0001 [a]
des Varannes et al. (2005)	In-subject simultaneous recording study Level: II Quality: 20/26	36 patients	Median: 23 (10th–90th percentile: 4–41)	Median: 40 (10th–90th percentile: 7–84)	0.0001 [a]
Ang et al. (2010)	Comparative study of diagnostic yield Level: III-2 Quality: 14/26	Catheter-free monitoring: 66 patients Catheter-based monitoring: 51 patients	(Mean ± SD) Day 1: 27 ± 27 Day 2: 31 ± 34	(Mean ± SD) 99 ± 111	<0.005 [b] <0.005 [b]
Schneider et al. (2007)	Non-randomised controlled trial Level: III-2 Quality: 9/26	Control group: 10 Oesophagitis: 68 catheter-free monitoring, 55 catheter-based monitoring After Nissen fundoplication: 43	Median in controls: 23 Median in patients: 247 Median post-surgery: 24	Median in controls: 18 Median in patients: 187 Median post-surgery: 18	0.6 [b] 0.09 [b] 0.07 [b]

[a] Analysis was done using a Wilcoxon signed rank test
[b] Analysis was done using the Mann-Whitney U-test

Diagnostic yield

Six studies report diagnostic yield in both wireless and catheter-based monitoring. Two of the studies were in-subject simultaneous recording studies, where both methods of pH monitoring were conducted at the same time in the same subjects. Both these studies reported no significant difference in diagnostic yield between the two tests. Only two comparative diagnostic yield studies of medium quality (Kushnir, Sayuk & Gyawali 2011; Sweis et al. 2009) reported significant differences—more patients were diagnosed with catheter-free monitoring compared with catheter-based monitoring (p-values <0.001) (Table 33). In the larger of the two comparative diagnostic yield studies, 357/462 (77.2%) patients were diagnosed with the wireless pH test, compared with 849/1605 (52.9%) with the catheter-based test (Kushnir, Sayuk & Gyawali 2011).

Table 33   Diagnostic yield using catheter-free monitoring compared to catheter-based monitoring

Study	Evidence level and quality	N	Wireless pH monitoring (%)	Catheter-based pH monitoring (%)	Cut-off value for positive test [a]	p-value
Wenner, Johnsson et al. (2007)	Randomised crossover trial Level: II Quality: 20/26	31 patients suspected of GORD	21/31 (68) (5 were discordant)	22/31 (71) (5 were discordant)	CFPM: 4.4% CBPM: 3.4%	-
Azzam et al. (2012)	In-subject simultaneous recording study Level: II Quality: 17/26	25 patients	19/25 (76)	16/25 (64)	Total: 4.5% Upright: 8.4% Supine: 3.5%	0.355 [b]
Hakanson et al. (2009)	Case-control study with in-subject simultaneous recording Level: II Quality: 14/26	92 subjects (45 volunteers and 47 patients)	42/92 (45.7)	53/92 (57.6)	CFPM: 2.0% CBPM: 4.0%	-
Sweis et al. (2009)	Comparative study of diagnostic yield Level: III-2 Quality: 15/26	129 patients in the wireless group (intolerant of the catheter) 102 patients in the catheter group	Cut-off 5.3%: 92/129 (71) Cut-off 4.2%: 98/129 (76)	Cut-off 4.2%: 49/102 (48)	CFPM: 5.3% CBPM: 4.2% Both: 4.2%	<0.001 [c] <0.001 [c]
Ang et al. (2010)	Comparative study of diagnostic yield Level: III-2 Quality: 14/26	Wireless group: 66 patients Catheter group: 55 patients	Overall: 26/66 (39.4) On day 1 only: 5/66 (7.6) On day 2 only: 8/66 (12.1) Both day 1 and 2: 13/66 (19.7)	Overall: 20/55 (36.4)	CFPM day 1: 5.8% CFPM day 2: 4.5% CFPM both days: 4.2% CBPM: 4.2%	-
Kushnir, Sayuk & Gyawali (2011)	Comparative study of diagnostic yield Level: III-2 Quality: 13.5/26	462 patients in the wireless group 1,605 patients in the catheter group	357/462 (77.2)	849/1,605 (52.9)	CFPM: 4.0% CBPM: 4.0%	<0.001 [c]

CFPM: catheter-free pH monitoring
CBPM: catheter-based pH monitoring
[a] Cut-off value is presented in the percentage of time the pH <4; the test result is positive when the percentage is higher than the cut-off value.
[b] Unclear which statistical test was used.
[c] Analysis was done using the Chi square test.

As the monitoring period can be lengthened with the wireless pH method (most tests are 48 hours instead of the 24 hours with the traditional catheter-based method), this is likely to increase diagnostic yield. Six studies reported an additional yield on day 2 of monitoring, varying from 2.0% to 15.6% with a median of 7.8% (Table 34). This means that 2.0%–15.6% of patients only get a positive test result above the cut-off value after day 2, and these patients would not have been diagnosed with GORD if they had only had a 24-hour monitoring period.

Table 34   Yield comparison between 24 hours and 48 hours with catheter-free monitoring

Study	Study design	N	Cut-off value [a]	Total yield day 1	Total yield day 2	Overall detection increase (%)
Prakash & Clouse (2005)	Retrospective case series	121 patients	5.33%	NA	NA	15 (7.5%)
Grigolon et al. (2011)	Matched pairs cohort study	51 patients	4.7%	24 patients (47.1%)	25 patients (49.0%)	2.0%
Garrean et al. (2008)	Prospective case series	40 patients	5.3%	10 patients (25.0%)	14 patients (35.0%)	10.0%
Pandolfino et al. (2003)	Diagnostic case-control study	37 patients	5.33%	12 patients (32.4%)	15 patients (40.5%)	8.1%
Scarpulla et al. (2007)	Retrospective case series	34 patients	4.2%	12 patients (35.3%)	14 patients (41.2%)	5.9%
Karamanolis et al. (2012)	Post-test case series	32 patients	Day 1: 5.8% Day 2: 4.5%	12 patients (37.5%)	17 patients (53.1%)	15.6%

[a] Cut-off value is presented in the percentage of time the pH <4; the test result is positive when the percentage is higher than the cut-off value.

Fourteen studies reported diagnostic yield with the wireless test (and no comparator). Since there is no consensus on what the threshold is for diagnosing GORD, studies used different cut-off points, different scores (composite pH score, acid exposure time, reflux events), different placements in the oesophagus (proximal, distal) and different analysis (an abnormal score on only one day, an abnormal score on both days, worst day analysis) to establish a diagnosis. Therefore, there is a large variation in diagnostic yield, varying from 6.1%—only abnormal on day 2 in patients suspected of GORD (Turner et al. 2007), to 92.9%—abnormal score overall in patients with strong clinical GORD evidence (Ayazi et al. 2009) (see Tables 81 and 82 in Appendix H).

Does it change patient management?

Summary—Does catheter-free (wireless) ambulatory oesophageal pH monitoring change clinical management, compared with no pH monitoring, for patients with symptoms of GORD who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate?

No studies were available that reported on a change in management based on catheter-free monitoring in the population who would otherwise have had no monitoring, i.e. those who could not tolerate catheter-based monitoring. Five studies reported on change in patient management after catheter-free monitoring in patients who could have potentially tolerated a catheter-based test. One comparative study reported that concordance between the results of the test and management of GORD was higher in the catheter-free monitoring group (78%) than in the catheter-based monitoring group (58%; p<0.05).

As with the comparative study, before-and-after case series also reported that the management received did not always correspond to the results of the test. One study reported that 42.2% of patients (38/90) continued to take anti-reflux medication despite a negative pH test result. Only 17 patients recalled being instructed to stop taking PPIs. In another case series, 12/38 patients underwent surgery for GORD even though only 9 of these had a positive diagnosis based on catheter-free monitoring. In a further case series, catheter-free monitoring led to a change in management in 63% of patients referred for monitoring; and in a paediatric study, catheter-free monitoring resulted in a change in management in 88% of patients. Patients with an abnormal study result were more likely to have a change in management than patients with a normal result.

From the literature it can be assumed that the results of catheter-free monitoring do influence subsequent management, although not all patients have management consistent with the results of pH monitoring. It is expected that the key changes based on results of catheter-free monitoring are that patients who are found *not* to have GORD have their (mostly ineffective) PPI treatment suspended, whereas patients who have their GORD *confirmed* have an additional option of surgical treatment.

Studies were included to assess change in management following catheter-free (wireless) oesophageal pH monitoring according to criteria outlined *a priori* in Box 5. No studies were identified that met the population inclusion criteria. Five studies that did not meet the *a priori* inclusion criteria reported on change in patient management after catheter-free monitoring in patients who were suspected of GORD but who could have potentially tolerated a catheter-based test. Furthermore, none of the studies had 'no pH monitoring' as a comparator. However, due to a lack of more relevant data, these studies are discussed below.

Box 5   Inclusion criteria for identification of studies relevant to an assessment of change in management following catheter-free (wireless) ambulatory oesophageal pH monitoring

Selection criteria	Inclusion criteria
Population	Patients with symptoms of GORD in whom a catheter-based test has failed or in whom catheter-based monitoring is anatomically inappropriate
Intervention	Catheter-free (wireless) ambulatory oesophageal pH monitoring
Comparators	No pH monitoring
Outcomes	% change in management plan including surgeries performed and changes in medication plans
	Time to diagnosis
Study design	Randomised or non-randomised controlled trials and cohort studies, case-control studies, case series, or systematic reviews of these study designs
Search period	Catheter-free monitoring was introduced in 2001, so the search period will be 2001 – 5/2013
Language	Studies in languages other than English were excluded unless they represented a higher level of evidence than that available in the English language evidence-base

One of the studies, a retrospective case series conducted in Chicago, Illinois (Gawron et al. 2012), reported on the number of patients who continued to take anti-reflux medication despite a negative pH test result. A total of 38 out of 90 patients (42.2%) reported current PPI use despite negative results from a pH study—66 catheter-free and 24 catheter-based monitoring—and 13 of these patients were on twice-daily PPIs. Only 17 patients (18.9%) recalled being instructed to stop taking PPIs after the test. Review data showed documented instructions to stop PPI therapy for only 15 patients (16.7%).

In a case series by Sweis et al. (2011), medical therapy or conservative advice was pursued by the physician in 68% (26/38) of the patients, with the outcome assessed at a median of 24 months (range 12–36 months) after a 96-hour wireless test. Anti-reflux surgery was performed in 12/38 (32%) patients, of which 9 had a positive diagnosis based on pathologic acid exposure, 10 had a positive reflux symptom association based on the wireless test, and 2 were operated on despite a negative pH test result. One of these patients had pathological supine reflux and a large hiatal hernia on endoscopy, and the other had ongoing symptoms of non-acid food and fluid regurgitation with normal endoscopy and manometry.

Another study, a matched-pairs retrospective cohort (Grigolon et al. 2011), compared the change in management after catheter-free monitoring with management changes following catheter-based monitoring. Anti-reflux medication was used in 43/50 (86%) and 38/50 (76%) patients before the wireless and traditional pH monitoring, respectively (baseline rates not significantly different). After the test, concordance between results of the test and treatment for GORD was higher in the wireless than in the traditional group. That is, those with GORD should be treated with medication or surgery, and those without GORD should

stop medication—78% versus 58% of the patients (p <0.05). There were 11 discordant cases in the wireless group and 21 in the catheter-based group: that is, 10 and 14 patients, respectively, who were taking PPIs before the test continued treatment despite a negative or indeterminate test; and 1 and 7, respectively, who were not taking PPIs before the test continued to do so despite a positive test result. Treatment for GORD at the time of a follow-up telephone interview 1–2 years after the test was discordant with the treatment recommended after the test in 12 (24%) and 15 (31%) patients in the wireless and traditional groups, respectively (not statistically significant). The majority of discordant cases, 11 in the former and 10 in the latter group, were patients who continued taking PPIs immediately after the test but who had stopped them by the time of the interview. The remaining discordant cases did not undergo GORD treatment after the test but were taking it at the time of the interview (Table 35).

Table 35   Concordance between results of the test and treatment for GORD

Discordant cases and treatment changes	Catheter-free monitoring group (n=50)	Catheter-based monitoring group (n=50)
Discordant cases (between results of test and treatment)	11/50 (22%)	21/50 (42%)
PPIs despite negative or indeterminate test	10/50 (20%)	14/50 (28%)
No treatment despite positive test	1/50 (2%)	7/50 (14%)
Discordant cases (between treatment and recommended treatment after test) 1–2 years after test	12/50 (24%)	15/50 (31%)
Stopped PPIs during follow-up	11/50 (22%)	10/50 (20%)
Started treatment during follow-up	1/50 (2%)	5/50 (10%)

Source: Grigolon et al. (2011)

Two case series by Lacy et al. (2009, 2011) in paediatric and adult populations, respectively, reported specific management changes after wireless pH monitoring (Table 36).

Table 36   Specific management changes after wireless oesophageal pH monitoring

Specific management changes	Lacy et al. (2011) n=167 (adult) n (%)	Lacy et al. (2009) n=44 (paediatric) n (%)
Referred for surgery	69 (41)	1 (2.3)
Increased dose of medication (PPI or H2RA)	32 (19)	1 (2.3)
Initiated medication (PPI)	-	24 (54.5)
Medication changed (PPI or from H2RA to PPI)	19 (11)	7 (15.9)
Medication stopped (PPI or H2RA)	28 (17)	9 (20.5)
Added medication (hyoscyamine, sucralfate, metoclopramide)	21 (13)	2 (4.6)

Specific management changes	Lacy et al. (2011) n=167 (adult) n (%)	Lacy et al. (2009) n=44 (paediatric) n (%)
Other reason (decreased dose of medication, medication change not otherwise specified)	19 (11)	-

In the larger of two case series by Lacy et al. (2011), conducted in adults, the results of the wireless pH-monitoring capsule directly led to a change in diagnosis in 22% of all patients referred for testing. The diagnosis changed most frequently in those patients undergoing evaluation for chronic cough (50%), followed by ENT symptoms (36%), acid reflux symptoms while off PPIs (29%), and chest pain (26%).The diagnosis was more likely to be changed in patients studied while *off* PPI therapy compared with being monitored while *on* PPI therapy. Wireless pH-testing results led to a change in management in 63% of all patients referred for monitoring. Clinical management changes were more likely to occur in patients studied off PPI (77%) compared with once daily (43%) or twice daily (45%) PPI therapy (p<0.001). In the earlier study by Lacy et al. (2009), conducted in children, management changed in 44/50 (88%) patients, and no follow up was available for two patients (4%). Patients with an abnormal study result, i.e. acid exposure time >6%, (n=30) were more likely to get a change in management than those with a normal result (n=14; p<0.0001).

Does change in management improve patient outcomes?

Summary—Do alterations in clinical management and treatment options have an impact on the health outcomes of patients with symptoms of GORD who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate?

For people initially suspected of having GORD who are given alternative diagnoses correctly after catheter-free monitoring (true negatives), it is assumed that their clinical management would be optimised as a result of obtaining the correct diagnosis. Patients who receive a false positive would likely continue to receive ineffective PPI treatment and suffer a delay in receiving their correct diagnosis. A false negative would likely result in suspension of PPI treatment and a delay in diagnosis while an alternative diagnosis is sought. No relevant studies were identified regarding the impact of false negatives; however, two studies on false positives were included. These studies reported that infants with infantile spasms with delayed treatment as a result of a false GORD diagnosis had a poorer outcome and worse treatment response than infants without a delay in diagnosis and treatment.

Furthermore, two systematic reviews on true positives were included to answer the question of whether anti-reflux surgery for GORD leads to better health outcomes compared with medical treatment in patients with GORD with a positive pH-test result. The first, medium-quality review reported that improved outcomes were more common after surgical than medical treatment, with significant differences in objective outcomes (pH-reflux duration, oesophagitis, LOS pressure etc.) in 5/6 included RCTs and 2/3 cohort studies. Subjective outcomes were also more common among surgical patients in 7/8 studies. The second, more recent, high-quality review compared medical management with laparoscopic fundoplication surgery. Four RCTs were included. Significant improvements in quality of life (QoL) were reported at 3 months and 1 year after surgery, compared with medical treatment. All studies reported significant increases in GORD-specific QoL postoperatively, compared with medical treatment. Post-operative complications were rare.

Two of the four RCTs included in the latter review published 5-year follow-up studies after the Cochrane review was conducted. Most patients achieve and remain in remission at 5 years with anti-reflux surgery, and fundoplication continued to give better pain relief than medical management. It should be noted, however, that these trials consisted solely of PPI responders, and their results do not generalise to patients who are refractory to PPI therapy.

The expected treatment changes that may result from having a catheter-free monitoring test are likely to benefit QoL.

To answer the question of whether a change in management leads to improved patient health outcomes, a second literature search was done based on the PICO criteria outlined *a*

priori in Box 6. For people initially suspected of having GORD but eventually given an alternative diagnosis (true negatives) and a subsequent change in management, it was assumed that their management has been optimised as a result of obtaining the correct diagnosis.

Box 6   Inclusion criteria for identification of studies relevant to an assessment of health outcomes following a change in management

Selection criteria	Inclusion criteria true positives	Inclusion criteria false positives	Inclusion criteria false negatives
Population	People with GORD	People who are wrongly diagnosed with GORD	People with undiagnosed GORD
Intervention	Optimal treatment for GORD (PPIs, H2 receptor antagonists, surgery)	(Unnecessary) treatment for GORD (PPIs, H2 receptor antagonists, surgery)	Possible cessation of PPI therapy and consideration and/or treatment trial for other diagnoses
Comparators	Alternative treatment	No GORD treatment and/or optimal treatment for true condition	Optimal treatment for GORD (PPIs, H2 receptor antagonists, surgery)
Outcomes	Adverse events from treatment, quality of life, reduction in progression to Barrett's oesophagus and/or oesophageal cancer, symptom resolution or symptom reduction	Adverse events from inappropriate use of GORD treatment, quality of life, adverse events from delay in appropriate treatment for differential diagnosis	Quality of life, progression to Barrett's oesophagus and/or oesophageal cancer, symptom resolution
Study design	Systematic reviews, meta-analysis, evidence-based clinical practice guidelines	Level 1 evidence if available, otherwise randomised or non-randomised controlled trials and cohort studies, case-control studies, case series, or systematic reviews of these study designs	Level 1 evidence if available, otherwise randomised or non-randomised controlled trials and cohort studies, case-control studies, case series, or systematic reviews of these study designs
Search period	Most recent studies found	No limits	No limits
Language	Studies in languages other than English were excluded unless they represented a higher level of evidence than that available in the English language evidence-base		

True positives

Patients suspected of GORD are usually given a trial of PPI therapy; where this is not completely successful in managing their symptoms, pH monitoring may be undertaken. If the patient has a (true) positive pH-test result, there may be a change in management in that the patient would have the option to undergo anti-reflux surgery. This is not generally recommended unless acid-reflux has been positively identified as the cause of symptoms. A combination of surgical and medical treatment was therefore included as the intervention (the optimal treatment), and medical treatment was included as the comparator (the alternative treatment). This raises the question of whether anti-reflux surgery for GORD

leads to better health outcomes, compared with medical treatment, in patients with GORD with a true positive pH-test result?

Out of 5,553 studies (3,708 after duplicates were removed), 114 full-text articles were studied. Three relevant systematic reviews based on these criteria were identified; however, the most recent review (Al Talalwah & Woodward 2013) was excluded as it provided no additional new data and was of lesser quality than the other identified systematic reviews. As a result, two systematic reviews were included (Allgood & Bachmann 2000; Wileman et al. 2010).

Allgood and Bachmann (2000) aimed to evaluate and summarise all comparative studies of the effectiveness of surgical and medical management of GORD. Six RCTs and three cohort studies were identified. The review was of medium quality and did not distinguish between the different medical and surgical treatments. Both laparoscopic and open anti-reflux surgical procedures were included, as well as both Nissen-Rossetti (360˚) and Toupet (180˚) fundoplication. Improved outcomes were more common after surgical than medical treatment, with significant differences in objective outcomes (pH-reflux duration, oesophagitis, LOS pressure etc.) in 5/6 included RCTs and 2/3 cohort studies. Five studies (4 RCTs and 1 cohort study) mentioned changes in oesophageal pH; and three studies (2 RCTs and 1 cohort study) reported a significantly better response to surgery, whereas one RCT only showed a significantly higher response to surgery than to medication for symptoms, and not for continuous medication. Subjective outcomes (symptoms and patients' satisfaction) were also more common among surgical patients in 7/8 studies. The overall conclusion by the authors was that surgery is consistently more effective across a range of patients and for a range of both short-term and long-term outcomes.

The other, more recent included systematic review is a high-quality Cochrane review (Wileman et al. 2010). The aim was to compare medical management (either PPIs or H2 receptor antagonists) with laparoscopic fundoplication surgery for the treatment of GORD in adults. Twelve papers were included that reported data on multiple time points of four clinical trials with a total of 1,232 randomised participants, published between 2005 and 2009. All four studies reported health-related QoL, and two studies were combined using fixed effect models. This showed that there were significant improvements in health-related QoL at 3 months and 1 year after surgery, compared with medical therapy (mean difference SF36 general health score: –5.23, 95%CI –6.83 to –3.62, I^2 = 0%). One other study reported higher vitality scores compared with medical therapy at 1 and 3 years postoperatively (p<0.001), and the fourth study reported a higher QoL at 3 months and 1 year after surgery,

compared with medical therapy (p<0.001). All included studies reported significant increases in GORD-specific QoL postoperatively, compared with medical treatment. The authors reported that symptoms of heartburn, reflux and bloating were improved after surgery, compared with medical therapy, but some patients had persistent postoperative dysphagia. Postoperative complications were rare, but surgery is not without risk and postoperative adverse events did occur. The authors suggest that there is evidence that surgery is more effective than medical management for the treatment of GORD, at least in the short to medium term.

Two of the four included RCTs in the Cochrane review (the 'LOTUS trial' and the 'REFLUX trial') published 5-year follow-up studies after the review was conducted, to give more information on the long-term results of surgery and medication for GORD (Galmiche et al. 2011; Grant, Cotton et al. 2013). The LOTUS randomised trial was conducted in academic hospitals in 11 different European countries between October 2001 and April 2009; it included 554 patients, of which 372 completed the 5-year follow-up (Galmiche et al. 2011). Initial recruitment for the REFLUX trial was done in 21 UK hospitals and 810 patients were included—357 patients were recruited to a randomised comparison and 453 to non-randomised preference groups (Grant, Cotton et al. 2013). The aim was to evaluate maintenance therapy with medication vs laparoscopic anti-reflux surgery. In the LOTUS trial, only esomeprazole was used in the medical management group.

The LOTUS trial reported that, at 5 years, an estimated 85% (95%CI, 81%–90%) of patients in the surgery group and 92% (95%CI, 89%–96%) in the esomeprazole group remained in remission (log-rank p=0.048). In the REFLUX trial QoL was measured on a disease-specific REFLUX questionnaire; among responders at 5 years the randomised surgery group scored significantly higher compared with the medication group, with a mean difference of 8.5 (95%CI 3.9, 13.1, p<0.001) for intention-to-treat analyses and 11.5 (95%CI 4.2, 18.7, p=0.002) adjusted for treatment received. Short Form-36 scores and mean EuroQol questionnaire 5 dimensions (EQ-5D) scores favoured the surgical group in all domains at all time points, although differences decreased over time.

The prevalence of GORD-related symptoms for both treatment groups in both trials can be found in Table 37. In the REFLUX trial heartburn, regurgitation and belching were reported less frequently in the group randomised to surgery compared with the medication group, with no significant differences in 'difficulty swallowing', 'wind from the bowel' and 'wanting to be sick but being unable' (Table 37). Galmiche et al. (2011) reported that most patients achieve and remain in remission at 5 years with either medication treatment or laparoscopic

surgery. However, the results of both trials are inconsistent, as the majority of the medication group in the REFLUX trial had heartburn at least once a week at the 5-year follow-up. Grant, Cotton et al. (2013; REFLUX trial) concluded that, after 5 years, laparoscopic fundoplication continued to give better relief of GORD symptoms than medical management. It should be noted that these trials consisted solely of (partial or complete) PPI responders, and their results do not generalise to patients who are refractory to PPI therapy. Therefore, these results, especially in the medication groups, need to be interpreted with caution, as the patients who would undergo catheter-free monitoring may be refractory to PPI treatment.

Table 37   Prevalence of symptoms at 5-year follow-up of LOTUS and REFLUX trials

Prevalence of symptoms	Esomeprazole group LOTUS (n=192)	Surgery group LOTUS (n=180)	p-value LOTUS	Medication group REFLUX (n)	Surgery group REFLUX (n)
Heartburn	16%	8%	0.14	78 (73.6%)	46 (41.4%)
Acid regurgitation	13%	2%	<0.001	41 (36.6%)	29 (24.6%)
Dysphagia	5%	11%	<0.001	28 (25.5%)	27 (22.9%)
Bloating	28%	40%	<0.001	-	-
Flatulence/wind from bowel	40%	57%	<0.001	96 (87.3%)	104 (88.1%)
Belching	-	-	-	83 (75.5%)	71 (60.7%)
Nausea but unable to be sick	-	-	-	19 (17.1%)	17 (14.4%)
Adverse events	24.1%	28.6%	-	-	-

Sources: Galmiche et al. (2011); Grant, Cotton et al. (2013)

False positives

People in the false positive group would be people who have a positive result on the pH test despite not truly having GORD. This would mean that they would be treated for GORD, either medical treatment or surgery, unnecessarily after their pH-test result. Studies on unnecessary GORD treatment and GORD misdiagnosis were searched and eight were identified (Auvin et al. 2012; Kessing, Bredenoord & Smout 2011; Napuri et al. 2010; Starr 2006, 2007; Susman 2009; Sweetman, Ng & Kerrigan 2007; Taddio et al. 2011). However, six of these studies had to be excluded. Three case reports reported on heart disease patients who were wrongly diagnosed with GORD but one had to be excluded because the patient died before he was able to undergo pH monitoring (Susman 2009), and two patients reported that they responded sufficiently to anti-reflux medication before further investigations (e.g. pH monitoring) were undertaken (Starr 2006, 2007). One study reported three case reports where achalasia was initially misdiagnosed as GORD; however, no

relevant outcomes were reported (Kessing, Bredenoord & Smout 2011). Finally, two studies on GORD misdiagnosis in infants were excluded because they did not report health outcomes (Taddio et al. 2011; Wileman et al. 2010).

Two studies on GORD misdiagnosis of small infants were included, as these studies gave some evidence on health-related consequences of delayed diagnosis of conditions initially misdiagnosed as GORD (Auvin et al. 2012; Napuri et al. 2010). Although these infants did not undergo pH monitoring, it gives an indication of the impact of a false positive GORD diagnosis on health outcomes in infants. Auvin et al. (2012) reported that in 83 infants diagnosed with infantile spasms, 7% were initially diagnosed with GORD. This misdiagnosis was an important reason for diagnostic delay, and a diagnostic delay of more than 30 days was associated with a poor outcome (e.g. seizures, mental delay) (RR=12.08 [1.52–96.3]). In the multivariate analysis a poor outcome 2 years after diagnosis was also predicted by diagnostic delay of more than 30 days (RR=31.70 [2.30–437.68]). Napuri et al. (2010) included 156 infants also diagnosed with infantile spasms. Of these patients, 17% were initially diagnosed with GORD. The authors reported that response to treatment was significantly better when treatment lag was less than 6 weeks (40% responders) than when it was longer (20% responders; p=0.021). These results show that infants with infantile spasms benefit from early diagnosis, and a false GORD diagnosis could unnecessarily delay treatment, leading to worse health outcomes.

False negatives

No studies were identified that met the criteria for false negatives in Box 6.

Other relevant considerations

Defining the population and the potential for 'leakage'

It can be difficult to determine whether a person is 'unable' to tolerate a catheter. The catheter-free monitoring test is viewed as more convenient and can lead to a higher diagnostic yield due to prolonged monitoring, compared with catheter-based monitoring. Physicians could be inclined to give more patients catheter-free monitoring to avoid discomfort and embarrassment, especially in children. Furthermore, it is shown that catheter-free monitoring has less impact on normal daily activities, which makes the recording more reliable. Therefore, it is possible that the population using catheter-free monitoring may expand beyond those who are currently considered unable to tolerate catheter-based monitoring, when the alternative is no monitoring. That is, the use of the wireless test could start 'leaking' into the population currently receiving the catheter-based test.

Because of the higher costs of catheter-free monitoring compared with catheter-based monitoring, this 'leakage' could lead to extra costs. This scenario is examined in the economics section of the report.

Catheter-based monitoring is normally inserted transnasally with manometric guidance, whereas catheter-free monitoring is usually inserted through the mouth using endoscopy, although manometric guidance is possible with catheter-free monitoring. However, as the population eligible for catheter-free monitoring does not tolerate a catheter through the nose, they are unlikely to tolerate manometry, leaving transoral endoscopy as the only option to guide insertion. The costs for endoscopy are higher than for manometry, and this is one of the reasons that the costs for the new MBS item are higher than for MBS item 11810 (catheter-based monitoring). Therefore, if 'leakage' takes place and catheter-free monitoring would be performed with manometric guidance through the nose, the fee should be reassessed.

What are the economic considerations?

Economic evaluation

Overview

The DAP identified two key economic questions. These were:

1. What is the cost-effectiveness of catheter-free (wireless) ambulatory oesophageal pH monitoring, compared with no pH monitoring, for patients with symptoms of GORD who have previously failed a catheter-based test or where catheter-based monitoring is anatomically inappropriate?

2. What is the cost-effectiveness of catheter-free (wireless) ambulatory oesophageal pH monitoring, compared with catheter-based oesophageal pH monitoring, in patients with symptoms of GORD?

To address these questions, two separate economic models are required.

> **Model 1 Scenario**
>
> An evaluation where the population is restricted, as per the proposed listing, to *only those patients who cannot tolerate catheter-based monitoring*, i.e. who have tried and failed, or are anatomically unsuitable. In this case *'no alternative pH monitoring' is the appropriate comparator*. Initially, the base-case assumes 100% sensitivity and specificity of catheter-free (wireless) pH monitoring on the grounds that catheter-based pH monitoring is an imperfect reference standard and the true accuracy of either test is unknown. However, the sensitivity and specificity of wireless pH monitoring relative to catheter-based monitoring, based on the results of the systematic review, are also modelled for a sensitivity analysis. This modelled evaluation is intended to answer the first economic question.

> **Model 2 Scenario**
>
> An evaluation where the population accesses wireless pH monitoring as an alternative to catheter-based pH monitoring (the comparator). Within this scenario the assumptions of relative accuracy associated with wireless pH monitoring, referenced to catheter-based monitoring, are also varied and presented as sensitivity analyses. This modelled evaluation is intended to answer the second economic question.

In reality, while the proposed listing restricts the population to that in Model 1, given the reported patient preference and convenience associated with wireless monitoring, it may be anticipated that there will be some leakage of wireless pH monitoring beyond the requested listing. Therefore, future MBS-funded clinical practice may include patients who could actually tolerate catheter-based pH monitoring but have a strong preference for wireless monitoring, such as the population in Model 2. It might be expected that the overall population of potential wireless pH-monitoring patients would be made up a combination of patients from each of the model populations. Therefore, results are also presented for <u>various weighted combinations of Models 1 and 2</u> such that a pragmatic estimate of cost-effectiveness incorporating varying levels of leakage beyond the listing can be made.

In both models it is assumed that the monitoring will result in changed health outcomes, and therefore a cost–utility analysis is undertaken that measures incremental differences in quality-adjusted life years (QALYs) and GORD-related resource costs between wireless pH monitoring and the nominated comparator. The analysis is conducted primarily from the Australian health system perspective and, given the lack of long-term clinical data, is assessed over a 15-year time horizon to minimise uncertainty cost-effectiveness. It should be noted that a range of time horizons is explored in sensitivity analyses; however, if longer time horizons are required to demonstrate cost-effectiveness, there will be increasing uncertainty associated with the extended extrapolation and analysis.

Economic literature review

A search to identify existing literature containing economic models relating to pH monitoring and management of GORD symptoms was undertaken (see Appendix J). Six citations specifically relating to economic analysis of pH monitoring were identified. Of these, only two related to the relevant clinical pathway (i.e. following endoscopy). Lee et al. (2008) used wireless monitoring but was not restricted to patients who could not tolerate catheter-based monitoring; and the other study (Kleiman et al. 2013) used catheter-based monitoring. Both were cost analyses that did not consider clinical outcomes. None of the existing literature identified was adequate to answer the economic questions posed in the DAP; therefore, further development of economic models was required.

In addition, a broader literature search to identify other economic models in the area of gastro-oesophageal reflux was undertaken, with the intent that identified references may provide potential sources of data or validation for various parameters/inputs required in the assessment-developed models.

Populations and settings

The populations modelled are intended to represent Australian patients who would be candidates for wireless pH monitoring funded by Medicare.

The clinical pathway describes both the currently available catheter-based pH monitoring and the proposed wireless pH monitoring as being used as a diagnostic tool in patients who, despite a normal endoscopy, have ongoing GORD symptoms that are refractory to PPI use. There is not a single definition in the literature of what it means to be 'refractory to PPIs'; both the accepted dose range and the acceptability of anything less than a complete response with respect to symptom relief vary.

Based on Medicare data[5], catheter-based pH monitoring is currently undertaken across all ages of the Australian population, from babies to 85+ year olds. The average and median age group is 45–54 year olds, but the age group with the single largest use is 55–64 year olds. Of all pH-monitoring patients, 61% are female. To approximate this population, each model assumes that suspected GORD patients receiving pH monitoring are 55 years of age and that there is a male:female ratio of 40%:60%. This assumption is used to estimate background mortality.

The setting of the model begins in the outpatient setting with patients in the care of a specialist gastroenterologist. However, on being assigned the proposed intervention, the patient would have the wireless pH-monitoring capsule placed by a gastroenterologist, generally in either an endoscopy suite or day facility, and temporarily become a hospital inpatient for this procedure—93% of endoscopy patients who subsequently receive pH monitoring were inpatients for their initial endoscopy[5]. In contrast, empirical medical treatment is assumed to be undertaken in the outpatient setting (physician rooms); and only around 4% of MBS claims for catheter-based pH monitoring are for hospital-based patients, as the catheter is generally inserted without sedation.

Given the patient definition, it is assumed that patients entering the model will have GORD symptoms and, on average, a QoL commensurate with ongoing symptoms that will be unchanged unless successful treatment is achieved. Likewise, it will be assumed that patients entering the model are, on average, consuming resources equivalent to a standard single daily dose of PPI treatment, although, in reality, pre-existing use and dosing of GORD

---

[5] Medicare data on Item 11810 processed July 2011 – June 2013 (https://www.medicareaustralia.gov.au)

medication is likely to be heterogeneous. This use is assumed to continue in the first cycle where the intervention (i.e. pH testing/monitoring) is allocated in Model 2, but in the comparator arm of Model 1 (i.e. no pH testing/monitoring plus empirical treatment) it is assumed that there would be a trial of high-dose PPIs immediately.

Structure and rationale of the economic evaluation

The structures of the economic models are consistent with the clinical pathway presented in Figure 3. A health state diagram showing potential transitions between health states is presented   in
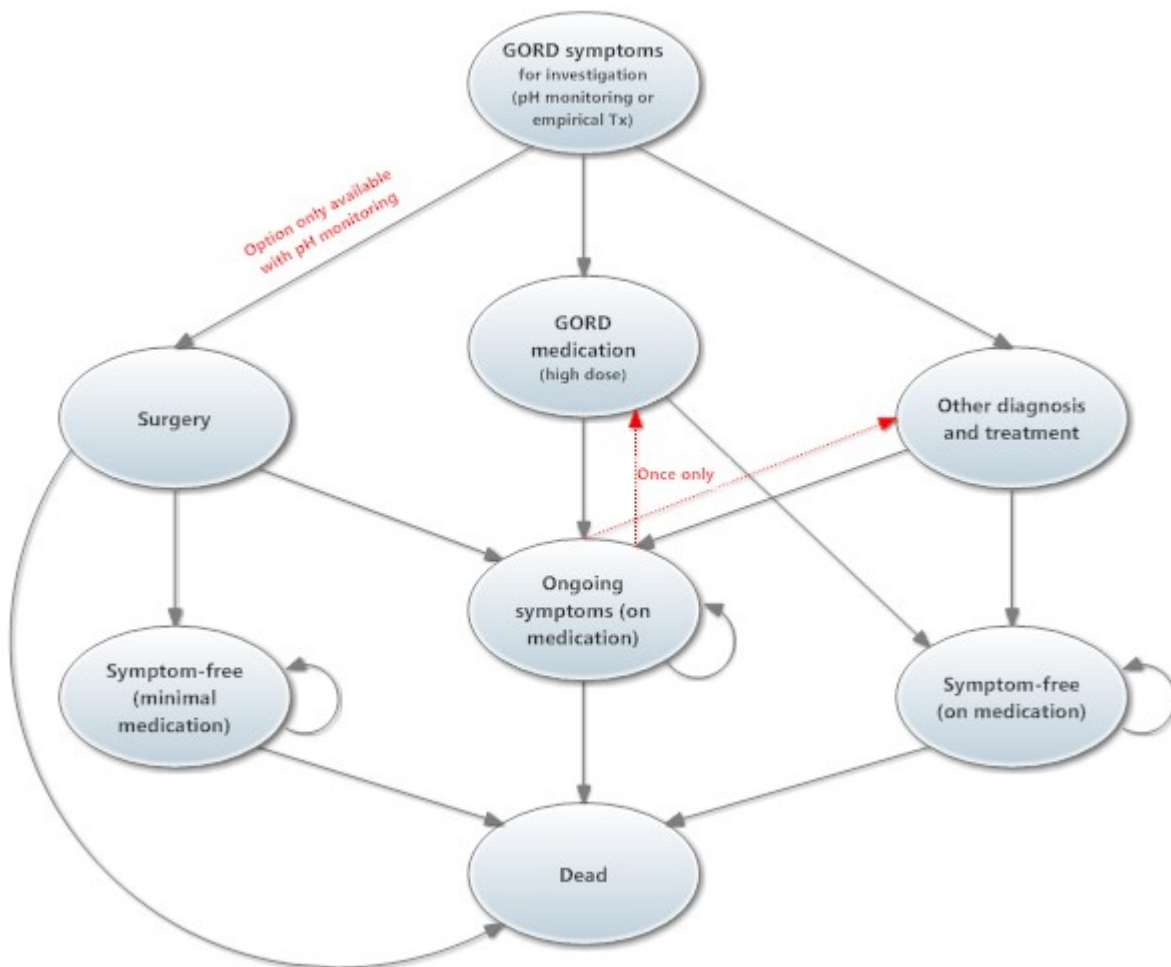
Figure 8. Immediately on entering the model, patients are assumed to be allocated to either receive pH monitoring (Model 1 intervention and Model 2 both arms) or empirical treatment with high-dose PPIs (Model 1 comparator).
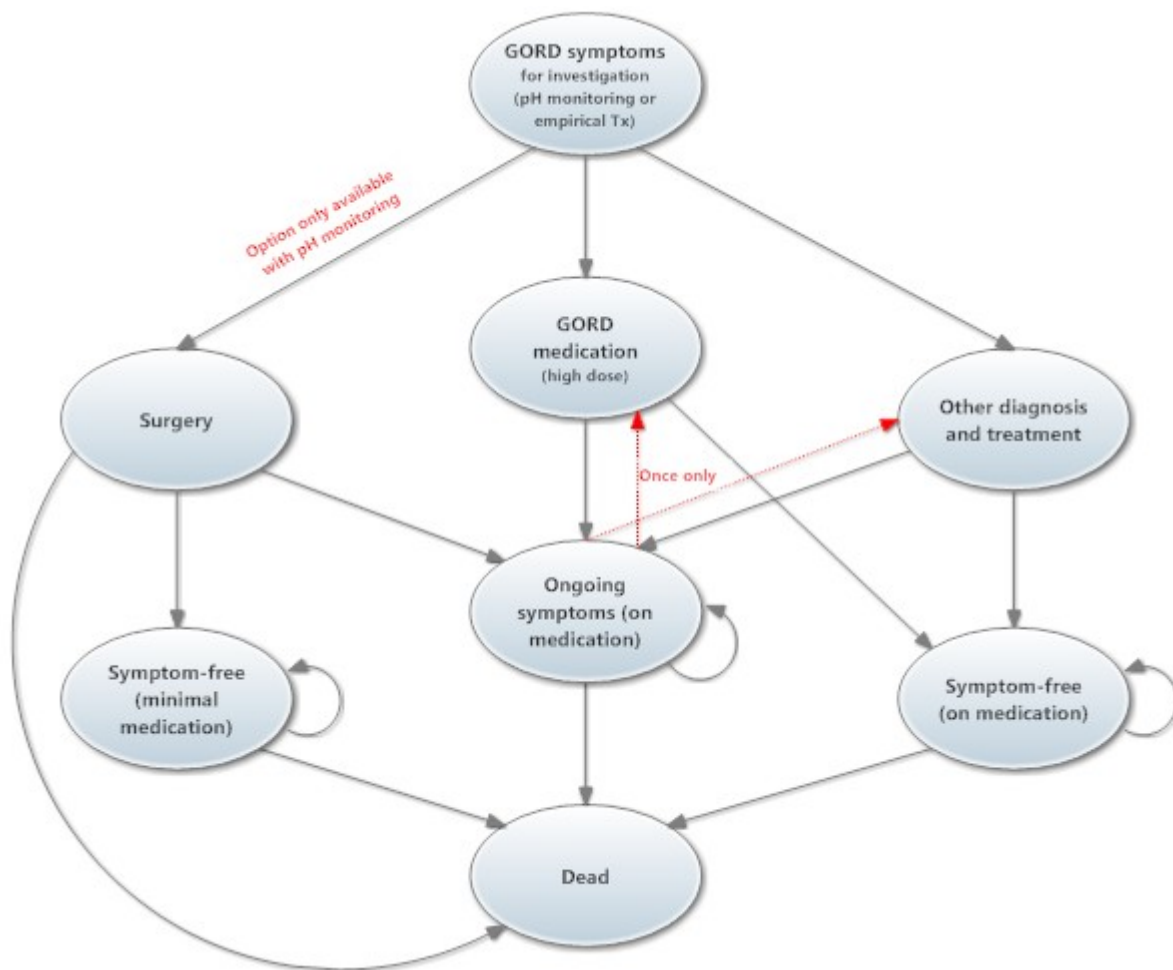
Figure 8   Health state diagram for the economic models

The model is constructed in two phases:

Phase 1: The first phase represents the first year following the decision to pH monitor or not. In this year patient diagnoses and management strategies are established. Patients are assumed to enter phase 1 of the model with symptoms and be receiving routine daily average-dose GORD medication. Depending on the model scenario, patients are then assigned to receive either:

- wireless pH monitoring **or** no monitoring (ongoing medical treatment) **[Model 1]**

or

- wireless pH monitoring **or** catheter-based monitoring **[Model 2]**

A summary of the structure of the mechanics of the economic model is presented in Table 38.

Table 38   Summary of the economic evaluation (applies to both models)

Time horizon	Base-case analysis of 15 years
Outcomes	Quality-adjusted life years
Methods used to generate results	Markov model (with half-cycle correction)
Cycle length	1 year
Discount rate	5% for both costs and outcomes
Software packages used	TreeAge Pro and Excel (hybrid)

It is assumed that only patients who have a positive pH test result will be eligible for surgery, and the subsequent uptake of surgery by a percentage of patients following pH monitoring is a key difference between the intervention and the comparator in Model 1. In patients who have a normal pH-test result it is assumed that other diagnoses are investigated and treated. Patients who do not get a pH-test result, due to test failure or allocation to the comparator group, and who do not respond to high-dose PPIs may be investigated for other diagnoses and receive other treatment. However, expert opinion suggests that this would not occur universally in all patients—a number would simply remain as unsuccessfully treated and refractory NERD cases. This assumption is a key driver in the model. The extent to which alternative diagnoses are sought without a negative pH test result is varied and tested in the sensitivity analysis. The specific values associated with all transitions between health states are discussed in the relevant sections under 'Inputs to the economic evaluation'.

The initial phase of each model has six cycles of 2 months, and at the end of the 1-year period it is assumed that diagnoses and treatment paths have been determined.

Phase 2: The second phase calculates the long-term costs and outcomes associated with the patient's diagnosis and treatment as established at the end of the first year of investigation and treatment (phase 1). It is assumed that no new diagnoses or changes in treatment are made in phase 2 and that all patients either remain in the health state in which they ended phase 1, or they die (at the background mortality rate—applied equally across all patients). Costs and outcomes associated with treatment options are calculated in yearly cycles with a 5% annual discount rate. The base-case accrues costs and outcomes for a further 14 years (i.e. total model duration of 15 years); however, timeframes from 1 year to 25 years are tested in the sensitivity analysis.

The decision analytic structure for the model pathways is shown in Figure 9 .
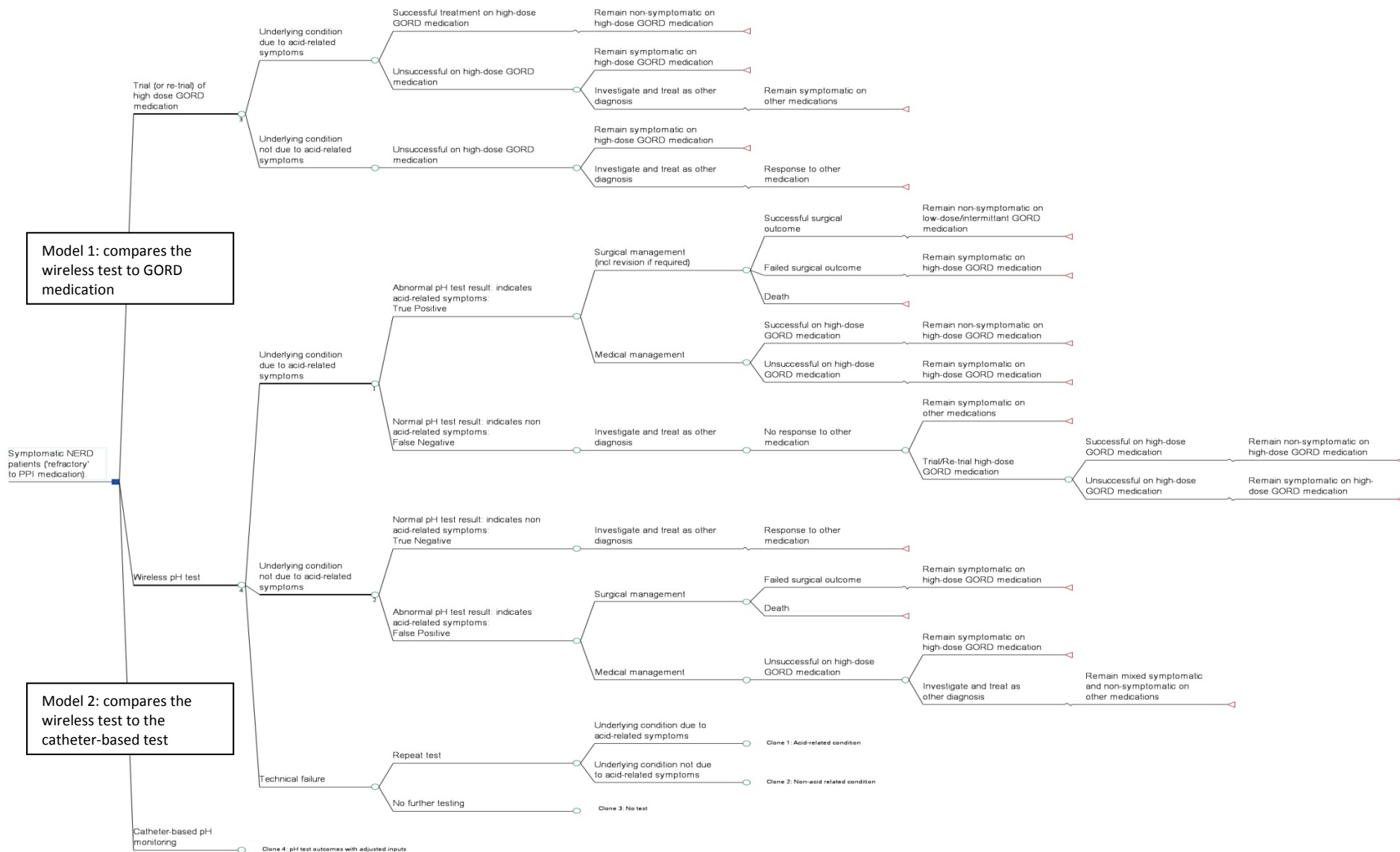
Figure 9   Decision analytic structure of the economic evaluation. Model 1 represents the decision between wireless pH monitoring and non-selective high-dose GORD PPI therapy. Model 2 represents the decision between wireless pH monitoring and catheter-based monitoring.

Assumptions incorporated into the model structure:

- Only one test failure—patients failing a pH-monitoring test obtain a result on a subsequent test.

- Patients would not tolerate a manometric pH test; therefore, this is not incorporated anywhere in the model as an alternative or subsequent diagnostic tool. Likewise, endoscopic procedures for management of GORD are not included as it is assumed that fundoplication is the standard surgical technique and there is a lack of evidence regarding efficacy of the alternative treatments.

- Underlying disease is delineated by acid reflux into the oesophagus or not. While this is not a true reflection of the real population, which has mixed/multiple disease aetiologies, it is considered a justifiable simplification for use in the model. Although segregated in the model, the clinical implications of mixed aetiologies are incorporated in the model results because it is assumed that there is an imperfect success rate associated with acid treatments, and the cost of some ongoing acid treatments (PPIs) is incorporated into the expected costs of treatment for alternative diagnoses.

- Surgery is repeated in 10% of patients and assumed to occur within the same year as the initial surgery. This is consistent with Australian MBS data on surgical revisions (Grant, Cotton et al. 2013). The diagrammatic structure suggests that only one repeat surgery may be undertaken per patient; however, this strict interpretation is not essential. The final outcome from surgery, either success or failure, is based on overall surgical population outcomes that capture all revisions.

- Patients that have a diagnosis of acid-related symptoms and who fail to resolve their symptoms with surgery remain on a double dose of PPIs for the remainder of the model. Likewise, patients that have a diagnosis of acid-related symptoms and elect medical management remain on a double dose of PPIs for the remainder of the model, irrespective of whether or not they have successful symptom control. The assumption of a high-dose requirement is consistent with the population initially being considered 'refractory' (Fass 2009).

Inputs to the economic evaluation

A summary list of all the variables used in the model is presented in Table 85 in Appendix K.

Population and test parameters

Prevalence, test diagnostic yield and failure rate

There is no direct evidence concerning the true underlying prevalence of acid-related GORD in the symptomatic, endoscopy-negative patient population who would be referred for pH monitoring in the Australian setting. Assuming that the catheter-free pH-monitoring test is accurate, prevalence is estimated based on its diagnostic yield. The clinical report provides a broad range of yields, between 25.7% (Lacy et al. 2011) and 78.3% (Domingues, Moraes-Filho & Domingues 2011), although none are in a population that exactly matches that defined in the model. For the purposes of the model, the prevalence of acid-related GORD is estimated to be 40% of the endoscopy-negative population for monitoring. This estimate is broadly consistent (within 10%) with numerous studies (Ayazi et al. 2009, 2011; Bhat, McGrath & Bielefeldt 2006; Grigolon et al. 2011; Scarpulla et al. 2007) and also with a study in a similarly defined population that used pH manometry as an alternative diagnostic test (Broeders et al. 2010). The systematic review results presented earlier identified that where comparative evidence was available between 48-hour wireless monitoring and 24-hour catheter-based monitoring, a (median) additional yield of 7.8% was associated with the longer duration of the wireless test (see Table 34). Clinical opinion supports the assumption that the additional yield is accurate and represents true positive patients. However, given that this cannot be verified with an alternative reference standard, it is only incorporated into a supplementary analyses for model 2 (see below).

The test is assumed to have a failure rate of 7.5%, based on the median technical failure rate reported in 'Is it effective – Technical efficacy'.

Test accuracy

The systematic review presents varying results to describe test accuracy, depending on the reference standard, pH cut-off used and duration of monitoring. Given the heterogeneity of the test accuracy studies, a meta-analysis was not considered appropriate. Although the DAP specifies that 24-hour catheter-based monitoring should be considered as the reference standard, the repeated and clinically accepted finding of an increased yield of accurate positive results from the wireless pH-monitoring studies run over 48 hours (see 'Results of assessment') does not support the use of catheter-based monitoring as a perfect reference standard.

In the economic models the base-case analysis assumes perfect accuracy (i.e. 100% sensitivity and specificity) associated with wireless pH monitoring, given the lack of a reliable reference standard. However, supplementary exploratory analyses were also undertaken:

(i) (for models 1 and 2) using test accuracy data where catheter-based monitoring is considered the reference standard. The highest quality study available (level II diagnostic evidence) indicated that, against this reference, the catheter-free test, with a cut-off point of 4.4% of the time that pH <4, had a sensitivity of 86.4% and specificity of 77.8% (see 'Is it accurate?')

(ii) (for model 2) assuming that catheter-free monitoring is 100% accurate, but the sensitivity of catheter-based monitoring is 80.5% and the specificity 100%. This assumes that catheter-free monitoring has better test performance than catheter-based monitoring, based on the 7.8% increased diagnostic yield (accepted to be true positives) associated with 48-hour catheter-free monitoring over catheter-based monitoring, i.e. subtracting the 7.8% additional yield from the diagnostic yield, to assume that 24-hour catheter-based monitoring would only identify 32.2% (rather than 40%) of patients as pH positive.

Implications for true (and false) positive and negative test results

Where patients are positively diagnosed with acid-positive NERD, either medical or surgical management options may be appropriate—both strategies are considered to offer an acceptable level of symptom resolution. Numerous clinical comparisons and cost-effectiveness analyses have been published in the literature, from varying perspectives and with varying findings, such that it would be beyond the scope of this report to conclusively determine a single 'preferred' treatment route. The model assumes that both treatment options result in an 89% success rate with respect to providing satisfactory symptom management. This effectiveness rate for surgery is supported in the literature (Beck et al. 2009; Broeders et al. 2010), and it may be considered a conservative assumption that appropriate medical management is equally effective, although of some inconvenience given the requirement for regular medication. This is broadly consistent with the findings of the REFLUX trial, which found similar but marginally higher QoL outcome scores in patients assigned to surgical management over medical management (Grant, Boachie et al. 2013).

In the base-case, medical management of confirmed acid reflux (e.g. a positive pH-monitoring test) is assumed to consist of high-dose (twice daily) PPIs and a once-daily H2 antagonist (i.e. ranitidine). The assumption that medical treatment is successful in refractory patients appears contradictory—it is contingent on the assumption that the initial patient

population has been selected using a 'loose' definition of 'refractory'. But this is consistent with the literature, which reports that while many GORD patients are labelled 'refractory' to standard doses of PPIs, many respond to twice-daily PPIs, and in rare cases (e.g. fast metabolisers) response rates can be achieved with even high doses and/or combinations with additional products (e.g. H2 antagonists, antacids). Nor it is uncommon that medication is not taken in an optimal dosing manner. If this is investigated and addressed, cases previously considered refractory may also have satisfactory medical treatment results (Fass 2009; Gunaratnam et al. 2006; Scarpignato 2012). Thus, it would appear that many supposedly 'refractory' patients can have successful medical treatment following a confirmation of their diagnosis.

The implications associated with test results for patients in the model are depicted in Table 39 and Table 40.

Table 39   Implications for true positive and true negative test results for patients in the economic evaluation (base-case)

True health state	pH-monitoring test result	Implications
NERD due to acidic reflux (positive)	Positive (true)	Option of medical management (ongoing high-dose PPIs +/– an H2 antagonist) or surgical management
Symptoms not due to acid reflux (e.g. FH, HO, OeO) (negative)	Negative (true)	Alternative diagnoses investigated and successfully treated, as specified in the DAP

NERD = non-erosive reflux disease; PPIs = proton pump inhibitors, H2 = histamine 2 (receptor); FH = functional heartburn; HO = hypersensitive oesophagus; OeO = oesinophilic oesophagitis

Table 40   Additional implications for false positive and false negative patients in the supplementary economic evaluation[6]

True health state	pH-monitoring test result	Implications
NERD due to acidic reflux (positive)	Negative (false)	Alternative diagnoses investigated and other treatments trialled Alternative treatment is assumed to be 0% effective in patients incorrectly identified to have an alternative diagnosis
Symptoms not due to acid reflux (e.g. FH, HO, OeO) (negative)	Positive (false)	Option of medical management (ongoing high-dose PPIs +/– an H2 antagonist) or surgical management GORD treatment is assumed to be 0% effective in patients incorrectly identified to be acid-symptom positive patients

NERD = non-erosive reflux disease; PPIs = proton pump inhibitors, H2 = histamine 2 (receptor); FH = functional heartburn; HO = hypersensitive oesophagus; OeO = oesinophilic oesophagitis

---

[6] Noting that, in the base-case analyses of Model 1 and Model 2, catheter-free monitoring is assumed to have perfect test performance, i.e. there will be no false positives or false negatives.

Test or treatment uptake rates

Catheter-free monitoring uptake

The model assumes that patients who do not wish to undertake wireless monitoring will not participate or be charged for it and, as such, they are not relevant or included in the cost-effectiveness model. Therefore, for the purposes of assessing cost-effectiveness, it is assumed that all patients in the intervention arm initially take up wireless monitoring. However, there is a technical failure rate of 7.5% (see 'Technical efficacy') and therefore some patients will not get a meaningful test result. The model assumes that half of the patients who experience a technical failure will re-attempt the test and subsequently get a result, and half will not re-attempt the test. This ratio is arbitrary and tested in the sensitivity analyses.

Medication uptake

The model assumes that where medical management is the selected treatment, patients adhere to recommended treatment regimens and purchase medication accordingly. This applies to situations both where medication is prescribed for suspected or confirmed GORD and where other medications are prescribed for a non-GORD diagnosis (see 'Healthcare resources' below for a description of assumed medication regimens). Strictly speaking, this assumption may overestimate medication usage but, given that patients are by definition symptomatic and apparently seeking relief, it would seem reasonable.

Surgery uptake

Based on the literature regarding fundoplication surgery for reflux, it is considered that this surgery is only appropriate for patients who have a positive pH-monitoring test result. The estimate that 25% of patients who have a positive result have subsequent surgery is based on a calibration of the estimated diagnostic yield (40%), with available Medicare data[7] indicating that in Australia 10% of patients who undergo pH monitoring have subsequent surgeries.

---

[7] Medicare data from financial year 2011–12 showed that Item 11810 (Catheter-based pH monitoring) was claimed 2,857 times, and a surgery item (30527, 30530, 31464, 31466, 43951 or 43954) was co-claimed for the same patient 323 times (i.e. ~10% of patients) in the subsequent 12 months. Of these surgeries, 31 are revision surgeries (Item 31466). *Unpublished data requested from the Australian Government Department of Health.*

Surgery revision (re-operation)

Medicare Australia data showed that 10% of fundoplication surgery was revised. This rate is incorporated into the model as the rate of uptake of repeat surgery. This re-operation rate is consistent with many literature reports (Beck et al. 2009; Grant et al. 2008: 11.3%; Broeders et al. 2010: 13.6%). Lower rates, as claimed in some publications (Anvari et al. 2011: 7.8%; Mahon et al. 2005: 3.7%; REFLUX trial (Grant, Cotton et al. 2013): 4.5%) are tested in a sensitivity analysis.

Investigations/treatment for alternative diagnosis

It is unknown what proportion of patients remain symptomatic while on PPIs and then, on returning to the doctor, further diagnostic investigations or alternative medications are prescribed. The literature suggests that many patients prescribed PPIs do not have follow-up re-evaluation, and ongoing inappropriate use of PPIs is continued (Heidelbaugh, Goldberg & Inadomi 2010; Heidelbaugh et al. 2012). The assumption that many patients remain on PPIs rather than have alternative diagnoses investigated is supported by HESP clinical opinion[8].

There is little evidence to quantify the extent of this occurrence. Considering that the patient group is already those seeking attention for refractory symptoms, and specialists are involved, a conservative estimate that only 10% of patients will not be reassessed and will continue high-dose PPIs indefinitely was applied in the economic models' base-case for the purposes of the assessment—i.e. 90% of patients that are non-responsive to high-dose PPIs will not continue medication and be prescribed alternative treatments. This assumption is tested in a sensitivity analysis.

Clinical outcomes

Treatment outcomes

Table 41 summarises the treatment outcomes associated with the treatment pathways in the model. For the purpose of modelling QALYs, effectiveness refers to the extent that the different treatments result in patients being free of, or satisfactorily managed for, the GORD or GORD-like symptoms for which they initially sought diagnosis and treatment.

---

[8] Prof. Batey, email 18 October 2013; Prof. Fraser, email 1 November 2013.

Table 41   Health outcomes associated with treatment pathways in the model

Management strategy	Effectiveness rate	Source
Surgical management (fundoplication) in acid-related NERD	89% (range tested:50%–100%)	(Beck et al. 2009; Broeders et al. 2010; Grant, Cotton et al. 2013)
Medical management (high-dose PPIs and H2 antagonists) in acid-related NERD	89% (range tested:50%–100%)	Conservatively assumed equivalent to surgery and broadly consistent with the REFLUX trial (Grant, Cotton et al. 2013)
Medical management: treatment for alternative diagnosis (non–acid related symptoms)	100%	DAP specified—assumed
Medical management (high-dose PPIs and H2 antagonists) or surgical management (fundoplication) in symptomatic medical conditions that are NOT due to acid-reflux (i.e. applied in scenario of false positive test result)	0%	Assumed
Medical management: using treatment for alternative diagnoses (non–acid related conditions) for symptoms of true acid-related NERD (i.e. applied in scenario of false negative test result)	0%	Assumed

PPIs = proton pump inhibitors; H2 = histamine 2 (receptor); NERD = non-erosive reflux disease

Mortality

A low rate of mortality directly due to fundoplication surgery is incorporated into the model. This rate is estimated at 5 deaths per 10,000 patients (0.05%), based on the REFLUX trial, chapter III (Grant et al. 2008). Although described as a rate/patient, this rate is assumed to be derived from patient episodes, and therefore the event of an additional revision surgery is assumed to carry additional mortality risk. The same mortality rate is applied to surgery revision (10% of patients), such that the overall surgical mortality risk is 0.055% per patient.

Some models do not include a surgical mortality rate at all (implicitly 0% rate), but Epstein et al. (2009) apply a rate of 4/3,397 (0.118%). These alternative estimates are tested as upper and lower limits in the sensitivity analysis.

A background mortality rate is applied equally to all patients in each year of the model. This is a time-dependent transition probability, derived from the Australian Bureau of Statistics general Australian population mortality rates, in a 40%:60% male:female ratio, beginning at age 55 years and advancing 1 year of age with each year of the model.

Healthcare resources

Healthcare resources are included at the full MBS or Pharmaceutical Benefits Scheme (PBS) fee, i.e. including standard minimum patient co-payments, on the grounds that the patient and any mandatory co-payment are part of the societal healthcare system and the full fee is the accepted 'resource cost'. Additional out-of-pocket patient contributions above the

published MBS fee are not included in the economic model; these are regarded as privately negotiated arrangements that are external to the healthcare system. Consideration of the impact of the listing of catheter-free pH monitoring on government costs (i.e. excluding patient co-payments) and total costs (i.e. including patient co-payment and gap costs) is considered in the financial analysis (see 'Financial and costing impact'). The modelled healthcare resource use and costs are presented in Table 42.

Table 42   Testing-related healthcare resources used in the economic model

Type of resource item	Unit cost	Components/source of unit cost	Application
Wireless pH monitoring (one test)	$1,682	Proposed fee $913.64 Plus Anaesthesia: (MBS items 17610; $43, 20740; $99, 23010; $19.80) Facility fee: (AR-DRG Z40Z) same day: $607	All patients receiving test, plus repeated cost in % of patients with initial test failure who elect to repeat test.
Regular-dose PPIs (per monthly prescription)	$27.66	Weighted average DPMQ of omeprazole 20mg, esomeprazole 20mg, esomeprazole 40mg and pantoprazole 40mg (based on Medicare usage data 2012–13)	Assumed regimen of patients entering the model. In the intervention arm, this regimen is maintained during the first 2 months while pH monitoring is carried out and before future treatment decisions are made on the basis of the test results
High-dose PPIs (per monthly prescription)	$48.69	Twice the DPMQ of the average normal dose (weighted combination of omeprazole, esomeprazole and pantoprazole—as above—less 1x pharmacist dispensing fee)	Assumed to be used as medical management strategy: (i) as empirical treatment in comparator arm, and (ii) in all patients diagnosed after pH monitoring as having an acid-related condition
Low-dose or intermittent-use PPIs (per monthly prescription)	$9.22	One-third of the cost of daily PPIs, on the basis that only one-third of patients require daily medication after surgery.	Following successful fundoplication surgery (no ongoing symptoms), an ongoing medication cost is still assumed
Treatment of alternative diagnosis (per monthly prescription(s))	$47.27	Average (unweighted) DPMQ cost of treatments reported as potentially used for alternative diagnoses (baclofen, TCA, SSRI, metoclopramide, domperidone etc.): $19.61, plus assumed to be taken with low-dose PPI: $27.66 = monthly drug costs: $47.27	Assumed to be used as management strategy in all patients considered to have a condition that is not acid-related (i.e. after negative pH monitoring and in a proportion of patients who fail high-dose PPIs and have reassessment.)

Type of resource item	Unit cost	Components/source of unit cost	Application
Fundoplication surgery (total costs per episode of surgery)	$6,583	Initial: Surgery MBS items – 119: $43, 31464: $871, 51303: $174 Anaesthesia (~2 h 30 min) MBS items – 17610: $43, 20706: $138.60, 23111: $218 Facility fee (AR-DRG G03C) ALOS 2.7 days: $5,138	Assumed to be used as the surgical management strategy in patients who choose this management strategy after they have been diagnosed as having an acid-related condition, confirmed with pH monitoring.
	$7,106 (in 10% patients)	Revision: Surgery MBS items – 119: $43, 31466: $1,307, 51303: $261 Anaesthesia (~2 h 30 min) MBS items – 17610: $43, 20706: $138.60, 23111: $218 Facility fee (AR-DRG G03C) ALOS 2.7 days: $5,138	
Catheter-based monitoring (one test)	$174.45	MBS item 11810 (includes results)	All patients receiving test, plus repeated cost in % of patients with initial test failure who elect to repeat test.

DPMQ = Dispensed price maximum quantity (for PBS prescription); PBS = Pharmaceutical Benefits Scheme; PPI = proton pump inhibitor; TCA = tricyclic antidepressant; SSRI = selective serotonin reuptake inhibitor; ALOS = average length of stay

Utility values

At all times, regardless of the specific management, the economic model assumes that patients are in one of three health utility states:

- experiencing GORD or GORD-like symptoms, i.e. unsuccessfully treated;
- free of GORD or GORD-like symptoms, i.e. successfully treated; or
- dead.

The utility weights associated with the health states and the sources of these utilities are presented in Table 43.

Table 43   Health state utilities used in the economic evaluation

Health state	Utility weight	Source
Unsatisfactorily managed (normal medication refractory) NERD (or apparent NERD)	0.56	Baseline in REFLUX trial (0.72) less decrement for symptoms in Heudebert et al. (1997) (0.18) Approach as per Bojke, Hornby & Sculpher (2007)
Satisfactorily managed NERD (or apparent NERD)	0.78 (2)	REFLUX trial post-surgery (average 2–5 years) (Grant, Cotton et al. 2013)
Death	0	Assumed

However, depending on the treatment pathway a patient follows, utility adjustments (decrements) are applied to account for the side effects or encumbrances relating to GORD treatment. These are described in Table 44.

Table 44   Utility adjustments used in the economic evaluation

Treatment event	Utility weight	Source(s)	Application
Surgery—acute inconvenience	–0.5 (decrement) for 2 weeks/surgery = a –0.02 QALY (decrement/episode of surgery)	Comay et al. (2008; Heudebert et al. (1997)	Given that each patient is assigned 1.1 episodes of surgery, to account for the 10% revision rate, each patient undergoing surgery acquires a decrement of –0.022 QALYs to account for the decreased utility associated with revision surgeries
Daily medication decrement	–0.01 (decrement each year on medication)	Arguedas et al. (2004; Comay et al. (2008)	Applied to all patients without GORD symptoms but requiring ongoing medication to maintain symptom relief, to account for inconvenience of regular medication requirement It is noted that this value is less than the difference between post-surgery and medical management groups in REFLUX (Grant, Cotton et al. 2013); therefore, this may be considered a conservative estimate.

QALY = quality-adjusted life years

It is assumed that:

- all patients with refractory NERD and refractory NERD-like symptoms[9] have the same utility weight while symptomatic, i.e. when entering the model or during the model if they are unsuccessfully treated. Likewise, all successfully treated patients, either true NERD or those with an alternative diagnosis, have the same symptom-free utility weight

- all symptomatic patients are taking medication, although without adequate effect, and that the disutility of medication is incorporated into the disutility associated with being symptomatic

- all patients with a negative pH-test result and/or diagnosis other than NERD will require other ongoing treatment for their alternative diagnosis.

Outputs from the economic evaluation

Incremental costs and effectiveness

The overall costs and outcomes, and incremental costs and outcomes, as calculated for the intervention and comparator in each model, with the base-case assumptions, are shown in Table 45 and Table 46.

---

[9] i.e. including those who have a negative acid test because symptoms are due to underlying conditions other than NERD, such as functional heartburn, hypersensitive oesophagus and oesinophilic oesophagitis.

Table 45   Model 1: Cost-effectiveness of wireless pH monitoring vs no monitoring (empirical treatment) in patients with refractory NERD

	Cost	Incremental cost	Effectiveness (QALYs)	Incremental effectiveness	ICER
Intervention	$8,705		7.874		
Comparator	$6,927	$1,778	7.751	0.123	$14,457

QALY = quality-adjusted life years; ICER = incremental cost-effectiveness ratio

Table 46   Model 2: Cost-effectiveness of wireless pH monitoring vs catheter-based pH monitoring in patients with refractory NERD

	Cost	Incremental cost	Effectiveness (QALYs)	Incremental effectiveness	ICER
Intervention	$8,705		7.874		
Comparator	$7,136	$1,569	7.878	–0.004	Dominated

QALY = quality-adjusted life years; ICER = incremental cost-effectiveness ratio

Although the base-case comparison of catheter-free (wireless) pH monitoring and catheter-based pH monitoring assumes equivalent efficacy, accuracy and subsequent treatment decisions following positive or negative test results, wireless pH monitoring results in fewer gains in health outcomes. This is due to the increased technical failure rate associated with wireless monitoring (7.5% vs 0.89%), and the subsequent increased number of 'no result' patients (50% of test failures) who revert to empirical treatment, i.e. the comparator in Model 1.

Weighted combinations of models 1 and 2

As neither Model 1 nor Model 2 are predicted to represent the actual patient market, the following analysis considers combinations of these models, which include all patients in the proposed listing combined with various extents of leakage into the existing catheter-based monitoring market. The relative sizes of these markets are discussed under 'Background' and 'Financial and costing impact', where it is estimated that the existing catheter-based market is approximately 10 times the proposed catheter-free market. Therefore, a 10% leakage of catheter-based-monitoring patients to catheter-free monitoring will result in a 50%:50% ratio of catheter-free patients representing Model 1 and 2. A summary of the cost-effectiveness of the intervention where this leakage occurs, and of further leakage scenarios is presented in Table 47.

Table 47   Societal ICERs following the proposed listing, allowing for various levels of leakage to include patients tolerant of catheter-based pH monitoring accessing wireless monitoring

Extent of catheter-based monitoring leakage (% of catheter market) [a]	Weighting of catheter-based [a]	Weighting of 'no test'	Weighted costs [b]	Weighted QALYs [c]	ICER [d]
0%	0%	100%	$6,927	7.751	$14,457
10%	50%	50%	$7,032	7.815	$28,129
20%	67%	33%	$7,066	7.836	$42,694
30%	75%	25%	$7,084	7.846	$58,429
40%	80%	20%	$7,094	7.853	$75,278
50%	83%	17%	$7,101	7.857	$93,436
60%	86%	14%	$7,106	7.860	$113,061
70%	88%	13%	$7,110	7.862	$134,339
80%	89%	11%	$7,113	7.864	$157,487
90%	90%	10%	$7,115	7.865	$182,764
100%	91%	9%	$7,117	7.866	$210,478

[a] The catheter-suitable market is estimated at 10x the size of the catheter-intolerant market; therefore, a 10% leakage in this market represents an equal number of patients to the intolerant market etc.

[b] Weighted combination of comparator, catheter-based: 7.878 QALYs; and no pH monitoring: 7.751 QALYs

[c] Weighted combination of comparator, catheter-based: $7,136; and no pH monitoring: $6,927

[d] In all cases the ICER is calculated against the proposed intervention of 100% wireless monitoring, which, on average, has a cost of $8,705 and accrues 7.874 QALYs (per patient).

The combination analysis above suggests that the cost-effectiveness of the proposed listing is highly dependent on limiting the leakage of use in patients who could otherwise tolerate catheter-based monitoring. If this leakage is relatively small (e.g. under 20%), the proposed listing is more likely to be considered cost-effective in typical Australian ICER interpretation, whereas if there is extensive leakage (e.g. greater than 50%), the proposed listing is unlikely to be considered cost-effective in broadly accepted terms.

Scenario and sensitivity analyses

Many variables in the models are uncertain, with estimates derived from mixed sources other than local Australian data, i.e. literature reports, expert opinion and arbitrary estimates; therefore, the impacts of these variables are examined in sensitivity analyses. The key sensitivity analyses relate to the accuracy of the wireless pH test, the extent that doctors will look for other diagnoses and discontinue high-dose PPIs in non-responders and the timeframe of the cost–utility analysis. Other variables and inputs are also tested.

Sensitivity analyses in Model 1: Wireless pH monitoring vs no monitoring (empirical treatment) in patients with refractory NERD

The cost-effectiveness of wireless monitoring vs no monitoring is dependent on the accuracy, particularly the specificity, of the test (Table 48). Using the accuracy data estimated in the report, which is known to be imperfect given the imperfect reference standard (catheter-based monitoring) and inconsistent test result interpretation criteria, wireless monitoring is dominated by the comparator in the sensitivity analysis, thus producing fewer health outcomes at a higher cost. This is found in all analyses conducted where the test specificity of catheter-free monitoring is less than 95%. However, the cost-effectiveness of the wireless test hardly changes, relative to the base-case, when the sensitivity of the test is reduced to 86% as long as 100% specificity is maintained.

Table 48　　Sensitivity analyses around wireless pH-test accuracy for Model 1

	Catheter-free test: Cost ($)	Catheter-free test: Outcomes (QALYs)	No pH test: Cost ($)	No pH test: Outcomes (QALYs)	ICER
Base-case (100% accurate)	$8,705	7.874	$6,927	7.751	$14,457
With sensitivity of 86% and specificity of 77%	$9,060	7.577	$6,927	7.751	Dominated
With sensitivity of 100% and specificity of 77%	$9,168	7.587	$6,927	7.751	Dominated
With sensitivity of 86% and specificity of 100%	$8,597	7.863	$6,927	7.751	$14,912
With sensitivity of 90% and specificity of 90%	$8,829	7.742	$6,927	7.751	Dominated
With sensitivity of 95% and specificity of 95%	$8,767	7.803	$6,927	7.751	$35,388

QALY = quality-adjusted life years (gained); ICER = incremental cost-effectiveness ratio

There are no data available to estimate the extent to which untested patients in this symptomatic population remain on PPIs despite the lack of efficacy and potentially an alternative diagnosis and treatment being more appropriate. It is suggested that the base-case estimate, where only 10% of patients remain inappropriately followed up, is conservative. Alternative assumptions are shown in Table 49. The cost-effectiveness of wireless monitoring increases substantially if it is assumed that more than 10% of patients remain on inappropriate treatment, although at no time does it become less expensive. However, if active follow-up of alternative diagnoses and discontinuation of inappropriate PPIs occurred in 99% of patients, pH monitoring would be unlikely to be considered cost-effective. If 100% of patients were followed up, the wireless pH test would be dominated, as empirical medication treatment would achieve greater outcomes at less cost.

Table 49   Sensitivity analyses around the assumption of PPI discontinuation for Model 1: Cost-effectiveness of wireless pH monitoring vs no monitoring (empirical treatment) in patients with refractory NERD

% of non-responders to high-dose PPIs that remain on inappropriate treatment	Catheter-free test: Cost ($)	Catheter-free test: Outcomes (QALYs)	No pH test: Cost ($)	No pH test: Outcomes (QALYs)	ICER
100% (all non-responders remain on high-dose PPI treatment)	$8,741	7.83	$7,982	6.593	$613
75%	$8,731	7.842	$7,624	6.915	$1,194
50%	$8,721	7.855	$7,356	7.237	$2,209
25%	$8,711	7.867	$7,088	7.558	$5,253
Base-case: 10% of non-responders remain on high-dose PPIs and 90% discontinue	$8,705	7.874	$6,927	7.751	$14,457
5%	$8,703	7.876	$6,874	7.816	$30,486
2.5%	$8,702	7.877	$6,847	7.848	$63,971
1%	$8,701	7.878	$6,831	7.867	$170,014
0% (all non-responders to high-dose PPIs investigated/treated for alternative diagnoses)	$8,701	7.879	$6,820	7.880	Dominated

NERD = non-erosive reflux disease; ICER = incremental cost-effectiveness ratio, PPI = proton pump inhibitors; QALY = quality-adjusted life years (gained)

An analysis of cost-effectiveness when considered across different time horizons was also undertaken (Table 50). The cost-effectiveness of pH monitoring is substantially improved as time horizons of a longer duration are considered. This is driven by the fact that the most significant costs are incurred in the first year (i.e. costs of the test and of subsequent surgery), as is the QALY decrement associated with surgery; however, both cost savings due to reduced expenditure on high-dose PPIs, and QALY benefits associated with discontinued use of medication, accrue every year thereafter. The ICER does not change substantially after 20 years, given that discounting and background mortality reduce the overall impact of distant years.

Table 50   Sensitivity analyses around the time horizon measured in Model 1: Cost-effectiveness of wireless pH monitoring vs no monitoring (empirical treatment) in patients with refractory NERD

	Catheter-free test: Cost ($)	Catheter-free test: Outcomes (QALYs)	No pH test: Cost ($)	No pH test: Outcomes (QALYs)	ICER
1 year	$3,011	0.73	$681	0.729	$2,330,150
5 years	$5,125	3.382	$3,000	3.336	$46,199
10 years	$7,178	5.957	$5,252	5.868	$21,642
Base-case: timeframe 15 years	$8,705	7.874	$6,927	7.751	$14,457
20 years	$9,814	9.265	$8,143	9.119	$11,446
25 years	$10,582	10.229	$8,987	10.067	$9,847

QALY = quality-adjusted life years (gained); ICER = incremental cost-effectiveness ratio

Additional sensitivity analyses undertaken on Model 1 showed that the model was not particularly sensitive to discounting, surgical mortality rates, technical failure rates, re-test uptake rate after technical failure, fundoplication surgery costs, or the cost of investigation and treatment of alternative diagnoses. The results of these analyses are presented in Table 86 in Appendix K.

Sensitivity analyses in Model 2: Wireless pH monitoring vs catheter-based monitoring

Sensitivity analyses around test accuracy were also undertaken for Model 2 (Table 51). The wireless test is dominated by the less expensive and more technically reliable catheter-based test, even when assuming perfect accuracy of the wireless test (compared with catheter-based monitoring); predictably, for all analyses where the wireless test has less-accurate test parameters, it remains dominated. Where the catheter-based test is assumed to be imperfect, with a sensitivity of 0.805, but other (including all wireless) test accuracy parameters are assumed perfect, the wireless test has a positive but relatively high ICER. This suggests that the transfer of patients from catheter-based to wireless monitoring when it is not needed is unlikely to be cost-effective, irrespective of whether or not the clinical yield and sensitivity of wireless pH monitoring is higher than catheter-based monitoring. This is consistent with the wording of the proposed listing to restrict funding to patients who are unable to tolerate catheter-based monitoring.

Table 51   Sensitivity analyses around the accuracy of wireless pH monitoring and catheter-based monitoring in Model 2

	Catheter-free test: Cost ($)	Catheter-free test: Outcomes (QALYs)	Catheter-based test: Cost ($)	Catheter-based test: Outcomes (QALYs)	ICER
Base-case (100% accurate)	$8,705	7.874	$7,136	7.878	Dominated
With wireless test sensitivity of 86% and specificity of 77% (catheter-based test reference standard, assumed 100% accurate)	$9,060	7.577	$7,136	7.878	Dominated
Where catheter-based monitoring is assumed to have a sensitivity of 80.5% and all other accuracy parameters are 100%	$8,705	7.874	$6,980	7.863	$156,818
With wireless sensitivity of 100% and specificity of 77%	$9,168	7.587	$7,136	7.878	Dominated
With wireless sensitivity of 86% and specificity of 100%	$8,597	7.863	$7,136	7.878	Dominated
With wireless sensitivity of 95% and specificity of 95%	$8,767	7.803	$7,136	7.878	Dominated

QALY = quality-adjusted life years (gained); ICER = incremental cost-effectiveness ratio

Financial and costing impact

The estimated financial costs associated with wireless pH monitoring have been determined using a combined market share and epidemiological approach. This involves consideration of the existing catheter-based pH-monitoring market, and also epidemiological estimates with regard to the extent of catheter intolerance (i.e. failure of catheter-based monitoring or anatomically inappropriate).

Data sources used in the financial analysis

Existing usage and costing data relating to existing catheter-based pH monitoring (item 11810) and endoscopy was provided by Medicare Australia, on request.

The estimated percentage of patients <1 or >70 years of age (5–10%) is based on Medicare Local data concerning item 11810.

The estimate of placement error is based on the analysis presented under 'Technical efficacy' on page 75.

Estimates of patient catheter intolerance (~10%) are based on expert opinion and are supported by the literature (Sweis et al. 2009).

Use and costs of the proposed wireless pH-monitoring strategy

Proposed fee

The proposed fee per occasion of wireless pH monitoring, as requested in the application, is $913.64. This fee has been calculated by the applicant on the basis that it includes:
- professional time (estimated at 50 minutes with a value of $350):
  - performing the procedure (~27 minutes)
  - downloading and reading of data (~20 minutes)
- equipment costs (estimated at $563.64):
  - consumable pH transmitter capsule (Bravo) $430.30
  - depreciation of capital—wireless receiver and recording system, including software (Bravo) $19,990, depreciated fully over 3 years/150 patients at $133.34/occasion.

A significant proportion (62%) of the proposed fee is for equipment. It is noted that even after adjustment for currency exchange, the quoted equipment prices remain considerably higher than the 2006 prices published by the American Society for Gastrointestinal Endoscopy Technology Committee (Chotiprashidi et al. 2005) presented in

Table 87 in Appendix L, despite a general trend in other medical technology prices to have decreased over this time.

Of concern is that the calculated depreciation of capital requires each doctor providing the service to perform 50 services each year to break even with respect to equipment costs. The predicted current patient demand for services, based on the restricted listing, estimates that less than 400 tests should be required per year. This would suggest that the market would only support 8 practitioners around Australia to purchase the equipment and provide the service whereas, currently, catheter-based pH monitoring is undertaken by at least 56 practitioners (Medicare Australia 2013). Reduced numbers of practitioners may either result in access difficulties or, if most practitioners who currently undertake catheter-based monitoring were to invest in a wireless pH-monitoring system, there would be economic pressure to either increase the patient cost beyond the schedule fee (i.e. have a large gap payment) or to over-service ('leakage').

In addition, a technical failure rate of approximately 7.5% is associated with the wireless pH-monitoring procedure, most commonly involving placement failure of the wireless capsule without results, such that an additional capsule is required. Clarification is needed as to who (e.g. doctor, patient, private insurer, manufacturer) bears this cost—capsule (~$430) or procedure ($914).

Other issues of note that relate to the proposed fee include the following:
- Automatic annual indexation may not be appropriate given the equipment component.
- If the Bravo capsule component funding were to be obtained through an alternative source (e.g. the Department of Health's Prostheses List), revision of the fee would be required.

It is anticipated that the service would be typically undertaken in the inpatient setting, such that the MBS would rebate 75% of the fee ($685.23) and the patient contribution would be $228.32.

Associated resources

In addition, with each occasion of wireless pH monitoring, additional resources, from both MBS and healthcare facility providers, beyond the procedure item will be incurred. These relate to the requirement for anaesthesia while the capsule insertion takes place, and are listed in Table 52 and Table 53. Both the placement of the capsule by endoscopy, and the later interpretation of results, are included in the proposed fee; therefore, neither an

endoscopy nor follow-up consultation for interpretation of results can be claimed with the proposed item.

Table 52   Additional MBS item resources expected to be used concurrently with all occasions of the proposed intervention

MBS item	Description	Full MBS cost (fee)	MBS benefit [a]
17610	ANAESTHETIST, PRE-ANAESTHESIA CONSULTATION (Professional attendance by a medical practitioner in the practice of ANAESTHESIA) A BRIEF consultation involving a targeted history and limited examination (including the cardio-respiratory system). (etc)	$43.00	$32.25
20740	INITIATION OF MANAGEMENT OF ANAESTHESIA for upper gastrointestinal endoscopic procedures	$99.00	$74.25
23010	ANAESTHESIA, PERFUSION OR ASSISTANCE AT ANAESTHESIA (a) administration of anaesthesia performed in association with an item in the range 20100 to 21997… (etc) For a period of: FIFTEEN MINUTES OR LESS	$19.80	$14.85
	Total MBS resources	$161.80	$121.35

[a] 75% benefit applied, assuming most claims are in the inpatient setting

Table 53   Other resources expected to be used concurrently with all occasions of the proposed intervention

Type of resource	Description	Average cost
Healthcare facility cost	Endoscopy and follow-up care (AR-DRG Z40Z)	$607

Furthermore, at least 5% of patients currently obtaining pH monitoring are 75 years or older (Medicare data), and it might be estimated that another 5% are infants or between the ages of 70 and 74 years, such that up to 10% of patients would be eligible for the anaesthesia age-modifier fee. Also, in cases where insertion is difficult or a technical failure occurs during insertion but can be rectified without abandoning the procedure, increased fees associated with an increased duration of anaesthesia might be expected. The relevant item numbers associated with these occasional additional resources are shown below in Table 54.

Table 54 Additional MBS item resources expected to be used concurrently in approximately 5%–10% of the proposed interventions

MBS item	Description	Full cost (fee)	MBS benefit [a]
25015	ANAESTHESIA, PERFUSION OR ASSISTANCE AT ANAESTHESIA - where the patient is less than 12 months of age or 70 years or greater	$19.80	$14.85
23021	[From 23010: ANAESTHESIA….(a) administration of anaesthesia performed in association with an item in the range 20100 to 21997… (etc)] For a period of: 16 MINUTES TO 20 MINUTES	$39.60	$29.70
	Total MBS resources	$59.40	$44.55
	Average additional MBS expenditure/patient (5–10%)	$2.97–$5.94	$2.23–$4.46

[a] 75% benefit applied, assuming most claims are in the inpatient setting

It is plausible that on very rare occasions serious complications or side effects associated with endoscopy and sedation may occur (e.g. oesophageal tear, haemorrhage). However, there is no quantitative report of the occurrence of these complications associated with wireless pH monitoring and, due to their unknown but expected very low incidence, they are not included in the costing assessment. Likewise, the MBS costs associated with potential downstream treatment decisions (e.g. fundoplication surgery), although included in the economic model, are not included in the costing analysis, given that they are not directly or routinely associated with the proposed intervention.

For the remainder of the costing analysis, an associated MBS expenditure of $166 per patient in addition to the proposed fee will be assumed, equal to the sum of $161 + $5—rounded values from Table 52 and Table 54. Of this, $125 is assumed to be paid as benefits and $41 as a patient co-payment.

A summary of the total costs and funding sources directly associated with the proposed intervention are presented in Table 55. Without knowledge of the actual fees that doctors intend to charge for the proposed listing, it is not possible to calculate the additional patient out-of-pocket payments.

Table 55 Total costs directly associated with a service of wireless pH monitoring under the proposed MBS listing

Description	MBS rebate (benefits paid)	Patient MBS co-payment	Private insurance or patient cost	Total
Proposed listing	$685.32	$228.32	$0	$913.64
Associated MBS items	$125	$41	$0	$166
Facility fee	$0	$0	$607	$607
Total contributions	$810.32	$269.32	$607	$1,686.64

Estimated volume per year

The proposed listing is restricted to use in patients who are indicated for but cannot tolerate a catheter-based pH test. It is estimated that this equates to approximately 5%–10% of the number of patients who receive catheter-based pH monitoring (Sweis et al. 2009). To estimate the number of patients eligible for wireless monitoring under the proposed listing, the upper end of the estimate (10%) is applied to the number of catheter-based patients from existing Medicare data. The number of claims for catheter-based monitoring has increased by an average of approximately 4% each year over the past 4 years (2009–13), and therefore in the base-case it is assumed that an ongoing annual growth rate of 4% will apply to all pH monitoring. An estimate of the number of pH tests expected to be undertaken using this calculation is presented in Table 56.

Table 56   Projected number of wireless pH tests annually under the proposed listing (base-case estimates)

	2012–13	2013–14 [a]	2014–15 [a]	2015–16 [a]	2016–17 [a]
Base-case					
Number of catheter-based tests	3,590	3,734	3,883	4,038	4,200
Estimated number of wireless tests (10% of catheter-based pH tests)	359	373	388	404	420
Total pH tests	3,949	4,107	4,271	4,442	4,620

[a] projected to increase annually at 4%

There is uncertainty around the estimated growth rate; if only the most recent data is considered (2011–13), catheter-based monitoring has increased annually by 8%. If this annual growth rate is maintained, costs associated with listing will increase more rapidly. The impact of an increased growth rate on the number of tests is tested in a sensitivity analysis in Table 57.

There is also considerable uncertainty with respect to whether 'leakage' beyond the eligibility criteria for wireless monitoring will occur. The incentives for use include:

- Patient preference: there are reports that some patients feel considerable embarrassment and discomfort with the transnasal catheter in situ, and they may avoid attending work or leaving their home. The wireless monitoring system, on the other hand, is relatively discreet, with most patients reporting that they can maintain normal daily activities during the monitoring interval. Therefore, this may result in some patients requesting wireless monitoring and claiming unsuitability for catheter-based monitoring. This 'unsuitability' would be based on patient preference rather than an 'absolute' intolerance (Hirano 2006; Mearin et al. 1998).

- Clinical claim: despite the uncertain conclusion with respect to comparative accuracy and effectiveness, if doctors or patients are convinced that the wireless system reports an additional 10% accurate positive yield with respect to providing a diagnostic result, there is a clinical incentive to opt for wireless monitoring.

- Financial incentive for doctors: the proposed fee structure includes a component to cover capital depreciation of the pH-monitoring system, assuming that doctors have invested a significant capital outlay. This inherently results in an economic incentive for the doctor to favour this service rather than a strictly time-funded one, as a means to recoup their equipment costs (see earlier discussion under 'Proposed fee')

Countering these incentives is the fact that wireless pH monitoring requires an endoscopy procedure with sedation. The need for sedation, and therefore admittance to a hospital and third-party observation plus transport home afterward, involves considerable additional risk, money and inconvenience for the patient, such that it will not always be the preferred alternative.

There are no data available to make a quantitative estimate as to the extent that leakage to wireless monitoring in place of catheter-based monitoring is likely to occur in practice. For the purposes of a sensitivity analysis, the financial analysis presents cost estimates where the respective ratio of wireless:catheter-based pH monitoring is reversed, arbitrarily making a 91%:9% ratio, with the intention that this be considered the upper estimate of plausible leakage. Estimated numbers of wireless tests under this scenario are also calculated in Table 57.

Table 57   Sensitivity analyses for the projected number of wireless pH tests under the proposed listing

	2012–13	2013–14 [a]	2014–15 [a]	2015–16 [a]	2016–17 [a]
1: Increased annual growth rate of 8%					
Number of catheter-based test [a]	3,590	3,877	4,187	4,522	4,884
Estimated number of wireless tests (10% of catheter-based pH tests)	359	388	419	452	488
2: Assuming extensive leakage beyond the restriction					
Number of catheter-based tests			388	404	420
Number of wireless tests (upper limit)			3,883	4,038	4,200
Total pH tests			4,271	4,442	4,620

[a] projected to increase annually at 8%

In the majority of patients it is anticipated that pH monitoring would only occur once in their lifetime. However, in a minority of patients who continue to have symptoms despite

different medication regimens or surgery, repeat monitoring (e.g. on/off treatment) may be indicated. The MBS claims data does not differentiate between individual one-off patient tests and repeat tests in the same patient; therefore, assuming that repeat monitoring would occur similarly across both pH-monitoring technologies, the extent of repeat monitoring will not affect the costing estimates.

There is a significant technical failure rate with wireless pH monitoring that is generally associated with capsule placement. It is assumed that where a capsule has not been placed successfully, an MBS claim would not be submitted.

Estimated MBS costs per year

The direct costs to the MBS per year associated with the proposed listing—specifically for the new item—for directly associated items and in total are calculated in Table 58. The government benefits paid and patient contributions are also estimated.

Table 58   Total cost of the proposed intervention to the MBS (base-case)

	2013–14 [a]	2014–15 [a]	2015–16 [a]	2016–17 [a]
Number of patients eligible for wireless test (Table 56)	373	388	404	420
Fees of the proposed item number ($913.64) [b]	$341,117	$354,761	$368,952	$383,710
MBS benefits (75%)	$255,871	$266,106	$276,750	$287,820
Patient co-payments	$85,246	$88,655	$92,202	$95,890
MBS fees of associated items ($166)	$61,978	$64,457	$67,035	$69,717
MBS benefits (75%)	$46,670	$48,537	$50,478	$52,497
Patient co-payments	$15,308	$15,920	$16,557	$17,219
Total MBS fees associated with listing	$403,094	$419,218	$435,987	$453,426
MBS benefits payable (75%)	$302,541	$314,643	$327,228	$340,318
Patient co-payments	$100,553	$104,575	$108,758	$113,109

[a] estimated projections
[b] MBS benefits payable per patient round up to $810.32, which results in a total benefits calculation slightly higher than 75% of the total fees.
MBS = Medicare Benefits Schedule

Sensitivity analyses around the MBS costs allowing for (i) increased growth and (ii) an estimate of the upper limit of plausible leakage are presented in Table 59 and Table 60, respectively.

Table 59   Sensitivity analysis: Projected MBS costs associated with wireless pH monitoring if patient numbers grow at increased annual rate of 8%

	2014–15	2015–16	2016–17
Estimated number of patients receiving wireless test with 8% annual growth rate	419	452	488

Fees of the proposed item number ($913.64) [a]	$382,575	$413,181	$446,236
MBS benefits (75%)	$286,969	$309,927	$334,721
Patient co-payments	$95,606	$103,255	$111,515
MBS fees of associated items ($166)	$69,510	$75,071	$81,077
MBS benefits (75%)	$52,342	$56,530	$61,052
Patient co-payments	$17,168	$18,542	$20,025
Total MBS fees associated with listing	$452,086	$488,253	$527,313
MBS benefits payable (75%)	$339,311	$366,456	$395,773
Patient co-payments	$112,774	$121,796	$131,540

[a] MBS benefits payable per patient round up to $810.32, which results in a total benefits calculation slightly higher than 75% of the total fees.
MBS = Medicare Benefits Schedule

Table 60   Sensitivity analysis: Projected MBS costs associated with potential leakage (upper plausible extent) of wireless pH monitoring beyond the restriction (i.e. use in 90% of patients who are tolerant to catheter-based monitoring)

	2014–15	2015–16	2016–17
Estimated number of patients receiving wireless test, including upper limit of plausible leakage	3,883	4,038	4,200
Fees of the proposed item number ($913.64) [a]	$3,547,664	$3,689,278	$3,836,849
MBS benefits (75%)	$2,661,098	$2,767,322	$2,878,015
Patient co-payments	$886,567	$921,956	$958,834
MBS fees of associated items ($166)	$644,578	$670,308	$697,120
MBS benefits (75%)	$485,375	$504,750	$524,940
Patient co-payments	$159,203	$165,558	$172,180
Total MBS fees associated with listing	$4,192,242	$4,359,586	$4,533,970
MBS benefits (75%)	$3,146,473	$3,272,072	$3,402,955
Patient co-payments	$1,045,770	$1,087,514	$1,131,015

[a] MBS benefits payable per patient round up to $810.32, which results in a total benefits calculation slightly higher than 75% of the total fees.
MBS = Medicare Benefits Schedule

Changes in use and cost of current monitoring strategy

Technically, given that the listing is restricted to patients who cannot tolerate catheter-based monitoring, the comparator is 'ongoing medical care without a diagnostic test and empirical treatment'. However, given the diversity of medical care and treatments that may be afforded (briefly described in 'Economic evaluation') and the lack of any evidence or suggestion that the proposed test will directly alter the actual number of consultations, the base-case comparator for the purposes of the financial analysis is simply 'no pH test'.

Therefore, in the base-case assumption for the proposed listing of wireless monitoring, public funding would be in addition to existing levels of catheter-based pH monitoring, and

the extent of catheter-based monitoring would remain unchanged, i.e. there would be no cost offsets associated with the new test listing. Similarly, market growth is independent and the rate of market growth is not relevant.

However, if wireless pH monitoring was to be undertaken in place of catheter-based monitoring, i.e. 'leakage' beyond the listing (as described under 'Use and costs of the proposed wireless pH-monitoring strategy'), the extent of use of catheter-based monitoring would be expected to decrease 1:1 with the extent of increased 'leaked' use of wireless monitoring. In this sensitivity analysis scenario, an upper plausible limit of leakage to wireless monitoring, the estimated change in use and associated costs of catheter-based monitoring are calculated in Table 61. The listed MBS fee for catheter-based monitoring (item 11810) is $174.75 (benefit: 75% = $130.85, 85% = $148.30). However, Medicare data from 2012–13 report that the average fee per service is $190.80, and the average benefit paid per service is $154.81, 96% being for out-of-hospital services and a significant quantity attracting additional Medicare Safety Net benefits; therefore, these values are used in the calculations.

Table 61   Sensitivity analysis: Changes in MBS costs of catheter-based monitoring associated with upper plausible limit of potential leakage to wireless pH monitoring beyond the restriction

	2013–14[a]	2014–15[a]	2015–16[a]	2016–17[a]
Extent of current monitoring without proposed listing, or with proposed listing and no leakage Projected use of catheter-based pH monitoring (number of claims)	3,734	3,883	4,038	4,200
Total projected fees:	$712,371	$740,876	$770,450	$801,268
Estimated Medicare benefits payable (based on existing benefit paid/test	$577,999	$601,127	$625,123	$650,128
Estimated patient contributions (based on existing contributions/test)	$134,372	$139,749	$145,328	$151,141
With wireless test MBS-listed and significant leakage Projected use of catheter-based pH monitoring (number of claims)	373	388	404	420
Total projected fees:	$71,237	$74,030	$77,083	$80,167
Estimated Medicare benefits payable	$57,800	$60,066	$62,543	$65,045
Estimated patient contributions	$13,437	$13,964	$14,540	$15,122
Estimated difference (decrease) in number of claims for catheter-based monitoring if listing and leakage occurs	–3,360	–3,495	–3,634	–3,779
Estimated difference (saving) in fees for catheter-based monitoring	$641,134	$666,846	$693,367	$721,102
Difference (saving) in Medicare benefits payable	$520,199	$541,061	$562,580	$585,083
Difference (saving) in patient contributions	$120,935	$125,785	$130,788	$136,019

[a] estimated projections

MBS = Medicare Benefits Schedule

Financial implications to the Medical Benefits Schedule (MBS)

In the base-case there is no evidence of direct or quantifiable cost offsets associated with the proposed wireless pH test. Therefore, assuming no leakage beyond the proposed restriction, the overall costs to the MBS remain those associated with the listing, as calculated in Table 58. The total value of the MBS fees charged in the restricted population is estimated to grow annually, from approximately $420,000 (in 2014–15) to $455,000 (in 2016–17), and would be expected to exceed $500,000 by 2020. In 2014–15 the MBS would be expected to fund approximately $323,000 of this amount, with patient contributions of $97,000, increasing by 2016–17 to MBS funding of approximately $350,000 and patient contributions of approximately $105,000.

Uncertainty scenarios

If the market growth rate increases annually by 8%, the overall MBS costs are as presented in Table 59, with annual MBS costs increasing by an additional approximately $50,000 (2014–15) to $75,000 (2016–17).

For the scenario analysis allowing for leakage beyond the restriction (i.e. from catheter-based monitoring) to an upper plausible extent, the overall MBS costs will increase substantially. These costs are calculated in Table 62 and indicate that if substantial leakage occurs, total MBS fees could be increased by over $3.5 million in the 2014–15 financial year, with further increases annually. Government-funded benefits would increase to approximately $2.7 million and patients would additionally contribute approximately $850,000, annually, with ongoing growth expected.

Table 62   Sensitivity analysis: Net change in MBS costs associated with the proposed listing of wireless pH monitoring, assuming the upper plausible extent of leakage beyond the restriction

	2014–15 [a]	2015–16 [a]	2016–17 [a]
MBS costs associated with proposed listing	$4,192,242	$4,359,586	$4,533,970
MBS benefits (75%) [b]	$3,146,473	$3,272,072	$3,402,955
Patient co-payments	$1,045,770	$1,087,514	$1,131,015
Savings in fees for substituted catheter-based monitoring (calculated in Table 61)	–$666,846	–$693,367	–$721,102
Difference (saving) in Medicare benefits payable	–$541,061	–$562,580	–$585,083
Difference (saving) in patient contributions	–$125,785	–$130,788	–$136,019
Net change in total MBS costs	$3,525,396	$3,666,219	$3,812,868
Net change in MBS benefits	$2,605,412	$2,709,492	$2,817,872
Net change in patient contributions	$919,985	$956,726	$994,996

[a] estimated projections; [b] MBS benefits payable per patient round up to $810.32, which results in a total benefits calculation slightly higher than 75% of the total fees.
MBS = Medicare Benefits Schedule

Clearly, the potential for leakage is significant, and the financial ramifications to the MBS are high should such leakage occur.

Other Australian healthcare system costs

Costs to the state and territory health systems

Although the proposed listing will occur in the hospital setting, it is assumed that MBS claims will only be made for patients admitted in private hospitals, and no additional public hospital funding is envisaged. It is currently unknown whether this technology is available to public hospital patients, and therefore accessibility may present an equity issue.

Costs to the private health insurer and/or patient

As identified in Table 53, the proposed listing does require utilisation of a hospital facility, at cost to either a private health insurer or the patient. The additional costs to the insurer or patient associated with hospitalisation are calculated in Table 63.

Table 63   Costs of hospitalisation incurred by private insurer or patient (base-case)

	2014–15[a]	2015–16[a]	2016–17[a]
Number of patients eligible for wireless test (Table 56)	388	404	420
Hospital costs associated with test ($607/patient)	$235,695	$245,122	$254,927

[a] estimated projections

Sensitivity analysis

The additional hospitalisation costs that would be borne by private health insurers and/or patients according to (i) the alternative annual market growth rate (8%) and (ii) significant leakage beyond the restriction (i.e. replacing catheter-based tests) are estimated in Table 64.

Table 64   Costs of hospitalisation incurred by private insurer or patient—alternative scenarios

	2014–15[a]	2015–16[a]	2016–17[a]
(i) Increased market growth rate			
Number of patients eligible for wireless test (Table 56)	419	452	488
Hospital costs associated with test ($607/patient)	$254,174	$274,508	$296,468
(ii) Significant market leakage			
Number of patients having wireless test (Table 56)	3,883	4,038	4,200
Hospital costs associated with test ($607/patient))	$2,356,981	$2,451,066	$2,549,109

MBS = Medicare Benefits Schedule

Total Australian healthcare system costs

The total costs to the Australian healthcare system are presented in Table 65.

Table 65   Total Australian healthcare system costs (base-case)

	2014–15[a]	2015–16[a]	2016–17[a]
Number of patients eligible for wireless test (Table 56)	388	404	420
MBS costs (Table 58)	$419,218	$435,987	$453,426
Private insurer/patient costs (Table 63)	$235,695	$245,122	$254,927
Total	$654,913	$681,109	$708,353

MBS = Medicare Benefits Schedule

Table 66 also calculates the total healthcare system costs (i) assuming an increased annual market growth rate of 8% (vs base-case of 4%) and (ii) assuming significant market leakage.

Table 66   Total Australian healthcare system costs—alternative scenarios

	2014–15[a]	2015–16[a]	2016–17[a]
(i) Increased market growth rate			
Number of patients eligible for wireless test (Table 56)	419	452	488
Net MBS costs (Table 58)	$452,086	$488,253	$527,313
Private insurer/patient costs (Table 64)	$254,174	$274,508	$296,468
Total healthcare costs	$706,260	$762,760	$823,781
(ii) Significant market leakage			
Number of patients having wireless test (Table 56)	3,883	4,038	4,200
Net MBS costs (Table 62)	$3,525,396	$3,666,219	$3,812,868
Private insurer/patient costs (Table 64)	$2,356,981	$2,451,066	$2,549,109
Total healthcare costs	$5,882,377	$6,117,285	$6,361,977

MBS = Medicare Benefits Schedule

Discussion

Safety

The population that would normally receive catheter-free monitoring in Australia comprises patients who cannot tolerate a transnasal catheter or in whom a catheter is anatomically inappropriate. Therefore, the appropriate comparator would be 'no pH monitoring', since this patient group does not receive pH monitoring in current practice. However, due to the absence of evidence on catheter-free monitoring in the right patient population and with the specified comparator, the selection criteria were amended to include patients suspected of GORD but who can possibly tolerate a transnasal catheter, which allowed the inclusion of catheter-based monitoring as a comparator.

Of the 47 comparative and non-comparative studies identified in the literature, only one compared catheter-free monitoring with no monitoring; and only one included patients who were unable to tolerate catheter-based monitoring, with a group undergoing catheter-based monitoring as a comparator. Francis et al. (2012) used 'no monitoring' as a comparator when investigating safety issues on catheter-free monitoring 16 cm above the gastro-oesophageal junction (GOJ); however, 48 hours before proximal capsule placement a more distal capsule was placed approximately 6 cm above the GOJ. This means that there is a large probability that the distal capsule was still in the oesophagus at the time of placement of the proximal capsule, possibly causing additional discomfort since most capsules stay in place for more than 48 hours. One conclusion from this study is that a proximal capsule placement in addition to a 'normal' capsule placement can cause additional discomfort.

Sweis et al. (2009) compared a group of patients who could not tolerate catheter-based monitoring and underwent catheter-free monitoring with patients suspected of GORD undergoing catheter-based monitoring. The catheter-free monitoring patients were significantly more satisfied with the test than the group undergoing catheter-based monitoring. This shows that patients who cannot tolerate a catheter are still capable of tolerating catheter-free monitoring. Therefore, even though there was only limited evidence on the population of interest that was outlined *a priori*, we can assume that this population is able to undergo catheter-free monitoring.

Most adverse events reported in the included studies were related to discomfort caused either by the catheter or by the capsule attaching to the oesophageal mucosa. More-severe complications from catheter-free monitoring, such as a rupture in the oesophageal mucosa,

are often caused by a technical failure or human error where the capsule fails to detach from the catheter while inserting the device. Capsule aspirations or dislodgements were reported in case reports and can be dangerous if they dislodge in the lung. This highlights the usefulness of case reports in describing safety outcomes; rare complications are unlikely to be reported by controlled studies or case series. Therefore, case reports can play an essential role in identifying complications from catheter-free monitoring, but caution should be exercised while commenting on rare complications based on case reports that have been presented for catheter-free monitoring but not for its comparator(s).

There was limited data on adverse events of catheter-free monitoring in children. It is not clear if children are less likely to tolerate the transnasal catheter but, if this is the case, data on the safety of catheter-free monitoring in children is essential since they will constitute a considerable share of the catheter-free monitoring users. Croffie et al. (2007) reported more chest discomfort with the new test but less overall discomfort in children, compared with the catheter-based test. However, in four case series with a total of 410 children, two mucosal tears were reported using catheter-free monitoring, so severe adverse events do occur in children when using catheter-free monitoring.

The percentage of patients who reported pain and discomfort attributable to catheter-free monitoring varied greatly between studies. In the non-comparative studies that reported chest pain or discomfort as an adverse event, rates varied between 0.27% and 45%. Some comparative studies reported a median discomfort score on a VAS, some used a different scale and others listed percentages of patients with adverse events. It is possible that patients would report some chest pain on a VAS, but would not report chest pain when the question 'Do you have any side effects from monitoring?' was asked. Results may be biased owing to different methods of reporting adverse events, variation in the physician's experience in performing the procedure, and small sample sizes within the studies included in this assessment.

Effectiveness

Two studies were identified that reported on the effectiveness of catheter-free (wireless) pH monitoring in terms of health outcomes. However, these studies did not meet the specific inclusion criteria *a priori*, as patients were those who could have potentially tolerated catheter-based monitoring. Health outcomes were similar between those monitored by the catheter-based test and those by the catheter-free test. Given the assumption that catheter-based monitoring is considered superior to no monitoring in those who can tolerate catheter-based monitoring, and the finding that the catheter-free system

appears as good as the catheter-based system, it is expected that catheter-free monitoring improves health outcomes compared with no monitoring. However, the available direct evidence was limited and of low quality, so this justified the use of a supportive linked evidence approach, i.e. assessing the diagnostic accuracy of catheter-free monitoring and its impact on patient management.

Nine studies compared technical failures in catheter-free monitoring with catheter-based monitoring failure rates. Even though only one study identified catheter-free monitoring as being at higher risk for technical failures, when a meta-analysis was conducted the relative risk was 3.3 (95%CI 1.63, 6.81; I^2 = 0%, p=0.906; k = 8 studies) for the occurrence of technical failures while using catheter-free monitoring compared with catheter-based monitoring. The most commonly reported technical failures that caused problems in catheter-free monitoring were early capsule detachments (i.e. in the wrong place in the oesophagus or where the capsule fails to attach to the oesophageal mucosa), causing incomplete data capture and errors in placement of the system.

A concern was raised that oesophageal acid exposure exhibited day-to-day variability. If this is the case, monitoring for 48 hours instead of 24 hours, as with traditional catheter-based monitoring, would give truer and more-accurate results. However, even though there were 20 studies reporting on day-to-day variability in adults, there was no consistency in their results. Fourteen studies reported no significant difference in oesophageal acid exposure when comparing the first day with the second day of monitoring, and three and two studies reported more acid exposure on day 1 and day 2, respectively. But there is evidence that increasing the monitoring period leads to an increased diagnostic yield: the incremental diagnostic yield when increasing the monitoring time from 24 hours to 48 hours varied from 2% to 15.6%, as reported by six studies. Because catheter-free monitoring has a standard monitoring period of 48 hours, it should be able to correctly diagnose more people than the standard catheter-based method, where a 24-hour monitoring period is used. Catheter-based monitoring was used as a reference standard for diagnostic accuracy in the current assessment; however, as there is an incremental diagnostic yield with catheter-free monitoring, the validity of this reference standard should be questioned.

Sensitivity, specificity, PPV, NPV, LR+ and LR– were used in some studies to determine how well catheter-free monitoring classifies patients in having or not having GORD based on an arbitrary cut-off point, and for some studies the sensitivity and specificity could be calculated based on the available data. Catheter-based monitoring and clinical diagnosis were used as reference standards in three and five studies, respectively. A diagnostic test is

ideal if its sensitivity, specificity, NPV and PPV all reach 100%, although this assumes a perfect reference standard and perfect concordance. Diagnostic tests in practice vary in this regard. For those compared against the potentially imperfect reference standard of catheter-based monitoring, lower sensitivity and specificity could be due to catheter-free monitoring being more accurate. Furthermore, the threshold for a positive/negative test result changes the characteristics of the index test. When the threshold for GORD diagnosis in a specific test is set lower, the sensitivity and NPV will be higher and the specificity and PPV will be lower. Decisions on determining the optimal balance of sensitivity and specificity require compromise.

Currently, there is a lack of consensus as to the optimal cut-off value for diagnosis in catheter-free and catheter-based monitoring, and therefore the eight studies reporting sensitivity and specificity vary greatly in thresholds for diagnosis, which in turn leads to differences in specificity, sensitivity, NPV and PPV. With no consistency in cut-off values, a meta-analysis on diagnostic accuracy was not conducted. The majority of patients are already being treated with PPIs prior to the test, so a high sensitivity and NPV are important: the true value of the test may be ruling out those who don't have the disease to reduce the number of patients with false negative results. Consensus on an appropriate cut-off value is therefore necessary to diagnose GORD consistently across different settings when using either catheter-free or catheter-based monitoring.

When establishing a GORD diagnosis, not only does a cut-off value in acid exposure have to be determined, but people also have to agree on how the cut-off value will be used. For instance, a positive diagnosis can be based on one positive day (day 1 or day 2), two positive days (having a score above the threshold during both days), a positive score on the 'worst day' of monitoring, or a reflux symptom correlation (symptom index, sensitivity index or associated probability). In addition, whatever is decided as the cut-off value, the result should be used to guide management, and there was evidence that this is not always done. Furthermore, there should be one standard location of the capsule in the oesophagus. Most physicians agree on placing the capsule approximately 5 cm above the LOS or 6 cm above the SCJ, but there is no evidence stating that this is the optimal location for diagnosing GORD. Historically, this position was chosen for catheter-based monitoring because it was close enough to the SCJ to sense acid reflux and yet sufficiently distant from the stomach so that the electrode could not migrate too low and end up in the stomach itself (Pandolfino et al. 2006). Often, reflux-related mucosal injuries, such as oesophagitis and Barrett's metaplasia, are localised closer to the SCJ. With catheter-free monitoring the capsule can be attached to the oesophageal mucosa at any desired level that maintains a consistent

position, so further discussion could be had regarding whether a more distal capsule placement would have a greater relevance to GORD diagnosis. A study by Wenner et al. (2008) showed improved diagnostic accuracy for pH monitoring at 1 cm, compared with 6 cm, above the GOJ in patients with typical reflux symptoms.

Nine studies reported on concordance of oesophageal acid exposure times comparing catheter-free monitoring with catheter-based monitoring. There was no clear consistency in acid exposure time concordance, and discrepancy between the two monitoring methods was particularly clear in studies that used in-subject simultaneous recording. This was unexpected because the acid exposure times should closely match when measured in the same subject and at the same time, adjusting for all confounders.

Pandolfino et al. (2005) investigated the differences in acid exposure and concluded that there was a consistent offset in pH measurement between the two tests in healthy volunteers. This offset was confirmed further by an in-vivo reference measurement using orange juice with a known pH. The dominant source of discrepancy between the two monitoring systems was an inaccuracy in catheter-free monitoring electrode calibration; after adjustment, the discrepancy improved by 40%. However, some differences still existed and were most pronounced during short reflux events.

This corresponds to the results of this assessment, where three out of four studies on concordance in reflux events reported more reflux events in the catheter-based system. One possible explanation is that short drops in pH can be caused by factors unrelated to acid reflux (Fox 2006); for example, due to the ingestion of mildly acidic fluids, movement of the catheter relative to the mucosa in catheter-based monitoring, and movement of the catheter relative to the GOJ during swallowing. On swallowing, the oesophagus shortens by a couple of centimetres due to muscle contractions. As it shortens, the catheter electrode in catheter-based monitoring moves towards the GOJ and may pass into the stomach where the pH is low. The catheter electrode can dip into the 'acid pocket' before the catheter returns to its original position. Other short pH drops can occur with peristaltic contractions or during transient LOS relaxation, and resolve without swallowing. Therefore, many of the short reflux events measured by catheter-based monitoring and not catheter-free monitoring could be artefacts that are not related to GORD.

When pH monitoring indicates that there is pathological acid reflux, patients could undergo anti-reflux surgery. Two systematic reviews were identified to answer the question whether anti-reflux surgery for GORD leads to better health outcomes than medical treatment. Improved health outcomes were more common after surgical treatment compared with

medical treatment in the short to medium term according to both systematic reviews. Two trials reported 5-year follow-up results and concluded that most patients remain in remission for 5 years after anti-reflux surgery, and this group continued to have better pain relief than the medication group. However, these trials consisted solely of patients who responded to PPI therapy, and it is not known if the results generalise to patients who do not respond to medical therapy.

There were no studies identified that provided relevant information on the implications of non-treatment on false negative results. False negative patients would probably go off PPI therapy, and would either not undergo surgery or would delay surgical treatment, even though they have GORD. They would likely undergo investigations for alternative diagnoses. It was assumed that the management of false positive patients is not likely to change, although two studies mentioned a patient receiving anti-reflux surgery after a false GORD diagnosis (Sweetman, Ng & Kerrigan 2007; Taddio et al. 2011). Furthermore, patients with a false positive result would probably have a delay in treatment for their 'real' diagnosis, as they are misdiagnosed with GORD. Two studies on infants have been identified regarding delayed treatment after GORD misdiagnosis that confirmed the assumptions: infants with infantile spasms and a delayed diagnosis/treatment had a poorer outcome and a worse treatment response. Although these children never underwent pH monitoring and it is not known if they were eligible for the procedure, it shows that early diagnosis in children suspected of GORD can lead to better health outcomes.

The body of evidence included in this assessment report was appraised according to the NHMRC's methodological guidelines (NHMRC 2008). This appraisal considered the evidence-base, consistency of the studies' results, clinical relevance of the safety and effectiveness data, generalisability of the evidence to the population of interest, and applicability to the Australian healthcare system.

Table 67 presents the results of the assessment of the body of evidence for catheter-free monitoring in GORD.

Table 67   Body of evidence matrix

Component	A	B	C	D
	Excellent	Good	Satisfactory	Poor
Evidence-base [a]			One or two level III studies with a low risk of bias, or level I or II studies with a moderate risk of bias	

Consistency [b]		Most studies consistent and inconsistency may be explained		
Clinical impact			Moderate (with regard to safety)	Slight or restricted (with regard to change in management)
Generalisability			Population(s) studied in body of evidence differ from target population for guideline but it is clinically sensible to apply this evidence to target population [c]	
Applicability		Applicable to Australian healthcare context with few caveats		

Source: adapted from NHMRC (2008)

[a] Level of evidence determined from the NHMRC evidence hierarchy (Table 9)

[b] If there is only one study, rank this component as 'not applicable'.

[c] For example, results in adults that are clinically sensible to apply to children OR psychosocial outcomes for one cancer that may be applicable to patients with another cancer.

Economic considerations

Estimated cost-effectiveness

The economic evaluation suggested that catheter-free (wireless) pH monitoring, while not cost-saving, was reasonably cost-effective vs no monitoring in patients who could not tolerate catheter-based monitoring (base-case ICER: $14,457/QALY). However, it was not cost-effective if patients could have used catheter-based monitoring as an alternative (ICER: Dominated—more costly and less effective). The economic evaluations required numerous assumptions and uncertain inputs. Where inputs could not be derived from the systematic review or published literature, expert opinion or plausible estimates were required, and a conservative approach, not favourable to a finding of cost-effectiveness, was employed. Most inputs, although uncertain, were not found to be drivers with respect to cost-effectiveness, but two key determinants of the cost-effectiveness were identified:

- The accuracy of the pH test in detecting acid-related conditions. If wireless monitoring is assumed to be of equivalent accuracy to a clinical trial of high-dose PPI treatment, and both are equivalent to catheter-based monitoring, the accuracy does not limit cost-effectiveness. However, where catheter-based pH monitoring is used as a perfect reference standard and the accuracy of wireless monitoring relative to this standard is considerably lower, wireless pH monitoring is dominated in all cost–utility analyses. Although catheter-based monitoring was the DAP-specified reference, the concerns with this approach are detailed in 'Discussion – Effectiveness'.

- The extent to which patients who do not receive pH monitoring to ascertain a refinement of their diagnosis would remain on high-dose PPIs, even in cases where they are not effective. If 100% of patients who do not obtain pH monitoring received active follow-up and review of their treatment and empirical investigations for alternative diagnoses, the economic analysis suggests that wireless monitoring would not be cost-effective; however, if even as few as 1 in 10 patients do not receive follow-up and inappropriately remain on high-dose PPI treatment, as assumed in the base-case, the option of wireless monitoring to assist in diagnosis becomes relatively cost-effective.

The proposed listing is for use of wireless pH monitoring only in patients that cannot tolerate or that fail catheter-based pH monitoring; however, the classification of patients as such may be subjective, and some patients may have a preference toward wireless monitoring. There is then a risk that, following a listing, some patients would receive wireless monitoring when they would otherwise have had a catheter-based pH test. This economic analysis suggests that this is not cost-effective, and were it to occur in any more than one in five patients who could otherwise undergo catheter-based monitoring, the overall cost-effectiveness of the proposed listing may be compromised. The extent to which leakage would occur in practice is highly uncertain.

Estimated financial impact

Should the listing proceed, it is estimated that, over the next few years, approximately 400–500 patients would be eligible for the procedure each year. This represents 10% of the number of patients who receive catheter-based monitoring, as this is the estimated catheter-intolerance rate. An annual growth rate in eligibility of 4%–8% is estimated, which is consistent with recent growth rates of catheter-based monitoring.

It is assumed that each episode represents a single patient; however, it is possible that some degree of repeat monitoring is undertaken such that the number of patients receiving monitoring in a year may be fewer than the number of tests, e.g. if on-medication and off-medication tests are required. The extent of this is not explored as it is not expected to change the overall financial impacts significantly.

If the estimate of intolerance to catheter-based monitoring is an underestimate, or if patients who could tolerate catheter-based monitoring 'leak' into the market for wireless monitoring, the number of patients could increase by approximately 10-fold. It is noted that some patients will have a preference for wireless monitoring (Hirano 2006; Mearin et al. 1998), and the proposed fee structure incentivises doctors to provide the wireless service

rather than the catheter-based one. Therefore, a degree of leakage is anticipated, although a realistic estimate of the extent to which this may occur is difficult to predict.

In addition to the cost directly associated with the MBS item listing, each wireless pH-test episode is associated with:

- anaesthesia (MBS items 17610, 20740, 23010) to the value of $161.80, plus an average of $5.94/patient (1 in 10 patients) to allow for additional anaesthesia costs associated with age and complication (MBS Items 25015, 23021)

- admission to a hospital/day-centre (AR-DRG Z40Z; average cost $607).

While the standard patient MBS co-contribution of 25% (for in-patients) of the MBS fees— the proposed listing and MBS anaesthetic fees—can be estimated, the extent to which doctors' bills will include an additional 'gap' charge payable by the patient is unknown. It is expected that hospital/day-facility fees will generally be paid for by private insurance companies but in some cases this may represent an additional out-of-pocket expense for the patient.

Where wireless monitoring is undertaken in patients who would not have tolerated catheter-based monitoring, there are no cost offsets to be considered. Where wireless monitoring leaks into the market for patients who would have otherwise undergone catheter-based monitoring, the cost of catheter-based monitoring (MBS item 11810; $174.75) is offset from the cost implications of the wireless test.

A summary of the overall financial impacts associated with the listing, in the scenarios of restricted use in the proposed population and allowing for a high level of leakage in patients who would have undergone catheter-based monitoring, are presented in Table 68 and Table 69.

Table 68   Financial impact on the MBS, patients and private health insurers, assuming use is strictly limited to patients who cannot tolerate catheter-based monitoring

	2014–15	2015–16	2016–17
Number of patients eligible for wireless test (restricted) (Table 61)	388	404	420
MBS cost of the proposed item number ($913.64/patient) [a]	$354,761	$368,952	$383,710
Total MBS costs associated with listing (no leakage)	$419,218	$435,987	$453,426
MBS benefits payable (75%)	$314,643	$327,228	$340,318
Patient co-payments	$104,575	$108,758	$113,109
Hospital costs associated with test ($607/patient) (Table 63)	$235,695	$245,122	$254,927
Total healthcare costs (no leakage)	$654,913	$681,109	$708,353

MBS = Medicare Benefits Schedule

Table 69   Financial impact on the MBS, patients and private health insurers, assuming extensive use of wireless monitoring in patients who would otherwise have a catheter-based pH test

	2014–15	2015–16	2016–17
Number of patients taking up wireless test including upper limit of plausible leakage (Table 60)	3,883	4,038	4,200
Total MBS costs associated with listing (with leakage)	$4,192,242	$4,359,586	$4,533,970
Cost-offset associated with replaced catheter-based test (Table 61)	–$666,846	–$693,367	–$721,102
Net change in total MBS costs (Table 62)	$3,525,396	$3,666,219	$3,812,868
Net change in MBS benefits	$2,605,412	$2,709,492	$2,817,872
Net change in patient contributions	$919,985	$956,726	$994,996
Hospital costs associated with test ($607/patient) (Table 64)	$2,356,981	$2,451,066	$2,549,109
Total healthcare costs with high level of leakage	$5,882,377	$6,117,285	$6,361,977

MBS = Medicare Benefits Schedule

While the leakage of wireless monitoring beyond the proposed population as presented in Table 69 is extreme and unlikely to occur to such an extent, it is clear that the level of leakage that may occur would have a very significant impact on the overall financial impacts of the proposed listing. While a strictly observed restriction would have a relatively low healthcare system cost (less than $1 million/year), even a modest amount of leakage is likely to increase this impact substantially.

Conclusions

Is catheter-free monitoring safe?

Adverse effects may occur from catheter-free monitoring as a result of the capsule being attached to the oesophageal wall. One study reported that chest pain was significantly more likely to occur in those undergoing catheter-free monitoring than in those undergoing no pH monitoring. Rare reported complications regarding the capsule include dislodgement and/or aspiration, retention somewhere in the gastrointestinal tract, and oesophageal perforation or tearing caused by capsule placement.

Twelve comparative studies reported complications of catheter-free monitoring compared with catheter-based monitoring. No deaths or life-threatening events were reported. In general, catheter-free monitoring causes less overall discomfort than catheter-based monitoring and has less impact on normal daily activities and work attendance. Nose and throat pain, dysphagia, eating and drinking difficulties, and headache were more prevalent in catheter-based monitoring. However, chest pain and foreign-body sensation were more likely to occur with catheter-free monitoring, compared with catheter-based monitoring.

In conclusion, catheter-free monitoring is usually better tolerated than catheter-based monitoring. Most reported complications are mild and do not require medical therapy but some (rare) complications could become severe when they are left untreated (e.g. mucosal tears).

Is catheter-free monitoring effective?

Direct evidence on the health impact of catheter-free monitoring versus no monitoring in the selected study population was not available. However, there was limited evidence that the health effects of catheter-free monitoring are similar to catheter-based monitoring: one study reported disappearance of the principal symptom in 73% of patients in the catheter-free monitoring group and 69% in the catheter-based monitoring group. On the other hand, catheter-free monitoring faced more technical problems compared with catheter-based monitoring according to a meta-analysis based on eight studies (RR=3.3, 95%CI 1.47 – 5.62). The most commonly reported technical failures that caused problems in catheter-free monitoring were early capsule detachments (i.e. in the wrong place in the oesophagus or where the capsule fails to attach to the oesophageal mucosa), causing incomplete data capture and errors in placement of the system.

Sensitivity and specificity could be calculated from eight studies—three used catheter-based monitoring as a reference standard, and in five the reference standard was clinical diagnosis. Sensitivity and specificity in the latter five studies varied from 67% to 100% and from 66% to 100%, respectively, versus around 86.4% and 77.8%, respectively, in the first three studies. However, these accuracy values are likely to be biased. First, as arbitrary cut-off points were used in these studies, the threshold for a positive/negative test result changes the characteristics of the index test. It has to be determined which cut-off value leads to optimal accuracy. Second, there is currently no consensus on how the cut-off values lead to a positive test result (i.e. worst day analysis, positive value on both days or reflux symptom correlations). Third, the optimal position of the capsule in the oesophagus should be determined to maximise accuracy results. Finally, the catheter-free monitoring period is 48 hours instead of only 24 hours, as with catheter-based monitoring. This increased monitoring period leads to an incremental diagnostic yield of 2.0%–15.6%, as reported in six studies. Catheter-based monitoring was used as a reference standard for diagnostic accuracy in the assessment; however, as there is an incremental diagnostic yield with catheter-free monitoring, the validity of this reference standard should be questioned. As a result of these uncertainties, no conclusion can be drawn regarding the accuracy of catheter-free monitoring compared with catheter-based monitoring. Consensus has first to be reached on a cut-off value, the use of a cut-off value, the position of the capsule and an appropriate reference standard.

There was no available evidence that met the population inclusion criteria on change in management; however, change occurred in around 60% of patients who could potentially tolerate catheter-based monitoring, and management changed in 88% of child patients after a pH test. Patients with an abnormal study result were more likely to get a change in management than patients with a normal result.

The evidence on whether a change in management leads to improved health outcomes was limited. Two systematic reviews and two RCTs on true positives were included to answer the question whether anti-reflux surgery for GORD leads to better health outcomes compared with medical treatment in patients with GORD (and a positive pH-test result). Anti-reflux surgery led to better health outcomes and had a higher decrease in symptoms, compared with medical management. Two RCTs published 5-year follow-up results; most patients achieved and remained in remission at 5 years with anti-reflux surgery, and fundoplication

continued to give better pain relief than medical management. However, these results cannot be generalised to patients who are refractory to PPI therapy, as these trials were conducted in patients who responded to PPI therapy. Furthermore, two studies on GORD misdiagnosis in small infants were included, to provide evidence on false positive results. A small percentage (7% and 17%) of infants with infantile spasms initially received a GORD diagnosis, which led to a delay in (true) diagnosis and treatment. Treatment delay was associated with a poor outcome 2 years after diagnosis and low treatment response. In conclusion, false GORD diagnosis could lead to a delay in correct treatment and suboptimal health outcomes. However, it is not known whether this patient population is eligible for oesophageal pH monitoring.

In summary, using catheter-free monitoring for GORD appears *likely* to be effective at benefiting the health of patients, compared with no monitoring; however, due to applicability concerns regarding the patient populations used in the studies, there is a large amount of uncertainty surrounding this conclusion.

Is catheter-free monitoring cost-effective?

Economic modelling suggests that catheter-free pH monitoring appears to be reasonably cost-effective in a patient population who have PPI-refractory, endoscopic negative GORD symptoms who cannot tolerate catheter-based pH monitoring, as it may result in improved diagnosis and management of symptoms, albeit at some additional cost—a base-case ICER of $14,457/QALY.

However, this conclusion only holds if the accuracy of catheter-free monitoring is assumed to be close to equivalent or better than catheter-based pH monitoring, and this assumption is uncertain given the lack of evidence in using an appropriate alternative reference standard.

Furthermore, the proposed listing may be used beyond the intended population in patients who would otherwise have had catheter-based monitoring; where this occurs, cost-effectiveness is substantially reduced.

Costing

The expected uptake of catheter-free pH monitoring in the intended catheter-intolerant population is estimated at approximately 400 tests annually. However, if monitoring occurs in a broader population, as many as 4,000 tests could be undertaken annually.

The total cost to the MBS for catheter-free pH monitoring in the intended catheter-intolerant population is estimated to be approximately $500,000 annually. This would increase to $3.6 million if the test is used in the broader population requiring pH monitoring.

The total cost to the Australian healthcare system including the MBS for catheter-free pH monitoring in the intended catheter-intolerant population is estimated to be $700,000 annually, increasing to over $6 million if the test is used in the broader population requiring pH monitoring.

Appendix A   Health Expert Standing Panel and Assessment Group

Health Expert Standing Panel (HESP)

Member	Expertise or affiliation
Prof. Robert Fraser	Gastroenterologist, Flinders University Department of Gastroenterology, Adelaide, Australia
Prof. Robert Batey	Gastroenterologist, Addiction Medicine, Central Clinical School, Royal Prince Alfred Hospital, the University of Sydney, Sydney, Australia

Assessment group

AHTA, University of Adelaide, South Australia

Name	Position
Ms Sharon Kessels	Research Officer
Ms Camille Schubert	Senior Health Economist
Ms Skye Newton	Team leader (Medical HTA)
Assoc. Prof. Tracy Merlin	Managing Director

Noted conflicts of interest

There were no conflicts of interest.

Appendix B   Search strategies

AUSTRALIA

Australian Safety and Efficacy Register of New Interventional Procedures – Surgical (ASERNIP-S)	http://www.surgeons.org/for-health-professionals/audits-and-surgical-research/asernip-s/
Centre for Clinical Effectiveness, Monash University	http://www.med.monash.edu.au/sphpm/divisions/mars/cce.html
Centre for Health Economics, Monash University	http://www.buseco.monash.edu.au/centres/che/

AUSTRIA

Institute of Technology Assessment / HTA unit	http://www.oeaw.ac.at/ita

CANADA

Agence d'Evaluation des Technologies et des Modes d'Intervention en Santé (AETMIS)	http://www.aetmis.gouv.qc.ca/site/home.phtml
Alberta Heritage Foundation for Medical Research (AHFMR)	http://www.ahfmr.ab.ca/publications.html
Alberta Institute of Health Economics	http://www.ihe.ca/
The Canadian Agency for Drugs And Technologies in Health (CADTH)	http://www.cadth.ca/index.php/en/
Canadian Health Economics Research Association (CHERA/ACRES) – Cabot database	
Centre for Health Economics and Policy Analysis (CHEPA), McMaster University	http://www.chepa.org
Centre for Health Services and Policy Research (CHSPR), University of British Columbia	http://www.chspr.ubc.ca
Health Utilities Index (HUI)	http://www.fhs.mcmaster.ca/hug/index.htm
Institute for Clinical and Evaluative Studies (ICES)	http://www.ices.on.ca
Saskatchewan Health Quality Council (Canada)	http://www.hqc.sk.ca

DENMARK

Danish Centre for Evaluation and Health Technology Assessment (DACEHTA)	http://www.sst.dk/english/dacehta.aspx?sc_lang=en
Danish Institute for Health Services Research (DSI)	http://dsi.dk/english/

FINLAND

Finnish Office for Health Technology Assessment (FINOHTA)	http://www.thl.fi/en_US/web/en

FRANCE

L'Agence Nationale d'Accréditation et d'Evaluation en Santé (ANAES)	http://www.anaes.fr/

GERMANY

German Institute for Medical Documentation and	http://www.dimdi.de/static/en/index.html

Information (DIMDI) / HTA

Institute for Quality and Efficiency in Health Care (IQWiG)   http://www.iqwig.de

THE NETHERLANDS

Health Council of the Netherlands Gezondheidsraad   http://www.gezondheidsraad.nl/en/

Institute for Medical Technology Assessment (Netherlands)   http://www.imta.nl/

NEW ZEALAND

New Zealand Health Technology Assessment (NZHTA)   http://www.otago.ac.nz/christchurch/research/nzhta/

NORWAY

Norwegian Knowledge Centre for the Health Services   http://www.kunnskapssenteret.no

SPAIN

Agencia de Evaluación de Tecnologias Sanitarias, Instituto de Salud "Carlos III"I/Health Technology Assessment Agency (AETS)   http://www.isciii.es/

Andalusian Agency for Health Technology Assessment (Spain)   http://www.juntadeandalucia.es/

Catalan Agency for Health Technology Assessment (CAHTA)   http://www.gencat.cat

SWEDEN

Center for Medical Health Technology Assessment   http://www.cmt.liu.se/?l=en&sc=true

Swedish Council on Technology Assessment in Health Care (SBU)   http://www.sbu.se/en/

SWITZERLAND

Swiss Network on Health Technology Assessment (SNHTA)   http://www.snhta.ch/

UNITED KINGDOM

National Health Service Health Technology Assessment (UK) / National Coordinating Centre for Health Technology Assessment (NCCHTA)   http://www.hta.ac.uk/

NHS Quality Improvement Scotland   http://www.nhshealthquality.org/

National Institute for Clinical Excellence (NICE)   http://www.nice.org.uk/

The European Information Network on New and Changing Health Technologies   http://www.euroscan.bham.ac.uk/

University of York NHS Centre for Reviews and Dissemination (NHS CRD)   http://www.york.ac.uk/inst/crd/

UNITED STATES

Agency for Healthcare Research and Quality (AHRQ)   http://www.ahrq.gov/clinic/techix.htm

Harvard School of Public Health   http://www.hsph.harvard.edu/

Institute for Clinical and Economic Review (ICER)	http://www.icer-review.org/
Institute for Clinical Systems Improvement (ICSI)	http://www.icsi.org
Minnesota Department of Health (US)	http://www.health.state.mn.us/htac/index.htm
National Information Centre of Health Services Research and Health Care Technology (US)	http://www.nlm.nih.gov/hsrph.html
Oregon Health Resources Commission (US)	http://www.oregon.gov/oha/OHPR/HRC/Pages/index.aspx
Office of Health Technology Assessment Archive (US)	http://fas.org/ota
U.S. Blue Cross/ Blue Shield Association Technology Evaluation Center (Tec)	http://www.bcbs.com/blueresources/tec/
Veteran's Affairs Research and Development Technology Assessment Program (US)	http://www.research.va.gov/default.cfm

Bibliographic databases

Electronic database	Time period
Cochrane Library – including, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, the Cochrane Central Register of Controlled Trials (CENTRAL), the Health Technology Assessment Database, the NHS Economic Evaluation Database	2001 – 9/2013
Current Contents	2001 – 9/2013
Embase	2001 – 9/2013
PubMed	2001 – 9/2013
Web of Science – Science Citation Index Expanded	2001 – 9/2013
Cinahl	2001 – 9/2013
Econlit	2001 – 9/2013
Scopus	2001 – 9/2013

Additional sources of literature

Source	Location
Internet	
NHMRC- National Health and Medical Research Council (Australia)	http://www.nhmrc.gov.au/
US Department of Health and Human Services (reports and publications)	http://www.hhs.gov/
New York Academy of Medicine Grey Literature Report	http://www.greylit.org/
Trip database	http://www.tripdatabase.com
Current Controlled Trials metaRegister	http://controlled-trials.com/
National Library of Medicine Health Services/Technology Assessment Text	http://text.nlm.nih.gov/
U.K. National Research Register	http://www.nihr.ac.uk/Pages/NRRArchive.aspx
Google Scholar	http://scholar.google.com/
Australian and New Zealand Clinical Trials Registry	www.anzctr.org.au
Pearling	
All included articles will have their reference lists searched for additional relevant source material	
Guidelines search (last step linked evidence)	
Guidelines International Network (G-I-N)	http://www.g-i-n.net/
NHMRC Clinical Guidelines Portal	http://www.clinicalguidelines.gov.au

Additional databases searched for economic evaluations

Electronic database
Cost-effectiveness Analysis (CEA) Registry
Database of Abstracts of Reviews of Effects or Reviews of Effects (DARE)
Health Technology Assessment database

NHS Economic Evaluation Database (NHS EED)
European Network of Health Economics Evaluation Databases (EURONHEED)

Specialty websites

International Foundation for Functional Gastrointestinal Disorders	http://www.aboutgerd.org/
American Gastroenterological Association	http://www.gastro.org/join-or-renew/join-aga/physician-scientists/international-physician-scientists
American college of Gastroenterology	http://gi.org/
Gastroenterological Society of Australia	http://www.gesa.org.au/

Appendix C   Studies included in the review

Study profiles of included studies on safety

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Safety outcomes assessed
Agrawal, Akerman & Rich (2009) Rhode Island Hospital, Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA	Case report Level: NA	1 patient 45 year old female Indication: Persistent heartburn despite maximal acid-suppressive therapy	Inclusion: NA Exclusion: NA	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the SCJ Duration: not stated	Mild epigastric pain Severe retrosternal chest discomfort
Ahlawat et al. (2006) Division of Gastroenterology, Department of Medicine, Georgetown University Hospital, Washington, DC, USA	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	90 patients 48 males and 42 females Age range: 18–80 years Indications: Being intolerant to the conventional naso-oesophageal pH catheter placement or had normal 24-hour oesophageal pH monitoring despite symptoms	Inclusion: Not stated Exclusion: Patients with bleeding diathesis, strictures, severe oesophagitis, varices, obstructions, pacemakers, implantable cardiac defibrillators, and surgical manipulations of the upper GI tract	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours Comparator: NA	Chest pain Foreign-body sensation
Andrews et al. (2012) Division of Gastroenterology, University of	Dual centre, randomised, non-blinded trial	86 patients (40 in Calgary and 46 in Edmonton) 61 females, 25 males Age range 18–75 years	Inclusion: Patients aged 18–75 years referred for ambulatory pH monitoring between August 2008 and August 2009	Index test: System: activated and calibrated Bravo™ wireless capsule system, Medtronic Inc., Minneapolis, MN	Manometry discomfort (nasal, throat, chest, overall) pH placement

Calgary, Canada	Level: II Quality: 19/26 CX P2	Indications: Reflux symptoms Chest pain Dysphagia	Exclusion: Patients with previous oesophageal surgery or achalasia	Insertion: transorally Placement: 5 cm above the proximal border of the LOS Duration: 24 (Calgary) or 48 (Edmonton) hours Comparator: Catheter: calibrated standard antimony-based pH catheter (Comfortec Plus catheter and ZepHR Sleuth recorder, Sandhill Scientific Inc., Highlands Ranch, CO) Insertion: transnasally Placement: 5 cm above the proximal border of the LOS Duration: 24 hours	discomfort (nasal, throat, chest, overall) pH-test discomfort (nasal, throat, chest, overall) Eating and drinking Ability to do usual daily activities
Azzam et al. (2012) Sao Paulo University Medical School, Sao Paulo, Brazil	Case series Level: IV Quality: Q1 CX (no comparator) P2	25 patients 21 females, 4 males Mean age: 52.4 years (range 34–73) Indications: All patients had as predominant symptom the typical GORD complaints; 64% also had atypical complaints and 76% had associated extra oesophageal complaints	Inclusion: Heartburn and/or regurgitation as the main clinical complaint; at least 18 years of age; recent upper GI endoscopy (within the past 2 months); interruption in the administration of PPIs for 7 days; and signature on the free and informed consent form Exclusion: Oesophageal diverticula; strictures and varices; hiatal hernia greater than or equal to 3 cm; erosive oesophagitis Los Angeles grades C or D; Barrett's oesophagus; neoplasms; obstructive diseases; previous surgery of the GI tract	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 3 cm above the superior border of the LOS Duration: 48 hours Comparator: NA	Chest pain (none)
Belafsky et al. (2004)	Prospective case series	46 patients 23 males and 23 females	Inclusion: Not stated	Index test: System: Bravo™ wireless capsule system	Self-limited epistaxis Laryngospasm

Scripps Center for Voice and Swallowing, La Jolla, California, USA	Level: IV Quality: Q2 CX (no comparator) P2	Mean age: 52 years Indications: GORD, chronic cough, laryngopharyngeal reflux	Exclusion: Not stated	Insertion: transnasally Placement: 5 cm above the manometric lower oesophageal high-pressure zone or 6 cm above the endoscopic SCJ Duration: >36 hours Comparator: NA	Vasovagal reaction
Bhat, McGrath & Bielefeldt (2006) Department of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA	Prospective case series Level: IV Quality: Q2 CX (no comparator) P2	217 patients 140 females, 77 males Mean age: 51 years (range 42–58) Indications: Preoperative workup for fundoplication, possible extra oesophageal manifestations of reflux disease, symptoms after fundoplication, refractory symptoms on medication, atypical symptoms	Inclusion: Adult patients undergoing endoscopy with wireless pH studies at the University of Pittsburgh Medical Center between 9 April 2004 and 31 March 2005 Exclusion: Individuals unable to give informed consent, pregnant women, prisoners, patients with oesophageal varices, patients requiring continuing anticoagulation, patients with cardiac defibrillators	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Chest discomfort Rash
Bothwell, Phillips & Bauer (2004) Department of Otolaryngology, University of Missouri-Columbia, Columbia, Missouri, USA	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	25 patients Mean age: 3 years (range 3 months – 11 years) Indications: Vomiting and dysphagia, 'noisy breathing', chronic cough, aspiration pneumonia, reactive airway disease, recurrent croup, known subglottic stenosis, nasal airway obstruction with rhinosinusitis, sleep apnoea, restrictive lung disease	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: just below the closed cricopharyngeal muscle in the upper oesophagus Duration: 48 hours Comparator: NA	Mucosal tear
Bradley et al. (2011)	Cohort study	341 patients (248 patients underwent Bravo	Inclusion: Outpatients undergoing ambulatory pH-	Index test: System: Bravo™ wireless capsule system	Chest discomfort Foreign-body sensation

Division of Gastroenterology and Hepatology, Mayo Clinic, Scottsdale, Arizona, USA	Level: III-2 Quality: 15/26 CX P2	monitoring) 62% females in the wireless group, 72% females in the catheter group Mean age: 58 years in wireless group, 57 years in catheter group	metry between October 2006 and October 2008 Exclusion: Known history of oesophageal strictures, obstructions or surgical manipulations of the upper GI tract (e.g. anti-reflux surgery or oesophagectomy)	Insertion: transorally Placement: 6 cm above the SCJ or 5 cm above the upper border of the LOS Duration: 48 hours Comparator (MII-pH): System: Sandhill Scientific Inc., Highland Ranch, USA Insertion: transnasally Placement: proximal channel 5 cm above upper border of the LOS Duration: 24 hours Comparator (pH-metry): System: Alpine Biomed, Corp, Fountain Valley, USA Insertion: transnasally Placement: 5 cm above the upper border of the LOS Duration: 24 hours	Chest and abdominal pain Difficulty swallowing Painful swallowing Eating and drinking Nausea and vomiting
Cabrera et al. (2011) Division of Pediatric Gastroenterology, Hepatology and Nutrition, Indiana University School of Medicine, Indianapolis, Indiana, USA	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	289 patients (At least 1 day of data was obtained in 278 patients; 2 days were obtained in 274 patients) 144 females, 145 males Age: range 4–22 years Indications: Epigastric abdominal pain (59.9%), vomiting (33.9%), regurgitation (23.9%), heartburn (18.3%)	Inclusion: Children who are at least 4 years of age and weigh at least 30 lb Exclusion: Children with a history of oesophageal surgery, a documented anatomic abnormality, or a history of coagulopathy or a bleeding diathesis	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: at 87% of the distance from the incisors to the Z line Duration: 48 hours Comparator: NA	Oesophageal mucosa tear
Calabrese et al.	Case series	24 patients (12 males, 12 females)	Inclusion:	Index test:	Occasional retrosternal

(2008) Department of Clinical Medicine, University of Bologna, Bologna, Italy	Level: IV Quality: Q2 CX (no comparator) P3	Mean age: 37.7 ± 11.9 years Indications: Typical symptoms of GORD (heartburn and/or regurgitation)	The presence of clinical symptoms such as heartburn and/or regurgitation without erosive oesophagitis, symptoms occurred more than twice weekly for at least 6 months, treatment with acid suppression had results in complete symptom relief or in reduction of the symptoms by more than 50% Exclusion: Coagulopathy, severe oeosphagitis/strictures, erosive oesophagitis, portal hypertension, a history of upper GI surgery, *Helicobacter pylori* infection, age <20 or >70 years, pacemaker	System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 96 hours Comparator: NA	discomfort on swallowing
Croffie et al. (2007) Section of Pediatric Gastroenterology, Hepatology and Nutrition, Indiana University School of Medicine, Indianapolis, Indiana, USA	Randomised controlled trial Level: II Quality:14.5/26 CX P2	66 patients (children) (but 5 patients were excluded due to operative failure, failed recording etc.) 32 males, 34 females Mean age: 9.4 years (range 4–16.5) Indications: Persistent epigastric or substernal pain, persistent vomiting, heartburn, chronic nocturnal cough or wheezing, persistent throat clearance, dental abnormalities suspected to be caused by reflux	Inclusion: Children between the ages of 4 and 18 years and weighing at least 30 lb who were undergoing oesophagogastroduodenoscopy and oesophageal pH monitoring for symptoms suggestive of gastro-oesophageal reflux. Exclusion: Children younger than 4 years of age or weighing <30 lb, having anatomic abnormalities of the oesophagus, a history of surgical procedure of the oesophagus, stomach or duodenum, a history of coagulopathy or bleeding diathesis, presence of oesophageal varices, significant medical illness	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: at 87% of the distance from the incisors to the Z line Duration: 48 hours Comparator: System: not stated (catheter-based) Insertion: transnasally Placement: tip at 87% of the distance from the nostril to the GOJ Duration: 24 hours	Vomiting Chest pain Throat pain Oesophageal discomfort Eating and drinking Activity Satisfaction
Crowell et al. (2009) Division of	Case series Level: IV	157 patients, predominantly (>95%) Caucasian Indications: heartburn/regurgitation	Inclusion: Patients with GORD symptoms Exclusion:	Index test: System: Bravo™ wireless capsule system Insertion: transorally	Foreign-body sensation Chest discomfort Chest pain

Gastroenterology and Hepatology, Mayo Clinic, Scottsdale, Arizona, USA	Quality: Q1 CX (no comparator) P2	symptoms, chest pain, abdominal pain/discomfort	History of bleeding diathesis, strictures, severe oesophagitis, varices, obstructions, pacemakers or implantable cardiac defibrillators	Placement: 6 cm above the endoscopically determined GOJ or 5 cm above the manometrically determined upper margin of the LOS Duration: 48 hours Comparator: NA	Diminished appetite
de Hoyos & Esparza (2010); de Hoyos, Esparza & Loredo (2009) Department of Gastroenterology, Angeles del Pedregal Hospital, Mexico City, Mexico	Prospective case series Level: IV Quality: Q2 CX (no comparator) P2	66 patients 27 males, 39 females Mean age: 41.7 years (range 11–73) Indications: Typical oesophageal reflux symptoms, atypical manifestations or a mixed symptomatology	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm from the SCJ Duration: 48 hours Comparator: NA	Intolerable chest pain Detachment failure
des Varannes et al. (2005) Department of Gastroenterology and Hepatology, Hopital Hotel Dieu, Nantes, France	Case series Level: IV Quality: Q1 CX (no comparison) P2	40 patients (probe successfully attached in 36 patients) 21 males, 19 females Mean age: 50 years Indications: Heartburn (n=7), regurgitation (n=6) or both symptoms (n=26)	Inclusion: Patients suggestive of GORD and referred to the functional laboratory of four French academic centres for pH monitoring Exclusion: Severe oesophageal motility disorders, severe oesophagitis (Los Angeles grade C or above), pregnancy, women who were not using reliable contraception	Index test: System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 5 cm above the Z line Duration: 48 hours Comparator: NA	Epistaxis Dizzy spells Sleep disorders Dysphagia for liquids Swallowing discomfort Thoracic discomfort
Domingues, Moraes-Filho & Domingues (2011) Motility laboratory (Gastro	Retrospective case series Level: IV Quality: Q2	100 patients (probe successfully attached in 95 patients) 39 female, 61 male Mean age: 53 years (range 18– 81)	Inclusion: Patients referred to a GI motility laboratory to investigate GORD with typical or atypical symptoms, persistent GORD symptoms on medical therapy or recurrent GORD symptoms after surgical	Index test: System: Bravo™ wireless capsule system Insertion: 5 cm above the upper border of the LOS (manometry) or 6 cm above the LOS (endoscopy)	Foreign-body sensation Chest pain

Resolucao Diagnostico, Laboratorio de Motilidade Digestive), Rio de Janeiro, Brazil	CX (no comparator) P2	Indications: 82% were referred to diagnose GORD, 18% were referred due to persistent symptoms despite PPI use	fundoplication, between 2004 and 2009 Exclusion: <18 years of age, history of bleeding tendency or coagulopathy, significant concomitant medical co-morbidities, severe GI bleeding in the past 6 months, a history of upper GI surgery, medication with PPIs, presence of oesophageal varices, Barrett's oesophagus, oesophageal stenosis, erosive oesophagitis, pacemaker or implantable cardiac defibrillator in situ	Placement: transorally Duration: 48 hours Comparator: NA	
Fajardo et al. (2006) Division of Gastroenterology and Hepatology, Mayo Clinic College of Medicine, Rochester, Minnesota, USA	Case report Level: NA	1 patient 76 year old woman Indication: Recurrent GORD, failing to respond to medical therapy	Inclusion: NA Exclusion: NA	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: not stated Duration: NA Comparator: NA	Mucosal tear with oesophageal perforation
Francis, DL (2008) Miles and Shirley Fiterman Center for Digestive Diseases, Mayo Clinic, Rochester, Minnesota, USA	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	76 patients	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Severe chest pain Capsule dislodgement in the pyriform sinus
Francis, DO et al. (2012)	Randomised, blinded controlled	22 patients were included (13 males, 9 females)	Inclusion: Patients >18 years of age, scheduled for	Index test: System: Bravo™ wireless capsule system	Chest pain (adj OR) Odynophagia (adj OR)

Division of Gastroenterology, Hepatology and Nutrition, Vanderbilt University Medical Center, Nashville, Tennessee, USA	trial Level: II Quality: 16/26 C1 P2	Indications: Cough, hoarseness, globus, heartburn, regurgitation	planned oesophagogastroduodenoscopy and wireless pH monitoring for physiologic assessment of oesophageal acid exposure for typical GORD symptoms or for extraoesophageal reflux (EER) symptoms Exclusion: History of upper oesophageal surgery, bleeding dyscrasia, recent cerebrovascular accident or transient ischemic attack, recent (<6 months) GI haemorrhage, oesophageal varices or significant medical illness, current pregnancy	Insertion: transorally Placement: 16 cm above the SCJ Duration: 48 hours Comparator: No monitoring (a sham capsule placement)	Dysphagia (adj OR)
Garrean et al. (2008) Division of Gastroenterology, Northwestern University Freiberg School of Medicine, Chicago, Illinois, USA	Case series (prospective?) Level: IV Quality: Q2 CX (no comparator) P2	60 patients (22 males and 38 females) Age: 19–80 years 40 patients had complete 4-day data acquisition Indications: acid reflux symptoms, chest pain, laryngeal symptoms that were poorly responsive to twice-daily PPI therapy for a minimum of 6 weeks	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the SCJ Comparator: NA	Back pain Globus sensation Chest pain
Gillies et al. (2007) Departments of Upper Gastrointestinal Surgery and Gastrointestinal Physiology, Royal Berkshire hospital,	Non-randomised controlled trial Level: III-2 Quality: 13.5/26 CX P2	185 patients, but 200 oesophageal pH studies 104 males, 81 females Mean age: Catheter group: 46 years (range 16–71) Bravo group: 45 years (range 13–75) Indication:	Inclusion: See indication Exclusion: No specific exclusion criteria exist	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours Comparator: System: Medtronic	Overall discomfort (during insertion and monitoring) Nasal discomfort Throat discomfort Oesophageal discomfort Interference with daily activities Eating

Reading, UK		Symptoms suggestive of GORD or documenting the outcome of laparoscopic anti-reflux surgery		Insertion: transnasally Placement: 5 cm above the proximal border of the LOS Duration: 24 hours	Sleeping Work attendance
Grigolon et al. (2007) Gastroenterology, Department of Medical Sciences, University of Milan, Milan, Italy	Pseudo-randomised controlled trial Level: III-1 Quality: 13.5/26 CX P2	133 patients Catheter group: 78 patients, 46% males, aged 53 ± 2 years Bravo group: 55 patients, 45% males, aged 44 ± 3 years) <u>Indication:</u> Suspected GORD, consecutively referred for oesophageal pH monitoring	<u>Inclusion:</u> See indication <u>Exclusion:</u> Previous GI surgery (excluding cholecystectomy and appendectomy) and known or suspected stricture of the GI tract	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours <u>Comparator:</u> System: Slimline mod 901P3001, Medtronic, Denmark Insertion: transnasally Placement: 6 cm above the point of sudden change in pH from acid to neutral as detected by means of continuous rapid withdrawal from the stomach, or 5 cm above the upper margin of the LOS Duration: 24 hours	Discomfort at placement Overall discomfort Chest discomfort Eating and drinking Interference with daily activities
Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008) Department of Paediatric Surgery, Lund University Hospital, Lund, Sweden	Case series Level: IV Quality: Q1 CX (no comparator) P2	58 children 37 males, 21 females Mean age: 8 ± 4 years <u>Indications:</u> Vomiting, abdominal and chest pain, GORD control, GI bleeding, respiratory symptoms	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Oesophageal strictures, coagulopathy, any suspicion of intestinal strictures	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: NA Placement: above the diaphragm valve at the width of two of the individual patient's vertebral bodies as visualised on a perioperative X-ray Duration: 24 or 48 hours (23 children had 48-hour measurement) <u>Comparator:</u> NA	Chest discomfort Coughing Dysphagia

Hakanson et al. (2009) Department of Surgery, Center of Gastrointestinal Disease, Ersta Hospital, Stockholm, Sweden	Case-control study with in-subject simultaneous recording Level: III-2 Quality: 14/26 CX P2	53 volunteers and 55 patients (45 volunteers and 47 patients had enough data to be included) 30 females and 15 males in volunteer group 27 females and 20 males in patient group Mean age: Volunteers: 47 years (range 21–68) Patients: 50 years (range 23–69) Indication: Patients had symptoms including mainly heartburn, acid regurgitation, and chest or epigastric pain	Inclusion: Patients: referred to the clinic for investigation of GORD Exclusion: Volunteers: history of GORD or any other pathology of the upper GI tract, taking acid-suppressing agents or medication known to affect GI motility	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline, Medtronic, Skovlunde, Denmark Insertion: transnasally Placement: 5 cm proximal to the upper border of the LOS Duration: 24 hours	Complications (none) Premature removal of the capsule (none)
Hirano et al. (2005) Division of Gastroenterology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA	Case series (prospective) Level: IV Quality: Q2 CX (no comparator) P2	18 patients (17 patients had enough data to be included) 6 males, 12 females Median age: 52 years Indications: Chest pain, refractory heartburn, laryngeal symptoms	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the endoscopically determined SCJ Duration: 4 days Comparator: NA	Chest pain (none) Dysphagia (none) Bleeding (none)
Hochman and Favaloro-Sabatier (2005) Children's Center for Digestive Health Care, Atlanta, Georgia, USA	Retrospective case series Level: IV Quality: Q2 CX (no comparator)	44 patients (children) 27 males, 17 females Mean age: 11.8 years (range 6–19)	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the proximal border of the LOS Duration: 48 hours Comparator: NA	Patient discomfort

	P2				
Hogan et al. (2009) Gastrointestinal Associates, Jackson, Mississippi, USA	Case report (letter to the editor) Level: NA	1 patient 72 year old male Indication: Symptoms of gastro-oesophageal reflux	Inclusion: NA Exclusion: NA	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: not stated Duration: not stated	2-year retention of the capsule in a giant colonic diverticulum
Karamanolis et al. (2012) 2nd Department of Internal Medicine, Propaedeutic Attikon University General Hospital, Athens, Greece	Case series (prospective?) Level: IV Quality: Q1 CX (no comparator) P2	32 patients 18 males, 14 females Mean age: 45.3 ± 12.5 years Indications: Three episodes of non-cardiac chest pain per week	Inclusion: Patients with at least three episodes of chest pain per week and normal oesophageal manometry (in order to exclude the presence of any oesophageal motility disorder) Exclusion: Patients that were using aspirin or NSAIDs, a history of upper GI surgery, gastric or duodenal ulcer, connective tissue disease, or severe liver, lung, renal or haematological disease	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours Comparator: NA	Throat pain Back pain Globus sensation
Kramer & Chokhavatia (2012) Department of Medicine, the Mount Sinai Medical Center, New York, New York, USA	Case report Level: NA	1 patient 44 year old male Indication: History of GORD symptoms, non-adherent to anti-reflux therapy	Inclusion: NA Exclusion: NA	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the GOJ Comparator: NA	Capsule dislodging into the pyriform sinus
Lacy, Chehade & Crowell (2011) Section of Gastroenterology	Case series Level: IV	358 patients (257 females, 101 males) Mean age: 51 ±14 years Predominantly (97%) Caucasian	Inclusion: Not stated Exclusion:	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the GOJ or 5 cm	Chest discomfort

and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA	Quality: Q1 CX (no comparator) P2	Indications: Symptoms thought secondary to GORD	Not stated	above the upper border of the LOS Duration: 48 hours Comparator: NA	
Lee et al. (2005) Department of Internal Medicine, Department of Emergency Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan	Prospective case series Level: IV Quality: Q3 CX (no comparator) P2	40 patients (27 males and 13 females) Age: range 28–83 years Indications: Typical heartburn with unsatisfactory response to therapy (15), erosive oesophagitis and undergoing preoperative evaluation (25)	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Retrosternal discomfort Mucosal lesion Foreign-body sensation
Marchese et al. (2006) Digestive Endoscopy Unit, Agostino Gemelli University Hospital, Catholic University of Rome, Rome, Italy	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	40 patients (36 patients had successful oesophageal placement of the capsule) 23 males, 17 females Mean age: 42.5 years Indications: Reflux symptoms that were refractory to PPI therapy	Inclusion: Patients with reflux symptoms refractory to PPI therapy who had normal or equivocal endoscopic findings Exclusion: Patients with oesophageal varices, previous GI surgery, known coagulopathy or bleeding diathesis, cardiac pacemakers, defibrillators or other electromedical devices, significant comorbidity, pregnant women and patients <18 years of age	Index test: System: Bravo™ wireless capsule system Insertion: transnasally Placement: 5 cm above the upper margin of the LOS Duration: 24 or 48 hours Comparator: NA	Mild discomfort Severe chest pain (Endoscopic removal of the capsule) Epistaxis Pharyngeal irritation
Nusrat, Roy & Bielefeldt (2012) University of Pittsburgh Medical Center, Pittsburgh,	Retrospective case series Level: IV Quality: Q1	356 patients (65% females) Mean age: 49.9 ± 0.8 years Indications: Evaluations of GORD, recurrent or persistent symptoms after	Inclusion: Not stated Exclusion: Patients with pacemakers, implanted cardiac defibrillators, known oesophageal	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the endoscopically localised SCJ or 5 cm above the	No significant adverse effects during the study

Pennsylvania, USA	CX (no comparator) P2	fundoplication, chest pain, unexplained cough or other atypical symptoms	varices, bleeding disorders or ongoing anticoagulation	manometrically determined upper border of the LOS Duration: 48 hours Comparator: NA	
Pandolfino et al. (2003) Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA	Cohort study Level: III-2 Quality: 11.5/26 CX P3	44 controls (13 males, 31 females, aged 23–53 years) 41 patients (26 males, 15 females, aged 32–72 years)	Inclusion: Not stated Exclusion: Controls: abdominal symptoms, use of antacids or antisecretory medication, abnormal endoscopy Patients: history of surgical manipulation of the upper GI tract, history of bleeding diathesis or coagulopathy, stroke or transient ischemic attack in the past 6 months, significant medical illness, oesophageal varices	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Comparator: System: not stated (catheter-based) Insertion: transnasally Placement: not stated Duration: 24 hours	Foreign-body sensation Chest pain Throat discomfort Late detachment Oesophageal discomfort Eating and drinking Activity Sleep patterns
Pandolfino et al. (2006) Division of Gastroenterology, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA	Case series Level: IV Quality: Q2 CX (no comparator) P3	10 normal control subjects (7 males and 3 females, aged 21–53 years) 10 patients with GORD (6 males and 4 females, aged 20–54 years)	Inclusion: Controls: no history of heartburn, reflux, chest pain or atypical symptoms Patients: oesophagitis (3) or an abnormal 24-hour oesophageal pH study prior to this study Exclusion: History of surgical manipulation of the upper GI tract, anticoagulation therapy, significant medical disease Controls: abdominal symptoms, use of antacids, abnormal upper endoscopy	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 1 cm proximal to SCJ Duration: 24 hours Comparator: NA	Chest discomfort Foreign-body sensation

Prakash et al. (2006) Division of Gastroenterology, Washington University School of Medicine, St Louis, Missouri, USA	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	8 patients (who required endoscopic dislodgement because of severe discomfort) out of 452 patients 6 females, 2 males Mean age: 44.5 years (range 29–67)	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: not stated Duration: not stated Comparator: NA	Severe discomfort Severe chest pain and/or odynophagia Foreign-body sensation
Remes-Troche et al. (2005) Department of Gastroenterology, Instituto Nacional de Ciencias Medicas y Nutricion, Salvador Zubiran, Mexico City, Mexico	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	84 patients (77 patients had enough data to be included) 49 females, 35 males Mean age: 44 ± 12 years (range 19–73) Indications: Persistent GORD symptoms on PPI, preoperative evaluation before surgery, previous failed transnasal pH monitoring, extraoesophageal GORD	Inclusion: Consecutive patients with GORD symptoms with an indication for 24-hour pH monitoring Exclusion: Patients with surgical manipulation of the upper GI tract, bleeding diathesis or coagulopathy, severe GI bleeding in the past 6 months, oesophageal varices, significant comorbidities	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Chest pain Foreign-body sensation Nausea
Renteln et al. (2008) Department of Gastroenterology, Hepatology and Oncology, Klinikum Ludwigsburg, Ludwigsburg, Germany	Case report Level: NA	One 44 year old patient Indication: Evaluation of atypical GORD symptoms	Inclusion: NA Exclusion: NA	Index test: System: Bravo™ wireless capsule system Insertion: transnasally Placement: 5 cm above the upper margin of the LOS Duration: NA Comparator: NA	Retch Heavy cough Dropped oxygen saturation Capsule aspiration

Scarpulla et al. (2007) Gastroenterology Division, M. Raimondi Hospital, Can Cataldo, Italy	Retrospective case series Level: IV Quality: Q2 CX P2	83 patients (complete 4-day recordings were available for 34/83) 44 females, 39 males Median age: 42 years (range 18–63) Indications: Typical reflux symptoms or atypical symptoms including chest pain and extra oesophageal complaints	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the Z line Duration: 24, 48, 72 and/or 96 hours Comparator: NA	Retrosternal discomfort
Schneider et al. (2007) Department of General, Visceral, and Transplant Surgery, University Hospital of Tubingen, Tubingen, Germany	Non-randomised controlled trial Level: III-2 Quality: 9/26 CX P3	123 patients Bravo group: 36 males and 32 females, mean age: 51 years (range 34–67) Catheter group: 29 males and 26 females, mean age 43 years (range 32–59)	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 5 cm above the upper border of the LOS Duration: 48 hours Comparator: System: not stated Insertion: transnasally Placement: not stated Duration: 24 hours	Throat discomfort Oesophageal discomfort Limitation in daily activities
Shahid & Fisher (2011) Gastroenterology, Section, Temple University Hospital, Philadelphia, Pennsylvania, USA	Case report	1 patient 61 year old female Indication: History of heartburn, regurgitation and intermittent chest pain, unresponsive to PPI therapy	Inclusion: NA Exclusion: NA	Index test System: Bravo™ wireless capsule system Insertion: transorally Placement: not stated Comparator: NA	Capsule dislodgement in the left mainstem bronchus
Sofi et al. (2011)	Retrospective	58 patients	Inclusion:	Index test:	Severe chest pain Throat pain

Department of Gastroenterology, University of Toledo Medical Center, Toledo, Ohio, USA	case series	Age range: 26– 87 years 17 males, 41 females	Not stated Exclusion: Not stated	System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the GOJ Comparator: NA	Dysphagia Cough Work attendance
Sweis et al. (2009) Functional GI Disease Unit, Clinic for Gastroenterology and Hepatology, University Hospital of Zurich, Zurich, Switzerland	Cohort study Level: III-2 Quality: 15/26 CX P1	134 patients Mean age: 60 years (range 18–76) 58 males, 76 females Compared with 110 consecutive catheter pH controls Mean age: 57 years (range 16–85) 44 males, 66 females Indications: Intolerance of catheter insertion, intolerance of catheter after intubation, or vomiting of catheter within 24 hours of the monitoring period	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours Comparator: System: Slimline, Medtronic Insertion: transnasally Placement: 5 cm above the lower oesophageal high-pressure zone Duration: 24 hours	Overall satisfaction
Sweis et al. (2011) Guy's and St Thomas Hospitals, London, UK	Case series Level: IV Quality: Q2 CX P2	38 patients 13 males, 25 females Median age: 41.6 years (range 17–75) Indication: Reflux symptoms and no diagnosis of GORD on catheter-based pH monitoring	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 96 hours Comparator: NA	No patient complained of severe symptoms
Tankurt et al. (2011) Kent Hospital, Department of	Case series Level: IV	64 patients (37 males, 27 females) Mean age: 37.9 years (range 25–58) Indications:	Inclusion: Not stated Exclusion:	Index test: System: Bravo™ wireless capsule system Insertion: NS	Retrosternal pain or discomfort

Gastroenterology, Izmir, Turkey	Quality: Q2 CX (no comparator) P2	Preoperative evaluation (17), evaluation of atypical reflux symptoms (40) and 'other' (7)	Not stated	Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	
Tharavej et al.[10] (2006) Division of Thoracic and Foregut Surgery, Department of Surgery, University of Southern California, Los Angeles, California, USA	Case series Level: IV Quality:Q1 CX (no comparator) P2	80 patients (40 in control group and 40 in study group) Study group: 15 males and 25 females; mean age 45.5 years (range 40–53) Control group: 15 males and 25 females; mean age 49 years (range 44–63)	Inclusion: Consecutive patients with gastro-oesophageal reflux symptoms Exclusion: Patients with a history of achalasia, benign or malignant stricture, or a previous oesophageal or gastric surgery	Index test: System: Bravo™ wireless capsule system Insertion: transorally (study group) or transnasally (control group) Placement: 6 cm above the upper limit of the rugal folds (study group) or 5 cm above the upper border of the LOS Duration: 48 hours Comparator: NA	High distal oesophageal contraction amplitude Chest pain Hypertensive oesophageal contractions
Triester et al. (2005) Division of Gastroenterology and Hepatology, Division of Hospital Internal Medicine, Mayo Clinic, Scottsdale, Arizona, USA	Case report Level: NA	1 patient 61 year old woman with a history of GORD and a laparoscopic Nissen fundoplication Indication: Returning heartburn, regurgitation and sore throat, and intermittent dysphagia	Inclusion: NA Exclusion: NA	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the GOJ Duration: Comparator: NA	Foreign-body sensation Dysphagia Sharp substernal chest pain
Tseng et al.	Retrospective	190 procedures and 186 patients	Inclusion:	Index test:	Severe gagging

[10] This is a matched pairs comparative study, but we extracted data as if it was a case series because the comparison they make in the study is not relevant for this systematic review.

(2005) Digestive Health Center, Department of Surgery and Division of Gastroenterology, Oregon Health and Science University, Portland, Oregon, USA	case series Level: IV Quality: Q1 CX (no comparator) P2	(some patients had the test performed twice) 72 males and 114 females Mean age: 51 years Indications: Being evaluated for GORD	Patients older than 18 years of age Exclusion: Not stated	System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 5– 6 cm above the LOS or 5–6 cm above the SCJ Duration: 48 hours Comparator: 24-hour Bravo monitoring	during placement Foreign-body sensation Chest pain
Tu et al. (2004) Department of Internal Medicine and Department of Emergency Medicine, National Taiwan University Hospital, Taipei, Taiwan	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	30 patients 20 males, 10 females Mean age: 57.6 ± 14.3 years (range 29–83) Indications: 24 patients with GORD were evaluated for drug effectiveness before surgery, 6 patients showed typical heartburn but their endoscopic findings were normal.	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the GOJ Duration: 48 hours Comparator: NA	Mucosal abrasion and minor haemorrhage
Turner et al. (2007) Division of Gastroenterology, Brigham and Women's Hospital, Boston, Massachusetts, USA	Case series Level: IV Quality: Q2 CX (no comparator) P2	198 pH studies (148 studies had enough data to be included) Off medical therapy: 115 patients, 76 females, 39 males; mean age: 50 ± 13 years On medical therapy: 33 patients, 26 females, 7 males; mean age: 52 ± 14 years Indications: Typical and atypical symptoms of GORD or persistent GORD-related	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: (during endoscopy) Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Chest pain requiring capsule removal

		symptoms despite medical therapy. Patients with planned anti-reflux surgery were also evaluated			
Wenner, Johnsson et al. (2007) Department of Surgery, Lund University Hospital, Lund, Sweden	Randomised crossover trial Level: II Quality: 20/26 CX P2	35 patients (31 patients had enough data to be included) 15 males, 16 females Median age: 52 years Median BMI: 27 (range 24–28) Indications: Typical reflux symptoms (55%) and atypical reflux symptoms (45%)	Inclusion: Patients referred for oesophageal pH monitoring Exclusion: Patients with coagulopathy, severe oesophagitis/stricture, portal hypertension, a pacemaker, age <20 or >70 years, previous gastro-oesophageal surgery, severe cardiopulmonary disease, symptoms suggestive of oesophageal motor disorders	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline (Medtronic) Insertion: transnasally Placement: 5 cm above the upper border of the LOS Duration: 24 hours	Nose and/or throat symptoms Swallow-induced symptoms Chest symptoms All adverse symptoms Normal work activities Physical activities Eating and drinking Normal daily life
Wong et al. (2005) The Neuro-Enteric Clinical Research Group, Section of Gastroenterology, Department of Medicine, Southern Arizona VA Health Care System and University of Arizona Health Sciences Center, Tucson, Arizona, USA	Randomised controlled trial Level: II Quality: 16.5/26 CX P2	50 patients Mean age: 50.2 years (range 21–79) 26 males and 24 females Indication: Failure in controlling symptoms while on PPI therapy or for evaluation before anti-reflux surgery	Inclusion: Not stated Exclusion: Patients with a history of bleeding tendency or coagulopathy, significant concomitant medical co-morbidity, severe GI bleeding within the past 6 months, a history of upper GI surgery, oesophageal varices, pacemaker, implantable cardiac defibrillator, being unable to complete the 24-hour pH monitoring, unable to report daily activity	Index test: System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 5 cm above the upper border of the LOS Duration: 24 hours Comparator: System: Digitrapper Mark III (Medtronic) Insertion: transnasally Placement: 5 cm above the proximal margin of the LOS Duration: 24 hours	Nose pain Runny nose Throat pain Throat discomfort Cough Chest discomfort Headache Nose bleeding Work Eating and drinking Daily activities Overall experience with the test Satisfaction with the test

Study profiles of included studies on direct effectiveness

Direct evidence

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Health outcomes assessed
Grigolon et al. (2011) Department of Gastroenterology, Universita Degli Studi and Fondazione IRCCS Ca Granada, Ospedale Maggiore Policlinico, Milan, Italy	Matched-pairs retrospective cohort study Level: III-2 Quality: 16.5/26 CX P2	102 patients Catheter group: 51 patients, 20 males, mean age 48 years (range 22–78) Bravo group: 51 patients, 22 males, mean age 48 years (range 23–82)	Inclusion: All consecutive patients who underwent wireless 96-hour pH monitoring between January and December 2007 off PPI therapy, and patients undergoing traditional 24-hour pH monitoring in the same period, after matching for the most troublesome symptom potentially related to GORD, sex and age Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48–96 hours Comparator: System: Slimline mod 901P3001, Medtronic, Denmark Insertion: transnasally Placement: 6 cm above the point of sudden change in pH from acid to neutral as detected by means of continuous rapid withdrawal from the stomach or 5 cm above the upper margin of the LOS Duration: 24 hours	Outcome of the principal symptom
Sweis et al. (2011) Guy's and St Thomas Hospitals, London, UK	Case series Level: IV Quality: Q2 CX P2	38 patients 13 males, 25 females Median age: 41.6 years (range 17–75) Indication: Reflux symptoms and no diagnosis of GORD on catheter-based pH monitoring	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 96 hours Comparator: (24-, 48-, 72-hour Bravo pH monitoring)	Improvement in symptoms / poor outcome (with 24-, 48-, 72-hour Bravo as a comparator)

Technical efficacy

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Ahlawat et al. (2006) Division of Gastroenterology, Department of Medicine, Georgetown University Hospital, Washington, DC, USA	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	90 patients (82 patients had enough data to be included) 48 males and 42 females Age range: 18–80 years Indications: Being intolerant to the conventional naso-oesophageal pH catheter placement or had normal 24-hour oesophageal pH monitoring despite symptoms	Inclusion: Not stated Exclusion: Patients with bleeding diathesis, strictures, severe oesophagitis, varices, obstructions, pacemakers, implantable cardiac defibrillators, and surgical manipulations of the upper GI tract	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours Comparator: NA	Technical efficacy: Technical failure Day-to-day variability (% time pH <4)
Andrews et al. (2012) Division of Gastroenterology, University of Calgary, Canada	Dual centre, randomised non-blinded trial Level: II Quality: 19/26 CX P2	86 patients (40 in Calgary and 46 in Edmonton) 61 females, 25 males Aged 18–75 years Indications: Reflux symptoms Chest pain Dysphagia	Inclusion: Patients aged 18–75 years referred for ambulatory pH monitoring between August 2008 and August 2009 Exclusion: Patients with previous oesophageal surgery or achalasia	Index test: System: activated and calibrated Bravo™ wireless capsule system, Medtronic Inc., Minneapolis, MN Insertion: transorally Placement: 5 cm above the proximal border of the LES Duration: 24 (Calgary) or 48 (Edmonton) hours Comparator: Catheter: calibrated standard antimony-based pH catheter (Comfortec Plus catheter and ZepHR Sleuth recorder, Sandhill Scientific Inc., Highlands Ranch, CO) Insertion: transnasally Placement: 5 cm above the proximal border of the LOS Duration: 24 hours	Technical efficacy: Equipment malfunction Calibration failure Early detachment Insertion intolerance
Ayazi et al. (2011) Division of Thoracic	Case series	310 patients (158 males and 152 females)	Inclusion: Patients with no history of foregut	Index test: System: Bravo™ wireless capsule system	Technical efficacy: Day-to-day variability (%

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Foregut Surgery, Department of Surgery, Keck School of Medicine, University of Southern California, California, USA	Level: IV Quality: Q1 CX (no comparator) P2	Median age: 52 years (IQR=42–63)	surgery who had the Bravo capsule placed transnasally without sedation and who were monitored while off acid-suppression therapy Exclusion: Not stated	Insertion: transnasally Placement: 5 cm above the manometrically determined border of the LOS Duration: 48 hours Comparator: NA	time pH <4 and reflux events)
Azzam et al. (2012) Sao Paulo University Medical School, Sao Paulo, Brazil	In-subject simultaneous recording study Level: II Quality: 17/26 CX P2	25 patients 21 females, 4 males Mean age: 52.4 years (range 34–73) Indications: All patients had as predominant symptom the typical GORD complaints; 64% also had atypical complaints and 76% had associated extra oesophageal complaints	Inclusion: Heartburn and/or regurgitation as the main clinical complaint; at least 18 years of age; recent upper GI endoscopy (within the past 2 months); interruption in the administration of PPIs for 7 days; and signature on the free and informed consent form Exclusion: Oesophageal diverticula; strictures and varices; hiatal hernia greater than or equal to 3 cm; erosive oesophagitis Los Angeles grades C or D; Barrett's oesophagus; neoplasms; obstructive diseases; previous surgery of the GI tract	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 3 cm above the superior border of the LOS Duration: 48 hours Comparator: System: Alacer (Brazil) pH-monitor catheter with Medtronic/Synetics (USA) recording device Insertion: transnasally Placement: distal sensor was placed 3 cm above the superior border of the LOS, proximal sensor was positioned 5 cm above the superior border of the LOS Duration: 24 hours	Technical efficacy: Early detachment
Bansal et al. (2009) Divisions of Gastroenterology and Hepatology, University of Kansas School of Medicine and Veterans Affairs Medical Center, Kansas City,	Case series Level: IV Quality: Q1 CX P2	48 patients (43 males, 5 females) and 22 controls (22 males) Mean age: Patients: 57 ± 13 years Controls: 59 ± 9 years Indication: Patients: reflux symptoms	Inclusion: Patients: answering 'yes' to either heartburn or regurgitation, having a score of >2 on the RDQ Controls: answering 'no' to heartburn and regurgitation, having a score of 0 on the RDQ Exclusion: Patients: cancer or mass lesion in the oesophagus, advanced chronic liver	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 1 cm above the GOJ Duration: 24 hours Index test 2: System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the GOJ	Technical efficacy: Early detachment Cardia placement Loss of data

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Missouri, USA			disease, severe uncontrolled coagulopathy, history of oesophageal or gastric surgery Controls: use of acid suppression, evidence of erosive oesophagitis	Duration: 24 hours Comparator: NA	
Bechtold et al. (2007) Division of Gastroenterology, University of Missouri Hospital and Clinics, Columbia, Missouri, USA	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	27 patients (26 patients had enough data to be included) 6 males and 20 females Mean age: 47.6 ± 12.2 years (range 15–67)	Inclusion: Patients who underwent 48-hour Bravo ambulatory oesophageal pH monitoring for suspected GORD off any anti-reflux medications Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the GOJ Comparator: NA	Technical efficacy: Early detachment Day-to-day variability (time pH <4 (upright, supine, total), number of reflux episodes, long refluxes etc.)
Belafsky et al. (2004) Scripps Center for Voice and Swallowing, La Jolla, California, USA	Prospective case series Level: IV Quality: Q2 CX (no comparator) P2	46 patients 23 males and 23 females Mean age: 52 years Indications: GORD, chronic cough, laryngopharyngeal reflux	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transnasally Placement: 5 cm above the manometric lower oesophageal high-pressure zone or 6 cm above the endoscopic SCJ Duration: >36 hours Comparator: NA	Technical efficacy: Failures due to tight nasal vault Early detachment rate Delivery system failure rate Lost recorder rate
Bhat, McGrath & Bielefeldt (2006) Department of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA	Prospective case series Level: IV Quality: Q2 CX (no comparator) P2	217 patients (203 patients had enough data to be included) 140 females, 77 males Mean age: 51 years (range 42–58) Indications: Preoperative workup for fundoplication, possible extra oesophageal manifestations of reflux disease, symptoms after	Inclusion: Adult patients undergoing endoscopy with wireless pH studies at the University of Pittsburgh Medical Center between 9 April 2004 and 31 March 2005 Exclusion: Individuals unable to give informed consent, pregnant women, prisoners, patients with oesophageal varices, patients requiring continuing	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Technical efficacy: Operative failure rates Early detachment rate Day-to-day variability

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
		fundoplication, refractory symptoms on medication, atypical symptoms	anticoagulation, patients with cardiac defibrillators		
Bradley et al. (2011) Division of Gastroenterology and Hepatology, Mayo Clinic, Scottsdale, Arizona, USA	Cohort study Level: III-2 Quality: 15/26 CX P2	341 patients (338 patients had enough data to be included, 248 patients underwent Bravo monitoring) 62% females in the wireless group, 72% females in the catheter group Mean age: 58 years in wireless group, 57 years in catheter group	Inclusion: Outpatients undergoing ambulatory pH-metry between October 2006 and October 2008 Exclusion: Known history of oesophageal strictures, obstructions or surgical manipulations of the upper GI tract (e.g. anti-reflux surgery or oesophagectomy)	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ or 5 cm above the upper border of the LOS Duration: 48 hours Comparator (MII-pH): System: Sandhill Scientific, Inc., Highland Ranch, USA Insertion: transnasally Placement: proximal channel 5 cm above upper border of the LOS Duration: 24 hours Comparator (pH-metry): System: Alpine Biomed, Corp, Fountain Valley, US Insertion: transnasally Placement: 5 cm above the upper border of the LOS Duration: 24 hours	Technical efficacy: Early detachment Insertion intolerance Endoscopic removal
Cabrera et al. (2011) Division of Pediatric Gastroenterology, Hepatology and Nutrition, Indiana University School of Medicine, Indianapolis, Indiana, USA	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	289 children (At least 1 day of data was obtained in 278 patients, 2 days were obtained in 274 patients) 144 females, 145 males, Age: range 4–22 years Main indications: Epigastric abdominal pain (59.9%),	Inclusion: Children who are at least 4 years of age and weigh at least 30 lb Exclusion: Children with a history of oesophageal surgery, a documented anatomic abnormality, or a history of coagulopathy or a bleeding diathesis	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: at 87% of the distance from the incisors to the Z line Duration: 48 hours Comparator: NA	Technical efficacy: Early detachment Operative success/failure Recording failures

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
		vomiting (33.9%), regurgitation (23.9%), heartburn (18.3%)			
Chander et al. (2012) Department of Medicine, Section of Digestive Diseases, Yale University School of Medicine, New Haven, Connecticut, USA	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	124 patients 87 females, 37 males Age range: 24–68 years Indications: Heartburn, regurgitation, chest pain, belching or cough. And a previous negative EGD and symptoms despite PPI treatment	Inclusion: See indications Exclusion: Patients with a history of pacemakers, implantable defibrillators, oesophageal varices, severe oesophagitis, bleeding disorders, intestinal obstruction	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the GOJ Duration: 48 hours Comparator: NA	Technical efficacy: Day-to-day variability in acid exposure
Croffie et al. (2007) Section of Pediatric Gastroenterology, Hepatology and Nutrition, Indiana University School of Medicine, Indianapolis, Indiana, USA	Randomised controlled trial Level: II Quality:14.5/26 CX P2	66 children (but 5 patients were excluded due to operative failure, failed recording etc.) 32 males, 34 females Mean age: 9.4 years (range 4–16.5) Indications: Persistent epigastric or substernal pain, persistent vomiting, heartburn, chronic nocturnal cough or wheezing, persistent throat clearance, dental abnormalities suspected to be caused by reflux	Inclusion: Children between the ages of 4 and 18 years and weighing at least 30 lb who were undergoing oesophagogastroduodenoscopy and oesophageal pH monitoring for symptoms suggestive of gastro-oesophageal reflux Exclusion: Children younger than 4 years of age or weighing <30 lb, having anatomic abnormalities of the oesophagus, a history of a surgical procedure of the oesophagus, stomach or duodenum, a history of coagulopathy or bleeding diathesis, presence of oesophageal varices, significant medical illness	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: at 87% of the distance from the incisors to the Z line Duration: 48 hours Comparator: System: not stated (catheter-based) Insertion: transnasally Placement: tip at 87% of the distance from the nostril to the GOJ Duration: 24 hours	Technical efficacy: Operative success Recording failures Early detachment Day-to-day variability in acid exposure
Crowell et al. (2009) Division of Gastroenterology and Hepatology,	Case series Level: IV Quality: Q1 CX (no	157 (>95%) patients, predominantly Caucasian Indications: Heartburn/regurgitation symptoms, chest pain,	Inclusion: Patients with GORD symptoms Exclusion: History of bleeding diathesis, strictures, severe oesophagitis, varices,	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the endoscopically determined GOJ or 5 cm above the	Technical efficacy: Day-to-day variability Early detachment

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Mayo Clinic, Scottsdale, Arizona, USA	comparator) P2	abdominal pain/discomfort	obstructions, pacemakers or implantable cardiac defibrillators	manometrically determined upper margin of the LOS Duration: 48 hours <u>Comparator:</u> NA	
de Hoyos & Esparza (2010); de Hoyos, Esparza & Loredo (2009) Department of Gastroenterology, Angeles del Pedregal Hospital, Mexico City, Mexico	Prospective case series Level: IV Quality: Q2 CX (no comparator) P2	66 patients 27 males, 39 females Mean age: 41.7 years (range 11–73) <u>Indications:</u> Typical oesophageal reflux symptoms, atypical manifestations or a mixed symptomatology	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Not stated	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm from the SCJ Duration: 48 hours <u>Comparator:</u> NA	<u>Technical efficacy:</u> Poor data reception Early detachment Transmission failure
des Varannes et al. (2005) Department of Gastroenterology and Hepatology, Hopital Hotel Dieu, Nantes, France	In-subject simultaneous recording study Level: II Quality: 20/26 CX P2	40 patients (probe successfully attached in 36 patients) 21 males, 19 females Mean age: 50 years <u>Indications:</u> Heartburn (n=7), regurgitation (n=6) or both symptoms (n=26)	<u>Inclusion:</u> Patients suggestive of GORD and referred to the functional laboratory of four French academic centres for pH monitoring <u>Exclusion:</u> Severe oesophageal motility disorders, severe oesophagitis (Los Angeles grade C or above), pregnancy, women who were not using reliable contraception	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 5 cm above the Z line Duration: 48 hours <u>Comparator:</u> System: Mark III or Digitrapper pH, Medtronic, Stockholm, Sweden Insertion: transnasally Placement: adjusted to the same level as the index test Duration: 24 hours	<u>Technical efficacy:</u> Insertion intolerance Operative failure Early detachment Day-to-day variability
Doma et al. (2010) Gastroenterology Section, Temple University School of Medicine, Philadelphia,	Case series (prospective?) Level: IV Quality: Q1 CX (no	161 patients, but 147 patients had enough data to be included (44 males, 103 females) Age: 45.6 ± 16 years <u>Most common indications:</u> Heartburn (44.2%), chest pain (19.7%), cough (4.1%),	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Prior upper GI tract surgery	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 6 cm above the SCJ <u>Comparator:</u> NA	<u>Technical efficacy:</u> Inaccurate placement of the capsule

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Pennsylvania, USA	comparator) P2	regurgitation (8.8%)			
Domingues, Moraes-Filho & Domingues (2011) Motility laboratory (Gastro Resolucao Diagnostico Laboratorio de Motilidade Digestive), Rio de Janeiro, Brazil	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	100 patients (probe successfully attached in 95 patients) 39 female, 61 male Mean age: 53 years (range 18–81) Indications: 82% of patients referred to diagnose GORD, 18% referred due to persistent symptoms despite PPI use	Inclusion: Patients referred to a GI motility laboratory to investigate GORD with typical or atypical symptoms, persistent GORD symptoms on medical therapy or recurrent GORD symptoms after surgical fundoplication, between 2004 and 2009 Exclusion: <18 years of age, a history of bleeding tendency or coagulopathy, significant concomitant medical co-morbidities, severe GI bleeding in the past 6 months, a history of upper GI surgery, medication with PPIs, presence of oesophageal varices, Barrett's oesophagus, oesophageal stenosis, erosive oesophagitis, pacemaker or implantable cardiac defibrillator in situ	Index test: System: Bravo™ wireless capsule system Insertion: 5 cm above the upper border of the LOS (manometry) or 6 cm above the LOS (endoscopy) Placement: transorally Duration: 48 hours Comparator: NA	Technical efficacy: Attachment failure Technical flaws (recording failure, downloaded data lost etc.) Early detachment Day-to-day variability in oesophageal acid exposure
Francis, DL (2008) Miles and Shirley Fiterman Center for Digestive Diseases, Mayo Clinic, Rochester, Minnesota, USA	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	76 patients	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Technical efficacy: Technical failure to capture all data Early detachment Attachment failure
Francis, DO et al. (2012) Division of Gastroenterology,	Randomised blinded controlled trial	22 patients were included (13 males, 9 females) Indications: Cough, hoarseness, globus,	Inclusion: Patients >18 years of age, scheduled for planned oesophagogastroduodenoscopy and	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 16 cm above the SCJ	Technical efficacy: Operative success

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Hepatology and Nutrition, Vanderbilt University Medical Center, Nashville, Tennessee, USA	Level: II Quality: 16/26 C1 P2	heartburn, regurgitation	wireless pH monitoring for physiologic assessment of oesophageal acid exposure for typical GORD symptoms or for extraoesophageal reflux (EER) symptoms Exclusion: History of upper oesophageal surgery, bleeding dyscrasia, recent cerebrovascular accident or transient ischemic attack, recent GI haemorrhage <6 months, oesophageal varices or significant medical illness, current pregnancy	Duration: 48 hours Comparator (other half of randomisation): No monitoring (a sham capsule placement) Comparator 2: Initial test (whole population): System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the SCJ Duration: 48 hours	
Garrean et al. (2008) Division of Gastroenterology, Northwestern University Freiberg School of Medicine, Chicago, Illinois, USA	Case series (prospective?) Level: IV Quality: Q2 CX (no comparator) P2	60 patients (22 males and 38 females) Age range: 19–80 years 40 patients had complete 4-day data acquisition Indications: Acid reflux symptoms, chest pain, laryngeal symptoms that were poorly responsive to twice-daily PPI therapy for a minimum of 6 weeks	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the SCJ Comparator: NA	Technical efficacy: Loss of data transmission Day-to-day variability
Gillies et al. (2007) Departments of Upper Gastrointestinal Surgery and Gastrointestinal Physiology, Royal Berkshire hospital, Reading, UK	Non-randomised Controlled trial Level: III-2 Quality: 13.5/26 CX P2	185 patients, but 200 oesophageal pH studies 104 males, 81 females Mean age: Catheter group: 46 years (range 16–71) Bravo group: 45 years (range 13–75) Indication:	Inclusion: See indication Exclusion: No specific exclusion criteria exist	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours Comparator: System: Medtronic Insertion: transnasally	Technical efficacy: Attachment failures Technical failures (recording failures etc.) Early detachment Day-to-day variability

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
		Symptoms suggestive of GORD or documenting the outcome of laparoscopic anti-reflux surgery		Placement: 5 cm above the proximal border of the LOS Duration: 24 hours	
Grigolon et al. (2007) Gastroenterology, Department of Medical Sciences, University of Milan, Milan, Italy	Pseudo-randomised controlled trial Level: III-1 Quality: 13.5/26 CX P2	133 patients Catheter group: 78 patients, 46% males, aged 53 ± 2 years Bravo group: 55 patients, 45% males, aged 44 ± 3 years) Indication: Suspected GORD, consecutively referred for oesophageal pH monitoring	Inclusion: See indication Exclusion: Previous GI surgery (excluding cholecystectomy and appendectomy) and known or suspected stricture of the GI tract	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline mod 901P3001, Medtronic, Denmark Insertion: transnasally Placement: 6 cm above the point of sudden change in pH from acid to neutral as detected by means of continuous rapid withdrawal from the stomach, or 5 cm above the upper margin of the LOS Duration: 24 hours	Technical efficacy: Early detachment Technical failure Recording time Day-to-day variability
Grigolon et al. (2011) Department of Gastroenterology, Universita Degli Studi and Fondazione IRCCS Ca Granada, Ospedale Maggiore Policlinico, Milan, Italy	One arm of matched pairs retrospective cohort study Level: IV Quality: Q2 CX (no comparator) P2	Bravo group: 57 patients, of which 51 had enough data to be included 22 males, mean age 48 years (range 23–82)	Inclusion: All consecutive patients who underwent wireless 96-hour pH monitoring between January and December 2007 off PPI therapy, and patients undergoing traditional 24-hour pH monitoring in the same period, after matching for the most troublesome symptom potentially related to GORD, sex and age Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48–96 hours Comparator: System: Slimline mod 901P3001, Medtronic, Denmark Insertion: transnasally Placement: 6 cm above the point of sudden change in pH from acid to neutral as detected by means of continuous rapid withdrawal from the stomach, or 5 cm above the upper margin of the LOS	Technical efficacy: Early detachment Incomplete data capture

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
				Duration: 24 hours	
Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008) Department of Paediatric Surgery, Lund University Hospital, Lund, Sweden	Case series Level: IV Quality: Q1 CX (no comparator) P2	58 children 37 males, 21 females Mean age: 8 ± 4 years Indications: Vomiting, abdominal and chest pain, GORD control, GI bleeding, respiratory symptoms	Inclusion: Not stated Exclusion: Oesophageal strictures, coagulopathy, any suspicion of intestinal strictures	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: above the diaphragm valve at the width of two of the individual patient's vertebral bodies as visualised on a perioperative X-ray Duration: 24 or 48 hours (23 children had 48-hour measurement) Comparator: NA	Technical efficacy: Attachment failure Technical failure Day-to-day variability in oesophageal acid exposure
Hakanson et al. (2009) Department of Surgery, Center of Gastrointestinal Disease, Ersta Hospital, Stockholm, Sweden	Case-control study with in-subject simultaneous recording Level: III-2 Quality: 14/26 CX P2	53 volunteers and 55 patients (45 volunteers and 47 patients had enough data to be included) 30 females and 15 males in volunteer group 27 females and 20 males in patient group Mean age: Volunteers: 47 years (range 21–68) Patients: 50 years (range 23–69) Indication: Patients had symptoms including mainly heartburn, acid regurgitation, and chest or epigastric pain	Inclusion: Patients: referred to the clinic for investigation of GORD Exclusion: Volunteers: history of GORD or any other pathology of the upper GI tract, taking acid-suppressing agents or medication known to affect GI motility	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline, Medtronic, Skovlunde, Denmark Insertion: transnasally Placement: 5 cm proximal to the upper border of the LOS Duration: 24 hours	Technical efficacy: Attachment failure Early detachment Technical failure (recording failures etc.) Day-to-day variability in oesophageal acid exposure
Hochman and Favaloro-Sabatier (2005) Children's Center	Retrospective case series Level: IV	50 patients (44 patients had enough data to be included) 27 males, 17 females Mean age: 11.8 years (range 6–	Inclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the proximal border	Technical efficacy: Early detachment Attachment failure Day-to-day variability in

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
for Digestive Health Care, Atlanta, Georgia, USA	Quality: Q2 CX (no comparator) P2	19)	Exclusion: Not stated	of the LOS Duration: 48 hours Comparator: NA	oesophageal acid exposure
Iqbal et al. (2007) Department of Surgery, University of Nebraska Medical Center, Omaha, Nebraska, USA	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	100 patients 58 females, 42 males Mean age: Women: 40 ± 14 years Men: 42 ± 16 years	Inclusion: Consecutive patients with GORD symptoms and undergoing Bravo pH monitoring over a 1-year period Exclusion: Patients with surgical history of the upper GI tract, bleeding diathesis or coagulopathy, oesophageal strictures, severe GI bleeding in the past 3 months, advanced cirrhosis or significant comorbidities	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Comparator: NA	Technical efficacy: Early detachment
Korrapati et al. (2011) Department of Gastroenterology and Hepatology, Winthrop University Hospital, Mineola, New York, USA	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	108 patients (47 male, 61 female) Mean age: 54.74 ± 14.67 years Indications: Heartburn (58%), chest pain (30%), chronic cough (8%), laryngitis (21%), bloating (35%)	Inclusion: Not stated Exclusion: Patients on anti-reflux medication	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the GOJ Duration: 48 hours Comparator: NA	Technical efficacy: Day-to-day variability (% time pH <4)
Lacy et al. (2009) Division of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	51 patients (50 patients had enough data to be included) 24 males, 26 females Mean age: 13 years (range: 5–17) Indications: Heartburn, regurgitation, abdominal pain, nausea and/or vomiting, chest pain	Inclusion: Not stated Exclusion: Prior surgery to the oesophagus or stomach, known intestinal obstruction or oesophageal varices, achalasia, inability to understand the procedure, inability to undergo oesophageal manometry or sedation	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Technical efficacy: Day-to-day variability in oesophageal acid exposure

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Lacy, Chehade & Crowell (2011) Section of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA	Case series Level: IV Quality: Q1 CX (no comparator) P2	358 patients (257 females, 101 males) Mean age: 51 ± 14 years Predominantly (97%) Caucasian <u>Indications:</u> Symptoms thought secondary to GORD	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Not stated	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the GOJ or 5 cm above the upper border of the LOS Duration: 48 hours <u>Comparator:</u> NA	<u>Technical efficacy:</u> Early detachment Loss of signal
Lee et al. (2005) Department of Internal Medicine, Department of Emergency Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan	Prospective case series Level: IV Quality: Q3 CX (no comparator) P2	40 patients (27 males and 13 females) Age: range 28–83 years <u>Indications:</u> Typical heartburn with unsatisfactory response to therapy (15), erosive oesophagitis and undergoing preoperative evaluation (25)	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Not stated	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours <u>Comparator:</u> NA	<u>Technical efficacy:</u> Attachment failure Early detachment
Marchese et al. (2006) Digestive Endoscopy Unit, Agostino Gemelli University Hospital, Catholic University of Rome, Rome, Italy	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	40 patients (36 patients had successful oesophageal placement of the capsule) 23 males, 17 females Mean age: 42.5 years <u>Indications:</u> Reflux symptoms that were refractory to PPI therapy	<u>Inclusion:</u> Patients with reflux symptoms refractory to PPI therapy who had normal or equivocal endoscopic findings <u>Exclusion:</u> Patients with oesophageal varices, previous GI surgery, known coagulopathy or bleeding diathesis, cardiac pacemakers, defibrillators or other electromedical devices, significant comorbidity, pregnant women and patients <18 years of age	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transnasally Placement: 5 cm above the upper margin of the LOS Duration: 24 or 48 hours <u>Comparator:</u> NA	<u>Technical efficacy:</u> Technical failure Capsule displacement
Martinez de Haro et	One arm of non-	40 patients	<u>Inclusion:</u>	<u>Index test:</u>	<u>Technical efficacy:</u>

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
al. (2008) Servicio de Cirugia General, Hospital Universitario Virgen de la Arrixaca, Murcia, Spain	randomised controlled trial Level: IV Quality: Q2 CX P2	Group 2 standard pH monitoring: mean age 45 years (range 17–73); 9 males and 6 females Group 2 Bravo monitoring: mean age 37 years (range 21–59); 8 males and 7 females Group 3 (both): mean age, 36.5 years (range 18–56); 7 males and 3 females	General criteria: >18 years of age, informed consent of the risks, benefits and alternative tests, committed to completing follow-up Control group: less than 2 episodes of pyrosis or regurgitation per month, no dysphagia or atypical reflux symptoms, no oesophageal motility disorders, not taken medication for acid reflux, no evidence of hiatal hernia or oesophagitis Patient group: typical reflux symptoms requiring treatment, inflammatory lesions due to reflux in the oesophageal mucosa in group 2 Exclusion: <18 years of age, oesophageal stenosis, oesophageal varix, lesion affecting the nostrils, severe oesophageal motor disorders, high anaesthetic risk, history of coagulopathy, haemorrhage, intake of anticoagulants or platelet antiaggregates, myocardial infarction or CVA in the past 6 months, pregnancy, history of radiotherapy in the thoracic region, digestive haemorrhage in the past 6 months, any known medical disorder that could alter the data, unable to accept study protocols	System: Bravo™ wireless capsule system Insertion: transnasally (n=25) and transorally (n=10) Placement: 5 cm above the LOS Duration: 48 hours <u>Comparator:</u> System: Digitrapper Mark III pH monitoring (Synetics, Stockholm, Sweden) Insertion: transnasally Placement: 5 cm above the LOS Duration: 24 hours	Incorrect placement Early detachment Immediate detachment
Nusrat, Roy & Bielefeldt (2012) University of Pittsburgh Medical Center, Pittsburgh,	Retrospective case series Level: IV Quality: Q1	356 patients (65% females) Mean age: 49.9 ± 0.8 years <u>Indications:</u> Evaluations of GORD, recurrent or persistent symptoms after	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Patients with pacemakers, implanted cardiac defibrillators, known	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the endoscopically localized SCJ or 5 cm above the	<u>Technical efficacy:</u> Early capsule detachment

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Pennsylvania, USA	CX (no comparator) P2	fundoplication, chest pain, unexplained cough or other atypical symptoms	oesophageal varices, bleeding disorders or ongoing anticoagulation	manometrically determined upper border of the LOS Duration: 48 hours Comparator: NA	
Pandolfino et al. (2003) Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA	Case-control study Level: IV Quality: Q1 CX (no comparator) P3	44 controls (13 males, 31 females, aged 23–53 years) 41 patients (26 males, 15 females, aged 32–72 years)	Inclusion: Not stated Exclusion: Controls: abdominal symptoms, use of antacids or antisecretory medication, abnormal endoscopy Patients: history of surgical manipulation of the upper GI tract, history of bleeding diathesis or coagulopathy, stroke or transient ischemic attack in the past 6 months, significant medical illness, oesophageal varices	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Comparator: NA	Technical efficacy: Early detachment Attachment failure Day-to-day variability in oesophageal acid reflux
Pandolfino et al. (2006) Division of Gastroenterology, Department of Medicine, the Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA	Case-control study Level: IV Quality: Q2 CX (no comparator) P3	10 normal control subjects (7 males and 3 females, aged 21–53 years) 10 patients with GORD (6 males and 4 females, aged 20–54 years)	Inclusion: Controls: no history of heartburn, reflux, chest pain or atypical symptoms Patients: oesophagitis (3) or an abnormal 24-hour oesophageal pH study prior to this study Exclusion: History of surgical manipulation of the upper GI tract, anticoagulation therapy, significant medical disease Controls: abdominal symptoms, use of antacids, abnormal upper endoscopy	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 1 cm proximal to SCJ Duration: 24 hours Comparator: NA	Technical efficacy: Operative success Early detachment Concordance capsule 1 cm vs 6 cm proximal to SCJ
Park et al. (2013) Divisions of Gastroenterology, Department of	Case series Level: IV Quality: Q1	230 patients (101 males, 129 females) Mean age: 49.7 ± 12.4 years Indication:	Inclusion: ≥18 years of age Exclusion: Previous oesophageal, gastric or	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to SCJ	Technical efficacy: Operative success Attachment failure Early detachment

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
internal Medicine, Seoul; St Mary's Hospital, the Catholic University of Korea, Seoul, Korea	CX (no comparator) P2	suspected GORD	duodenal surgery, GI organic disease, significant comorbidity or GI bleeding	Duration: 48 hours Comparator: NA	Recording failure
Prakash & Clouse (2006) Division of Gastroenterology, Washington University School of Medicine, St Louis, Missouri, USA	Case series Level: IV Quality: Q1 CX (no comparator) P2	62 patients (41 females, 21 males) Mean age: 47 ± 2 years	Inclusion: Not stated Exclusion: Oesophagitis, prior anti-reflux surgery, failure to maintain a complete diary of activities and symptoms, having <14 hours of good-quality tracings on each of the study days	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Comparator: NA	Technical efficacy: Day-to-day variability
Remes-Troche et al. (2005) Department of Gastroenterology, Instituto Nacional de Ciencias Medicas y Nutricion, Salvador Zubiran, Mexico City, Mexico	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	84 patients (77 patients had enough data to be included) 49 females, 35 males Mean age: 44 ± 12 years (range 19–73) Indications: Persistent GORD symptoms on PPI, preoperative evaluation before surgery, previous failed transnasal pH monitoring, extraoesophageal GORD	Inclusion: Consecutive patients with GORD symptoms with an indication for 24-hour pH monitoring Exclusion: Patients with surgical manipulation of the upper GI tract, bleeding diathesis or coagulopathy, severe GI bleeding in the past 6 months, oesophageal varices, significant comorbidities	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Technical efficacy: Attachment failure Early detachment Day-to-day variability in acid exposure
Scarpulla et al. (2007) Gastroenterology Division, M. Raimondi Hospital, Can Cataldo, Italy	Retrospective case series Level: IV Quality: Q2 CX P2	83 patients (complete 4-day recordings were available for 34/83) 44 females, 39 males Median age 42 years (range 18–63) Indications: Typical reflux symptoms or	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the Z line Duration: 24, 48, 72 and/or 96 hours Comparator: NA	Technical efficacy: Early detachment Technical failure

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
		atypical symptoms including chest pain and extra oesophageal complaints			
Sofi et al. (2011) Department of Gastroenterology, University of Toledo Medical Center, Toledo, Ohio, USA	Retrospective case series	58 patients Age range: 26–87 years 17 males, 41 females	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the GOJ Comparator: NA	Technical efficacy: Early detachment (none) Device malfunction Recording failure
Tankurt et al. (2011) Kent Hospital, Department of Gastroenterology, Izmir, Turkey	Case series Level: IV Quality: Q2 CX (no comparator) P2	64 patients (37 males, 27 females) Mean age: 37.9 years (range 25–58) Indications: Preoperative evaluation (17), evaluation of atypical reflux symptoms (40) and 'other' (7)	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Technical efficacy: Early detachment
Tseng et al. (2005) Digestive Health Center, Department of Surgery and Division of Gastroenterology, Oregon health and Science University, Portland, Oregon, USA	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	190 procedures and 186 patients (some had the test performed twice) 72 males and 114 females Mean age: 51 years Indication: Being evaluated for GORD	Inclusion: Patients older than 18 years Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 5–6cm above the LOS or 5–6 m above the SCJ Duration: 48 hours Comparator: 24-hour Bravo monitoring	Technical efficacy: Loss of signal/data Early detachment Day-to-day variability
Tu et al. (2004) Department of Internal Medicine and Department of Emergency Medicine, National	Prospective case series Level: IV Quality: Q1 CX (no	30 patients 20 males, 10 females Mean age: 57.6 ± 14.3 years (range 29–83) Indications: 24 patients with GORD were	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the GOJ Duration: 48 hours	Technical efficacy: Attachment failure

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
Taiwan University Hospital, Taipei, Taiwan	comparator) P2	evaluated for drug effectiveness before surgery, 6 patients showed typical heartburn but their endoscopic findings were normal		Comparator: NA	
Turner et al. (2007) Division of Gastroenterology, Brigham and Women's Hospital, Boston, Massachusetts, USA	Case series Level: IV Quality: Q2 CX (no comparator) P2	198 pH studies (148 studies had enough data to be included) Off medical therapy: 115 patients, 76 females, 39 males. Mean age: 50 ± 13 years On medical therapy: 33 patients, 26 females, 7 males. Mean age: 52 ± 14 years Indications: Typical and atypical symptoms of GORD or persistent GORD-related symptoms despite medical therapy. Patients with planned anti-reflux surgery were also evaluated	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: (during endoscopy) Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Bravo™ wireless capsule system Insertion: transnasally Placement: Duration: 24 hours	Technical efficacy Early detachment Insufficient data capture Day-to-day variability in oesophageal acid exposure
Ward et al. (2004) Division of Gastroenterology and Hepatology, Mayo Clinic, Jacksonville, Florida, USA	Case series Level: IV Quality: Q2 CX (no comparator) P2	60 patients (58 patients had enough data to be included) 26 males, 34 females Mean age: 54 years (range 27–82) Indications: Document GORD before surgery, possible GORD with negative PPI trial, evaluate response to PPI, possible supra-oesophageal GORD, non-cardiac chest pain, evaluate response to Stretta	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: NA	Technical efficacy: Attachment failure Hardware malfunction

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Technical efficacy outcomes assessed
		procedure, previous failed pH test			
Wenner, Johnsson et al. (2007) Department of Surgery, Lund University Hospital, Lund, Sweden	Randomised crossover trial Level: II Quality: 20/26 CX P2	35 patients (31 patients had enough data to be included) 15 males, 16 females Median age: 52 years Median BMI: 27 (range 24–28) Indications: Typical reflux symptoms (55%) and atypical reflux symptoms (45%)	Inclusion: Patients referred for oesophageal pH monitoring Exclusion: Patients with coagulopathy, severe oesophagitis/stricture, portal hypertension, a pacemaker, age <20 or >70 years, previous gastro-oesophageal surgery, severe cardiopulmonary disease, symptoms suggestive of oesophageal motor disorders	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline (Medtronic) Insertion: transnasally Placement: 5 cm above the upper border of the LOS Duration: 24 hours	Technical efficacy: Technical problems
Wenner et al. (2008); Wenner. Johansson et al. (2007) Department of Surgery, Lund University Hospital, Lund, Sweden	Case-control study Level: IV Quality: Q2 CX P3	70 patients, but 64 had enough data to be included (39 males, 25 females) and 55 controls (27 males and 28 females) Not all patients are included in the distal capsule outcomes Indications: Patients with typical reflux symptoms	Inclusion: Patients: presence of typical reflux symptoms such as heartburn and regurgitation. Symptoms had to occur more than twice weekly. Treatment with medical acid suppression had to result in complete symptom relief or a more than 50% reduction of the symptoms Exclusion: Coagulopathy, severe oesophagitis/stricture, portal hypertension, pacemaker, <20 or >70 years of age, history of gastro-oesophageal surgery, severe cardiopulmonary disease, symptoms suggestive of oesophageal motor disorders	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ	Technical efficacy: Early detachment Capsule unable to pass through oesophagus Transmission failure / loss of signal Incorrect capsule placement

Study profiles of included studies on linked evidence

Diagnostic accuracy

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Ayazi et al. (2009)[11] Department of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California, USA	Diagnostic case-control study Level: III-3 Quality: Patient selection: ? Index test: ? Reference standard: ☺ Flow and timing: ? CX P3	25 asymptomatic subjects and 25 patients Indications: Patients had heartburn and were responsive to PPI therapy. All had hiatal hernia	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system, Medtronic Inc. Minneapolis Minnesota Insertion: transnasally Placement: 5 cm above the LOS Duration: 48 hours Comparator: NA Reference standard: Clinical diagnosis (heartburn, responsive to PPI therapy and evidence of a hiatal hernia)	Sensitivity Specificity Accuracy
Ayazi et al. (2009)[11] Department of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California, USA	Diagnostic case-control study Level: III-3 Quality: Patient selection: ☺ Index test: ? Reference	115 symptomatic patients who were tested for GORD (28 patients with strong clinical evidence, 77 with intermediate evidence, and 10 with minimal evidence, of GORD	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system, Medtronic Inc. Minneapolis Minnesota Insertion: transnasally Placement: 5 cm above the LOS Duration: 48 hours Comparator: NA	Sensitivity Specificity PPV NPV Accuracy

[11] This article consisted of multiple studies, of which two were included.

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
	standard: ☺ Flow and timing: ☺ CX P3			Reference standard: Clinical diagnosis (primary symptoms of heartburn and/or regurgitation, response to PPI therapy >50%, hiatal hernia >2 cm, oesophageal mucosal injury)	
Bansal et al. (2009) Divisions of Gastroenterology and Hepatology, University of Kansas, School of Medicine and Veterans Affairs Medical Center, Kansas City, Missouri, USA	Diagnostic case-control study Level: III-3 Quality: Patient selection: ☹ Index test: ? Reference standard: ☹ Flow and timing:? C1 P2	48 patients (43 males, 5 females) 22 controls (22 males) Mean age: Patients: 57 ± 13 years Controls: 59 ± 9 years Indication: Reflux symptoms	Inclusion: Patients: answering 'yes' to either heartburn or regurgitation, having a score of >2 on the RDQ Controls: answering 'no' to heartburn and regurgitation, having a score of 0 on the RDQ Exclusion: Patients: cancer or mass lesion in the oesophagus, advanced chronic liver disease, severe uncontrolled coagulopathy, history of oesophageal or gastric surgery Controls: use of acid suppression, evidence of erosive oesophagitis	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 1 cm above the GOJ Duration: 24 hours Index test 2: System: Bravo™ wireless capsule system Insertion: NS placement: 6 cm above the GOJ Duration: 24 hours Reference standard (clinical diagnosis): Answering 'yes' to either heartburn or regurgitation and having a score of at least 2 on the RDQ	Sensitivity Specificity
des Varannes et al. (2005) Department of Gastroenterology and Hepatology, Hopital Hotel Dieu, Nantes, France	In-subject simultaneous recording study Level: II Quality: Patient selection: ? Index test: ☺ Reference standard: ? Flow and timing: ☺ C1 P2	40 patients (probe successfully attached in 36 patients) 21 males, 19 females Mean age: 50 years Indications: Heartburn (n=7), regurgitation (n=6) or both symptoms (n=26)	Inclusion: Patients suggestive of GORD and referred to the functional laboratory of four French academic centres for pH monitoring Exclusion: Severe oesophageal motility disorders, severe oesophagitis (Los Angeles grade C or above), pregnancy, women who were not using reliable contraception	Index test: System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 5 cm above the Z line Duration: 48 hours Comparator: System: Mark III or Digitrapper pH, Medtronic, Stockholm, Sweden Insertion: transnasally Placement: adjusted to the same level as the index test Duration: 24 hours	Sensitivity Specificity PPV NPV LR+ LR−

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Hakanson et al. (2009) Department of Surgery, Center of Gastrointestinal Disease, Ersta Hospital, Stockholm, Sweden	Case-control study with in-subject simultaneous recording Level: III-3 Quality: Patient selection: ? Index test: ☹ Reference standard: ? Flow and timing: ☺ C1 P2	53 volunteers and 55 patients (45 volunteers and 47 patients had enough data to be included) 30 females and 15 males in volunteer group 27 females and 20 males in patient group Mean age: Volunteers: 47 years (range 21–68) Patients: 50 years (range 23–69) <u>Indication:</u> Patients had symptoms including mainly heartburn, acid regurgitation, and chest or epigastric pain	<u>Inclusion:</u> Patients: referred to the clinic for investigation of GORD <u>Exclusion:</u> Volunteers: history of GORD or any other pathology of the upper GI tract, taking acid-suppressing agents or medication known to affect GI motility	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48 hours <u>Comparator:</u> System: Slimline, Medtronic, Skovlunde, Denmark Insertion: transnasally Placement: 5 cm proximal to the upper border of the LOS Duration: 24 hours	Sensitivity Specificity PPV NPV LR+ LR–
Pandolfino et al. (2003) Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA	Diagnostic case-control study Level: III-3 Quality: Patient selection: ☹ Index test: ? Reference standard: ? Flow and timing: ☹ Q1 CX (clinical diagnosis)	44 controls (13 males, 31 females, aged 23–53 years) 41 patients (26 males, 15 females, aged 32–72 years)	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Controls: abdominal symptoms, use of antacids or antisecretory medication, abnormal endoscopy Patients: history of surgical manipulation of the upper GI tract, history of bleeding diathesis or coagulopathy, stroke or transient ischemic attack in the past 6 months, significant medical illness, oesophageal varices	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ <u>Comparator:</u> NA <u>Reference standard (diagnosis of patients in study population):</u> 25 patients had erosive oesophagitis, 6 had an abnormal 24-hour pH study, 10 were diagnosed clinically with typical GORD symptoms that resolved with PPI therapy	Sensitivity Specificity PPV NPV LR+ LR–

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
	P3				
Wenner, Johnsson et al. (2007) Department of Surgery, Lund University Hospital, Lund, Sweden	Randomised crossover trial Level: II Quality: Patient selection: ☺ Index test: ☺ Reference standard: ? Flow and timing: ☺ C1 P2	35 patients (31 patients had enough data to be included) 15 males, 16 females Median age: 52 years Median BMI: 27 (range 24–28) Indications: Typical reflux symptoms (55%) and atypical reflux symptoms (45%)	Inclusion: Patients referred for oesophageal pH monitoring Exclusion: Patients with coagulopathy, severe oesophagitis/stricture, portal hypertension, a pacemaker, age <20 or >70 years, previous gastro-oesophageal surgery, severe cardiopulmonary disease, symptoms suggestive of oesophageal motor disorders	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline (Medtronic) Insertion: transnasally Placement: 5 cm above the upper border of the LOS Duration: 24 hours	Sensitivity Specificity PPV NPV LR+ LR–
Wenner et al. (2008); Wenner, Johansson et al. (2007) Department of Surgery, Lund University Hospital, Lund, Sweden	Diagnostic case-control study Level: III-3 Quality: Patient selection: ☹ Index test: ? Reference standard: ☺ Flow and timing: ? CX P3	70 patients, but 64 had enough data to be included (39 males, 25 females) 55 controls (27 males and 28 females) Not all patients are included in the distal capsule outcomes Indications: Patients with typical reflux symptoms	Inclusion: Patients: presence of typical reflux symptoms such as heartburn and regurgitation. Symptoms had to occur more than twice weekly. Treatment with medical acid suppression had to result in complete symptom relief or a more than 50% reduction of the symptoms Exclusion: Coagulopathy, severe oesophagitis/stricture, portal hypertension, pacemaker, <20 or >70 years of age, history of gastro-oesophageal surgery, severe cardiopulmonary disease, symptoms suggestive of oesophageal motor disorders	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Reference standard (for sensitivity and specificity): Clinical diagnosis: symptoms suggestive of GORD and responding to acid suppression	Sensitivity Specificity LR+ LR–

Other accuracy—(incremental) diagnostic yield and concordance outcomes

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Ang et al. (2010) Division of Gastroenterology, Department of Medicine, Changi General Hospital, Singapore	Comparative diagnostic yield study Level: III-2 Quality: 14/26 C1 P2	125 patients (only 66 with catheter-free monitoring and 51 with catheter-based monitoring were analysed, 4 were unable to complete the catheter study due to intolerance) 51 males and 70 females Mean age: 45.14 years (range 22–82) Race: 76.9% Chinese 5.8% Malay 10.7% Indian 7.4% Other Indication: Typical or atypical reflux symptoms in the absence of endoscopic features of reflux oesophagitis	Inclusion: Patients who were referred for investigation of persistent reflux-like symptoms suspected to be related to NERD despite the presence of a normal gastroscopy over a 50-month period from January 2004 to February 2009 Exclusion: A history of bleeding tendency, coagulopathy, peptic ulcer disease, previous oesophageal or gastric surgery, known oesophageal or fundic varices, active GI bleeding, endoscopic features of oesophagitis based on Los Angeles classification, Barrett's oesophagus, peptic strictures, herpetic or Candida oesophagitis, Zollinger-Ellison syndrome, progressive systemic sclerosis, oesophageal or upper small intestinal Chrohn's disease, active solid tumor neoplastic tissue, pregnancy, presence of a pacemaker or implantable cardiac defibrillator, a history of severe cardiopulmonary disease or symptoms suggestive of oesophageal motor disorders	Index test: System: Bravo™ wireless capsule system, Medtronic Inc. Minneapolis, Minnesota Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: Catheter: calibrated antimony pH electrode (Slimline, Medtronic, Shoreview, Minnesota, USA) Insertion: transnasally Placement: 5 cm above the proximal border of the LOS Duration: 24 hours	Accuracy: Comparative diagnostic yield Concordance (acid exposure time and reflux events)
Ayazi et al. (2011) Division of Thoracic Foregut Surgery, Department of Surgery, Keck School of Medicine, University of	Diagnostic yield study Level: IV	310 patients (158 males and 152 females) Median age: 52 years (IQR=42–63)	Inclusion: Patients with no history of foregut surgery who had the Bravo capsule placed transnasally without sedation and who were monitored while off acid suppression therapy Exclusion:	Index test: System: Bravo™ wireless capsule system Insertion: transnasally Placement: 5 cm above the manometrically determined border of the LOS Duration: 48 hours	Accuracy: Diagnostic yield

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Southern California, California, USA			Not stated		
Azzam et al. (2012) Sao Paulo University Medical School, Sao Paulo, Brazil	In-subject simultaneous recording study Level: II Quality: 17/26 C1 P2	25 patients 21 females, 4 males Mean age: 52.4 years (range 34–73) Indications: All patients had as predominant symptom the typical GORD complaints, 64% had atypical complaints and 76% had associated extraoesophageal complaints	Inclusion: Heartburn and/or regurgitation as the main clinical complaint; at least 18 years of age, recent upper GI endoscopy (within the past 2 months), interruption in the administration of PPIs for 7 days and signature on the free and informed consent form Exclusion: Oesophageal diverticula; strictures and varices; hiatal hernia greater than or equal to 3 cm; erosive oesophagitis Los Angeles grades C or D; Barrett's oesophagus; neoplasms; obstructive diseases; previous surgery of the GI tract	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 3 cm above the superior border of the LOS Duration: 48 hours Comparator: System: Alacer (Brazil) pH-monitor catheter with Medtronic/Synetics (USA) recording device Insertion: transnasally Placement: Distal sensor was placed 3 cm above the superior border of the LOS, proximal sensor was positioned 5 cm above the superior border of the LOS Duration: 24 hours	Accuracy: Concordance (acid exposure time) Comparative diagnostic yield
Bansal et al. (2009) Divisions of Gastroenterology and Hepatology, University of Kansas, School of Medicine and Veterans affairs Medical Center, Kansas City, Missouri, USA	Diagnostic yield study Level: IV	48 patients (43 males, 5 females) 22 controls (22 males) Mean age: Patients: 57 ± 13 years Controls: 59 ± 9 years Indication: Patients: reflux symptoms	Inclusion: Patients: answering 'yes' to either heartburn or regurgitation, having a score of >2 on the RDQ Controls: answering 'no' to heartburn and regurgitation, having a score of 0 on the RDQ Exclusion: Patients: cancer or mass lesion in the oesophagus, advanced chronic liver disease, severe uncontrolled coagulopathy, history of oesophageal or gastric surgery Controls: use of acid suppression, evidence of erosive oesophagitis	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: 1 cm above the GOJ Duration: 24 hours	Accuracy: Diagnostic yield

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Bhat, McGrath & Bielefeldt (2006) Department of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA	Diagnostic yield study Level: IV	217 patients (203 patients had enough data to be included) 140 females, 77 males Mean age: 51 years (range 42–58) Indications: Preoperative workup for fundoplication, possible extra oesophageal manifestations of reflux disease, symptoms after fundoplication, refractory symptoms on medication, atypical symptoms	Inclusion: Adult patients undergoing endoscopy with wireless pH studies at the University of Pittsburgh Medical Center between 9 April 2004 and 31 March 2005 Exclusion: Individuals unable to give informed consent, pregnant women, prisoners, patients with oesophageal varices, patients requiring continuing anticoagulation, patients with cardiac defibrillators	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours	Accuracy: Diagnostic yield
Croffie et al. (2007) Section of Pediatric Gastroenterology, Hepatology and Nutrition, Indiana University School of Medicine, Indianapolis, Indiana, USA	Randomised controlled trial Level: II Quality:14.5/26 C1 P2	66 patients (but 5 patients were excluded due to operative failure, failed recording etc.) 32 males, 34 females Mean age: 9.4 years (range 4–16.5) Indications: Persistent epigastric or substernal pain, persistent vomiting, heartburn, chronic nocturnal cough or wheezing, persistent throat clearance, dental abnormalities suspected to be caused by reflux	Inclusion: Children between the ages of 4 and 18 years and weighing at least 30 lb who were undergoing oesophagogastroduodenoscopy and oesophageal pH monitoring for symptoms suggestive of gastro-oesophageal reflux Exclusion: Children younger than 4 years of age or weighing <30 lb, having anatomic abnormalities of the oesophagus, a history of a surgical procedure of the oesophagus, stomach or duodenum, a history of coagulopathy or bleeding diathesis, presence of oesophageal varices, significant medical illness	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: at 87% of the distance from the incisors to the Z line Duration: 48 hours Comparator: System: not stated (catheter-based) Insertion: transnasally Placement: tip at 87% of the distance from the nostril to the GOJ Duration: 24 hours	Accuracy: Concordance (acid exposure time)
des Varannes et al. (2005)	In-subject simultaneous	40 patients (probe successfully	Inclusion: Patients suggestive of GORD and	Index test: System: Bravo™ wireless capsule system	Accuracy: Concordance (acid

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Department of Gastroenterology and Hepatology, Hopital Hotel Dieu, Nantes, France	recording study Level: II Quality: 20/26 C1 P2	attached in 36 patients) 21 males, 19 females Mean age: 50 years <u>Indications:</u> Heartburn (n=7), regurgitation (n=6) or both symptoms (n=26)	referred to the functional laboratory of four French academic centres for pH monitoring <u>Exclusion:</u> Severe oesophageal motility disorders, severe oesophagitis (Los Angeles grade C or above), pregnancy, women who were not using reliable contraception	Insertion: transnasally or transorally Placement: 5 cm above the Z line Duration: 48 hours <u>Comparator:</u> System: Mark III or Digitrapper pH, Medtronic, Stockholm, Sweden Insertion: transnasally Placement: adjusted to the same level as the index test Duration: 24 hours	exposure time and reflux events)
Domingues, Moraes-Filho & Domingues (2011) Motility laboratory (Gastro Resolucao Diagnostico, Laboratorio de Motilidade Digestive), Rio de Janeiro, Brazil	Diagnostic yield study Level: IV	100 patients (probe successfully attached in 95 patients) 39 female, 61 male Mean age: 53 years (range 18 –81) <u>Indications:</u> 82% of patients referred to diagnose GORD, 18% referred due to persistent symptoms despite PPI use	<u>Inclusion:</u> Patients referred to a GI motility laboratory to investigate GORD with typical or atypical symptoms, persistent GORD symptoms on medical therapy or recurrent GORD symptoms after surgical fundoplication, between 2004 and 2009 <u>Exclusion:</u> <18 years of age, history of bleeding tendency or coagulopathy, significant concomitant medical co-morbidities, severe GI bleeding in the past 6 months, a history of upper GI surgery, medication with PPIs, presence of oesophageal varices, Barrett's oesophagus, oesophageal stenosis, erosive oesophagitis, pacemaker or implantable cardiac defibrillator in situ	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: 5 cm above the upper border of the LOS (manometry) or 6 cm above the LOS (endoscopy) Placement: transorally Duration: 48 hours	<u>Accuracy:</u> Diagnostic yield
Garrean et al. (2008) Division of Gastroenterology, Northwestern	Diagnostic yield study Level: IV	60 patients (22 males and 38 females) Age: 19–80 years 40 patients had complete	<u>Inclusion:</u> Not stated <u>Exclusion:</u> Not stated	<u>Index test:</u> System: Bravo™ wireless capsule system Insertion: NS Placement: 6 cm above the SCJ	<u>Accuracy:</u> Diagnostic yield Incremental diagnostic yield

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
University Freiberg School of Medicine, Chicago, Illinois, USA		4-day data acquisition Indications: Acid reflux symptoms, chest pain, laryngeal symptoms that were poorly responsive to twice-daily PPI therapy for a minimum of 6 weeks		Comparator: NA	
Gillies et al. (2007) Departments of Upper Gastrointestinal Surgery and Gastrointestinal Physiology, Royal Berkshire Hospital, Reading, UK	Diagnostic yield study Level: IV	95 patients Indication: Symptoms suggestive of GORD or documenting the outcome of laparoscopic anti-reflux surgery	Inclusion: See indication Exclusion: No specific exclusion criteria exist	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours	Accuracy: Diagnostic yield
Grigolon et al. (2011) Department of Gastroenterology, Universita Degli Studi and Fondazione IRCCS Ca Granada, Ospedale Maggiore Policlinico, Milan, Italy	Matched-pairs retrospective cohort study Level: III-2 Quality: 16.5/26 C1 P2	102 patients Catheter group: 51 patients, 20 males, mean age 48 years (range 22–78) Bravo group: 51 patients, 22 males, mean age 48 years (range 23–82)	Inclusion: All consecutive patients who underwent wireless 96-hour pH monitoring between January and December 2007 off PPI therapy, and patients undergoing traditional 24-hour pH monitoring in the same period, after matching for the most troublesome symptom potentially related to GORD, sex and age Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48–96 hours Comparator for concordance: System: Slimline mod 901P3001, Medtronic, Denmark Insertion: transnasally Placement: 6 cm above the point of sudden change in pH from acid to neutral as detected by means of continuous rapid withdrawal from the stomach, or 5 cm above the upper margin of the LOS Duration: 24 hours	Accuracy: Diagnostic yield Concordance (acid exposure time) Incremental diagnostic yield (of prolonged monitoring)
Gunnarsdottir,	Diagnostic yield	58 children	Inclusion:	Index test:	Accuracy:

Catheter free ambulatory oesophageal pH monitoring for GORD – MSAC 1203   Page 210 of 270

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Stenstrom & Arnbjornsson (2007, 2008) Department of Paediatric Surgery, Lund University Hospital, Lund, Sweden	study Level: IV	37 males, 21 females Mean age: 8 ± 4 years Indications: Vomiting, abdominal and chest pain, GORD control, GI bleeding, respiratory symptoms	Not stated Exclusion: Oesophageal strictures, coagulopathy, any suspicion of intestinal strictures	System: Bravo™ wireless capsule system Insertion: NA Placement: above the diaphragm valve at the width of two of the individual patient's vertebral bodies as visualised on a perioperative X-ray Duration: 24 or 48 hours (23 children had 48-hour measurement) Comparator: NA	Diagnostic yield (based on DeMeester score)
Hakanson et al. (2009) Department of Surgery, Center of Gastrointestinal Disease, Ersta Hospital, Stockholm, Sweden	Case-control study with in-subject simultaneous recording Level: II Quality: 14/26 C1 P2	53 volunteers and 55 patients (45 volunteers and 47 patients had enough data to be included) 30 females and 15 males in volunteer group 27 females and 20 males in patient group Mean age: Volunteers: 47 years (range 21–68) Patients: 50 years (range 23–69) Indication: Patients had symptoms including mainly heartburn, acid regurgitation, and chest or epigastric pain	Inclusion: Patients: referred to the clinic for investigation of GORD Exclusion: Volunteers: history of GORD or any other pathology of the upper GI tract, taking acid-suppressing agents or medication known to affect GI motility	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline, Medtronic, Skovlunde, Denmark Insertion: transnasally Placement: 5 cm proximal to the upper border of the LOS Duration: 24 hours	Accuracy: Concordance (acid exposure time) Comparative diagnostic yield
Karamanolis et al. (2012) 2nd Department of Internal Medicine, Propaedeutic	Diagnostic yield study Level: IV	32 patients 18 males, 14 females Mean age: 45.3 ± 12.5 years	Inclusion: Patients with at least 3 episodes of chest pain per week and normal oesophageal manometry (in order to	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ	Accuracy: Incremental diagnostic yield (of prolonged monitoring)

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Attikon University General Hospital, Athens, Greece		Indications: 3 episodes of non-cardiac chest pain per week	exclude the presence of any oesophageal motility disorder) Exclusion: Patients that were using aspirin or NSAIDs, a history of upper GI surgery, gastric or duodenal ulcer, connective tissue disease and severe liver, lung, renal or haematological disease	Duration: 48 hours Comparator: (24-hour Bravo pH monitoring)	
Kushnir, Sayuk & Gyawali (2011) Division of Gastroenterology, Washington University School of Medicine, St Louis, Missouri, USA	Comparative study of diagnostic yield Level: IV Quality: 13.5/26 C1 P2	462 patients (286 females and 176 males) in the wireless group 1,605 patients (1,092 females and 513 males) in the catheter group	Inclusion: Adult patients referred for ambulatory pH monitoring for all indications at the medical centre over a 7-year period Exclusion: Subjects who had undergone prior anti-reflux procedures or oesophageal surgery, incomplete studies with <14 hours of interpretable data, studies contaminated by artefact	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: NS Duration: 28 hours Comparator: System: NS (catheter-based) Insertion: NS Placement: NS Duration: 24 hours	Accuracy: Comparative diagnostic yield
Lacy, Chehade & Crowell (2011) Section of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA	Diagnostic yield study Level: IV Quality: Q1 CX (no comparator) P2	358 patients (257 females, 101 males) Mean age: 51 ± 14 years Predominantly (97%) Caucasian Indications: Symptoms thought secondary to GORD	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the GOJ or 5 cm above the upper border of the LOS Duration: 48 hours Comparator: NA	Accuracy: Diagnostic yield
Martinez de Haro et al. (2008) Servicio de Cirugia General, Hospital Universitario Virgen	Non-randomised controlled trial Level: III-2 Quality: 11/26	40 patients Group 2 standard pH monitoring: mean age 45 years (range 17–73) 9 males and 6 females	Inclusion: General criteria: >18 years of age, informed consent of the risks, benefits and alternative tests, committed to completing follow-up	Index test: System: Bravo™ wireless capsule system Insertion: transnasally (n=25) and transorally (n=10) Placement: 5 cm above the LOS	Accuracy: Concordance (acid exposure time)

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
de la Arrixaca, Murcia, Spain	C1 P2	Group 2 Bravo monitoring: mean age 37 years (range 21–59), 8 males and 7 females Group 3 (both): mean age 36.5 years (range 18–56), 7 males and 3 females	Control group: less than 2 episodes of pyrosis or regurgitation per month, no dysphagia or atypical reflux symptoms, no oesophageal motility disorders, not taken medication for acid reflux, no evidence of hiatal hernia or oesophagitis Patient group: typical reflux symptoms requiring treatment, inflammatory lesions due to reflux in the oesophageal mucosa in group 2. Exclusion: <18 years of age, oesophageal stenosis, oesophageal varix, lesion affecting the nostrils, severe oesophageal motor disorders, high anaesthetic risk, history of coagulopathy, haemorrhage, intake of anticoagulants or platelet antiaggregates, myocardial infarction or CVA in the past 6 months, pregnancy, history of radiotherapy in the thoracic region, digestive haemorrhage in the past 6 months, any known medical disorder that could alter the data, unable to accept study protocols	Duration: 48 hours Comparator: System: Digitrapper Mark III pH monitoring (Synetics, Stockholm, Sweden) Insertion: transnasally Placement: 5 cm above the LOS Duration: 24 hours	
Martinez de Haro et al. (2008) Servicio de Cirugia General, Hospital Universitario Virgen de la Arrixaca, Murcia, Spain	Diagnostic yield study Level: IV CX (no comparison) P2	40 patients Group 2 standard pH monitoring: mean age 45 years (range 17–73) 9 males and 6 females Group 2 Bravo monitoring: mean age 37 years (range 21–59),	Inclusion: General criteria: >18 years of age, informed consent of the risks, benefits and alternative tests, committed to completing follow-up Control group: less than 2 episodes of pyrosis or regurgitation per month, no dysphagia or atypical reflux symptoms,	Index test: System: Bravo™ wireless capsule system Insertion: transnasally (n=25) and transorally (n=10) Placement: 5 cm above the LOS Duration: 48 hours	Accuracy: Concordance (acid exposure time)

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
		8 males and 7 females Group 3 (both): mean age 36.5 years (range 18–56), 7 males and 3 females	no oesophageal motility disorders, not taken medication for acid reflux, no evidence of hiatal hernia or oesophagitis Patient group: typical reflux symptoms requiring treatment, inflammatory lesions due to reflux in the oesophageal mucosa in group 2. Exclusion: <18 years of age, oesophageal stenosis, oesophageal varix, lesion affecting the nostrils, severe oesophageal motor disorders, high anaesthetic risk, history of coagulopathy, haemorrhage, intake of anticoagulants or platelet antiaggregates, myocardial infarction or CVA in the past 6 months, pregnancy, history of radiotherapy in the thoracic region, digestive haemorrhage in the past 6 months, any known medical disorder that could alter the data, unable to accept study protocols		
Pandolfino et al. (2003) Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA	Diagnostic yield study Level: IV	44 controls (13 males, 31 females, age range 23–53 years) 41 patients (26 males, 15 females, age range 32–72 years)	Inclusion: Not stated Exclusion: Controls: abdominal symptoms, use of antacids or antisecretory medication, abnormal endoscopy Patients: history of surgical manipulation of the upper GI tract, history of bleeding diathesis or coagulopathy, stroke or transient ischemic attack in the past 6 months, significant medical illness, oesophageal	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ	Accuracy: Incremental diagnostic yield (of prolonged monitoring) Diagnostic yield

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
			varices		
Park et al. (2013) Divisions of Gastroenterology, Department of Internal Medicine, Seoul; St Mary's Hospital, the Catholic University of Korea, Seoul, Korea	Diagnostic yield study Level: IV	230 patients (101 males, 129 females) Mean age: 49.7 ± 12.4 years Indication: Suspected GORD	Inclusion: ≥18 years of age Exclusion: Previous oesophageal, gastric or duodenal surgery, GI organic disease, significant comorbidity, or GI bleeding	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to SCJ Duration: 48 hours	Accuracy: Diagnostic yield
Prakash & Clouse (2005) Division of Gastroenterology, Washington University, School of Medicine, St Louis, Missouri, USA	Diagnostic yield study Level: IV	157 patients (100 females, 57 males) Mean age: 47 ± 1 years 121 subjects were off therapy, 36 subjects were on therapy Indications: Typical reflux symptoms (55.4%) or atypical symptoms potentially attributable to GORD (44.6%)	Inclusion: Patients who had been referred for pH monitoring to evaluate a predefined, suspected reflux symptom and who recorded this symptom at least once during the 2-day monitoring period were identified Exclusion: Oesophagitis	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ	Accuracy: Incremental diagnostic yield (of prolonged monitoring)
Scarpulla et al. (2007) Gastroenterology Division, M. Raimondi Hospital, Can Cataldo, Italy	Study of diagnostic yield Level: IV	83 patients (complete 4-day recordings were available for 34/83) 44 females, 39 males Median age 42 years (range 18–63) Indications: Typical reflux symptoms or atypical symptoms including chest pain and extra oesophageal	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the Z line Duration: 24, 48, 72 and/or 96 hours	Accuracy: Diagnostic yield Incremental diagnostic yield

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
		complaints			
Schneider et al. (2007) Department of General, Visceral, and Transplant Surgery, University Hospital of Tubingen, Tubingen, Germany	Non-randomised controlled trial Level: III-2 Quality: 9/26 CX P3	123 patients Bravo group: 36 males and 32 females, mean age 51 years (range 34–67) Catheter group: 29 males and 26 females, mean age 43 years (range 32–59)	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 5 cm above the upper border of the LES Duration: 48 hours Comparator: System: not stated Insertion: transnasally Placement: not stated Duration: 24 hours	Accuracy: Concordance (acid exposure, number of reflux episodes)
Sweis et al. (2009) Functional GI Disease Unit, Clinic for Gastroenterology and Hepatology, University Hospital of Zurich, Zurich, Switzerland	Diagnostic yield study Level: IV	134 patients Mean age: 60 years (range 18–76) 58 males, 76 females Compared with 110 consecutive catheter pH controls Mean age: 57 years (range 16–85) 44 males, 66 females Indications Patients: Intolerance of catheter insertion, intolerance of catheter after intubation, or vomiting of catheter within 24 hours of the monitoring period	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 48 hours	Accuracy: Diagnostic yield
Sweis et al. (2011) Guy's and St	Diagnostic yield study	38 patients 13 males, 25 females	Inclusion: Not stated	Index test: System: Bravo™ wireless capsule system	Accuracy: Incremental diagnostic yield

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
Thomas Hospitals, London, UK	Level: IV	Median age: 41.6 years (range 17–75) Indication: Reflux symptoms and no diagnosis of GORD on catheter-based pH monitoring	Exclusion: Not stated	Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 96 hours Comparator: 24-, 48-, 72-hour Bravo pH monitoring	(of prolonged monitoring)
Ward et al. (2004) Division of Gastroenterology and Hepatology, Mayo Clinic, Jacksonville, Florida, USA	Diagnostic yield study Level: IV Quality: Q2 CX (no comparator) P2	60 patients (58 patients had enough data to be included) 26 males, 34 females Mean age 54 years (range 27–82) Indications: Document GORD before surgery, possible GORD with negative PPI trial, evaluate response to PPI, possible supra-oesophageal GORD, non-cardiac chest pain, evaluate response to Stretta procedure, previous failed pH test	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours	Accuracy Diagnostic yield
Wenner, Johnsson et al. (2007) Department of Surgery, Lund University Hospital, Lund, Sweden	Randomised crossover trial Level: II Quality: 20/26	35 patients (31 patients had enough data to be included) 15 males, 16 females Median age: 52 years Median BMI: 27 (range 24–28) Indications: Typical reflux symptoms (55%) and atypical reflux	Inclusion: Patients referred for oesophageal pH monitoring Exclusion: Patients with coagulopathy, severe oesophagitis/stricture, portal hypertension, a pacemaker, age <20 or >70 years of age, previous gastro-oesophageal surgery, severe cardiopulmonary disease, symptoms	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Duration: 48 hours Comparator: System: Slimline (Medtronic) Insertion: transnasally Placement: 5 cm above the upper border of	Accuracy: Concordance in acid exposure time Comparative diagnostic yield

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / Exclusion criteria	Diagnostic tests	Diagnostic accuracy outcome assessed
		symptoms (45%)	suggestive of oesophageal motor disorders	the LOS Duration: 24 hours	
Wenner et al. (2008); Wenner & Johansson et al. (2007) Department of Surgery, Lund University Hospital, Lund, Sweden	Comparative diagnostic yield study Level: III-2 Quality: 11.5/26 CX P3	70 patients, but 64 had enough data to be included (39 males, 25 females) 55 controls (27 males and 28 females). Not all patients are included in the distal capsule outcomes Indications: Patients with typical reflux symptoms	Inclusion: Patients: presence of typical reflux symptoms such as heartburn and regurgitation. Symptoms had to occur more than twice weekly. Treatment with medical acid suppression had to result in complete symptom relief or a more than 50% reduction of the symptoms Exclusion: Coagulopathy, severe oesophagitis/stricture, portal hypertension, pacemaker, <20 or >70 years of age, history of gastro-oesophageal surgery, severe cardiopulmonary disease, symptoms suggestive of oesophageal motor disorders	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm above the SCJ Reference standard (for sensitivity and specificity): Clinical diagnosis: symptoms suggestive of GORD and responding to acid suppression	Accuracy: Comparative diagnostic yield
Wong et al. (2005) The Neuro-Enteric Clinical Research Group, Section of Gastroenterology, Department of Medicine, Southern Arizona VA Health Care System and University of Arizona Health Sciences Center, Tucson, Arizona, USA	Randomised controlled trial Level: II Quality: 16.5/26	50 patients Mean age: 50.2 years (range 21–79) 26 males and 24 females Indication: Failure in controlling symptoms while on PPI therapy or for evaluation before anti-reflux surgery	Inclusion: Not stated Exclusion: Patients with a history of bleeding tendency or coagulopathy, significant concomitant medical co-morbidity, severe GI bleeding within the past 6 months, a history of upper GI surgery, oesophageal varices, pacemaker, implantable cardiac defibrillator, being unable to complete the 24-hour pH monitoring, unable to report daily activity	Index test: System: Bravo™ wireless capsule system Insertion: transnasally or transorally Placement: 5 cm above the upper border of the LOS Duration: 24 hours Comparator: System: Digitrapper Mark III (Medtronic) Insertion: transnasally Placement: 5 cm above the proximal margin of the LOS Duration: 24 hours	Accuracy Concordance (acid exposure time)

Change in management

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / exclusion criteria	Diagnostic tests	Change in management outcome assessed
Gawron et al. (2012) Division of Gastroenterology and Hepatology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	The final sample consisted of 90 patients with a negative test, 66 Bravo monitoring and 24 impedance (75.6% female) Mean age: 48.4 years	Inclusion: Patients who had undergone Bravo monitoring from January 2006 to January 2010 while withholding PPI therapy, had total acid exposure <5% and a negative symptom index on both days Exclusion: Being on PPI therapy at the time of the test, prior surgery, Barrett's oesophagus, oesinophilic oesophagitis	Index test: System: Bravo™ wireless capsule system Insertion: NS Placement: NS Duration: 48 hours Comparator: NA	Management: % of patients who use PPIs
Grigolon et al. (2011) Department of Gastroenterology, Universita Degli Studi and Fondazione IRCCS Ca Granada, Ospedale Maggiore Policlinico, Milan, Italy	Matched-pairs retrospective cohort study Level: III-2 Quality: 16.5/26 CX P2	102 patients Catheter group: 51 patients, 20 males, mean age 48 years (range 22–78) Bravo group: 51 patients, 22 males, mean age 48 years (range 23–82)	Inclusion: All consecutive patients who underwent wireless 96-hour pH monitoring between January and December 2007 off PPI therapy, and patients undergoing traditional 24-hour pH monitoring in the same period, after matching for the most troublesome symptom potentially related to GORD, sex and age Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: NA Placement: 6 cm above the SCJ Duration: 48–96 hours Comparator: System: Slimline mod 901P3001, Medtronic, Denmark Insertion: transnasally Placement: 6 cm above the point of sudden change in pH from acid to neutral as detected by means of continuous rapid withdrawal from the stomach, or 5 cm above the upper margin of the LOS Duration: 24 hours	Management: Concordance between results of the test and treatment
Lacy et al. (2009) Division of Gastroenterology and Hepatology, Dartmouth-	Prospective case series Level: IV	51 patients (50 patients had enough data to be included) 24 males, 26 females Mean age: 13 years (range 5–17)	Inclusion: Not stated Exclusion: Prior surgery to the oesophagus or stomach, known intestinal obstruction	Index test: System: Bravo™ wireless capsule system Insertion: not stated Placement: 6 cm above the SCJ	Management: Change in management (%)

Study setting	Study design / Quality appraisal	Study population	Inclusion criteria / exclusion criteria	Diagnostic tests	Change in management outcome assessed
Hitchcock Medical Center, Lebanon, New Hampshire, USA	Quality: Q1 CX (no comparator) P2	Indications: Heartburn, regurgitation, abdominal pain, nausea and/or vomiting, chest pain	or oesophageal varices, achalasia, inability to understand the procedure, inability to undergo oesophageal manometry or sedation	Duration: 48 hours Comparator: NA	
Lacy, Chehade & Crowell (2011) Division of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA	Prospective case series Level: IV Quality: Q2 CX (no comparator) P2	309 patients Mean age: 48 ± 15 years 107 males, 202 females Indications: Symptoms of acid reflux, evaluation of chest pain, preoperative testing, chronic cough, ENT symptoms	Inclusion: Adult patients (>18 years of age) referred for wireless pH-monitoring from October 2005 to December 2006 Exclusion: Women who were pregnant, history of major facial or nasal trauma, presence of any serious medical condition that would prevent safe deployment of the wireless capsule	Index test: System: Bravo™ wireless capsule system Insertion: transorally or transnasally Placement: 6 cm above the SCJ (orally) or 5 cm above the upper border of the LOS (nasally) Duration: 48 hours Comparator: NA	Management: Change in diagnosis (%) Change in management (%)
Sweis et al. (2011) Guy's and St Thomas Hospitals, London, UK	Case series Level: IV Quality: Q2 CX (no comparator) P2	38 patients 13 males, 25 females Median age: 41.6 years (range 17–75) Indication: Reflux symptoms and no diagnosis of GORD on catheter-based pH monitoring	Inclusion: Not stated Exclusion: Not stated	Index test: System: Bravo™ wireless capsule system Insertion: transorally Placement: 6 cm proximal to the SCJ Duration: 96 hours Comparator: NA	Management Change in treatment

Does management change health outcomes?

True positives

Study setting	Study design / Quality appraisal	Study population / Included studies	Inclusion criteria / exclusion criteria	Intervention and comparator	Health outcomes assessed
Allgood & Bachmann (2000) Department of Social Medicine,	Systematic review Level: III-2 Quality: 14/27	Six randomised controlled trials and three cohort studies that compared medical and surgical	Inclusion: All controlled trials of medical and surgical management of objectively diagnosed chronic GORD.	Intervention: Surgical treatment for GORD Comparator:	Time to treatment failure Mean lower oesophageal pressure Mean DeMeester acid reflux score

Study setting	Study design / Quality appraisal	Study population / Included studies	Inclusion criteria / exclusion criteria	Intervention and comparator	Health outcomes assessed
MRC Health Services Research Collaboration, University of Bristol, Bristol, UK		management of GORD were included, study population ranged from 15 to 155 patients per arm, 7/9 studies had fewer than 100 patients per arm	Randomised trials were regarded as being most relevant, but non-randomised trials were also included, with an emphasis on assessing potential bias. There were no age or sex restrictions Exclusion: Not stated	Medical treatment for GORD	Reduction in grade of endoscopic oesophagitis Mean % time pH <4 / absence of reflux Decrease in pulmonary medication Development of atrophic gastritis
Galmiche et al. (2011) Department of Gastroenterology and Hepatology, Nantes University, Nantes, France	Randomised controlled trial Level: II Quality: 21/26	554 patients 288 randomised to receive surgery (199 males, mean age 45 years), 248 underwent surgery, 180 completed 5-year follow-up 266 randomised to esomeprazole (199 males, mean age 45 years), 192 completed 5-year follow-up	Inclusion: Patients aged 18–70 years with chronic symptomatic GORD, diagnosed on the basis of typical clinical history and presence of oesophageal mucosal breaks at endoscopy and/or pH monitoring. All patients had to be eligible for both treatments Exclusion: Not stated	Intervention: Standard laparoscopic anti-reflux surgery Comparator: Esomeprazole, 20–40 mg/d, allowing for dose adjustments	Time-to-treatment failure Acid regurgitation Post-operative symptoms Endoscopic findings Health-related quality of life Safety
Grant, Cotton et al. (2013) Health Services Research Unit, University of Aberdeen, Aberdeen, UK	Randomised controlled trial Level: II Quality: 17/26	810 original participants, 357 were randomised 178 patients allocated to surgery, 111 received surgery, 127 had 5-year follow-up 179 patients allocated to medication, 10 received surgery, 119 had 5-year follow-up	Inclusion: Patients had had >12 months of maintenance treatment with PPIs (or alternative), reasonable control of GORD symptoms, evidence of GORD (endoscopy or pH monitoring), were suitable for both treatments and the recruiting doctor was uncertain which management policy to follow Exclusion: Not stated	Intervention: Anti-reflux fundoplication (type of surgery was chosen by the surgeon) Comparator: Optimised continued medical management (reviewed and adjusted as judged best by local gastroenterologist)	Health-related quality of life (REFLUX) Health status (SF-36, EuroQol) Use of anti-reflux drugs Surgical complications Individual symptoms
Wileman et al. (2010) Cochrane	Systematic review Level: I	Four randomised controlled trials were included. Two studies were performed in	Inclusion: All randomised controlled trials and quasi-randomised controlled trials	Intervention: Surgery: laparoscopic fundoplication surgery (all	Health-related quality of life GORD-specific quality of life

Study setting	Study design / Quality appraisal	Study population / Included studies	Inclusion criteria / exclusion criteria	Intervention and comparator	Health outcomes assessed
Collaboration, the Cochrane Library	Quality: 23/27	multiple studies in the UK, one study was conducted in 11 European countries, and one study was conducted in a centre in Canada. Sample sizes ranged from 104 to 554 participants, with a total of 1,232 randomised patients	comparing medical management with laparoscopic fundoplication surgery Exclusion: Non-randomised studies	types) <u>Comparator:</u> Medical management: either PPIs or histamine receptor antagonists (H2RAs)	Heartburn Regurgitation (acid reflux) Dysphagia Other GORD-related symptoms Conversion to open surgery Intraoperative complications Postoperative complications Admission to intensive treatment unit / high-dependency unit Reoperation Mortality during index admission Medication use Length of operation Length of hospital stay Time off work Visits to GP due to GORD Visits to hospital due to GORD pH monitoring (% time pH <4 / DeMeester score) Manometry (LOS pressure) Endoscopy Other complications or adverse events

False positives

Study setting	Study design	Study population / Included studies	Inclusion criteria / exclusion criteria	'True diagnosis' and misdiagnosis	Health outcomes assessed
Auvin et al. (2012) Insitut National de la Santé et de la Recherche Medicale, Paris, France	Retrospective case series	83 infants diagnosed with infantile spasms were included (19 in Paris, 20 in Naples, 44 in Baltimore) Median age of onset symptoms: 6 months 28 patients (34%) were diagnosed more than 30 days after initial presentation	Inclusion: Not stated Exclusion: Patients for whom data on diagnostic delay were unavailable	True diagnosis: West syndrome or infantile spasms Misdiagnosis: No diagnosis (behavioural symptoms, not organic): 301/362 physicians, 83.3% GORD: 7% Constipation: 7% Colitis: 3%	Seizures at age 2 years Psychomotor development at age 2 years
Napuri et al. (2010) Department of Child and Adolescent Medicine, Rennes, France	Retrospective case series	156 infants diagnosed with infantile spasms were included 87 male (55%), 69 female (45%) 45 (29%) symptomatic spasms Median age onset: 20 weeks, mean age onset: 22.4 weeks (SD 13.3 weeks) Time lag from first symptom to diagnosis ranged from a few days to 44 weeks with a peak at 4 weeks	Inclusion: A combination of clusters of spasms, altered psychomotor development and paroxysmal EEG activity, as defined by the International League Against Epilepsy Exclusion: Not stated	True diagnosis: Infantile spasms Misdiagnosis: Epilepsy: 96 (62%) GORD: 26 (17%) Healthy: 29 (18%) Miscellaneous: 5 (3%)	Response to treatment

NA = not available; SCJ = squamo-columnar junction; LOS = lower oesophageal sphincter; GORD = gastro-oesophageal reflux disease; GI = gastrointestinal; PPIs = proton pump inhibitors;

GOJ = gastro-oesophageal junction; RDQ = Reflux Disease Questionnaire; NS = not significant

Appendix D   Excluded studies

Conference abstracts

Ang, D, Ang, TL, Lim, KB, Kiong, TE & Fock, KM 2009, 'To Bravo or not? wireless pH monitoring versus conventional nasopharyngeal pH catheter monitoring in the evaluation of gastroesophageal reflux disease in a multiracial Asian cohort', *Gastroenterology*, vol. 136, no. 5, pp. A295–A296.

Ang, D, Ang, TL, Lim, KB, Teo, EK & Fock, KM 2010, 'Efficacy and safety of the wireless capsule pH monitoring system in the evaluation of gastroesophageal reflux disease (GERD): results from a multiracial Asian cohort', *Gastroenterology*, vol. 138, no. 5, p. S494.

Ang, D, Ang, TL, Xu, Y, Poh, CH, Law, NM, Tan, M & Fock, KM 2012, 'Wireless oesophageal pH monitoring: feasibility, safety and establishing values in a multiracial cohort of asymptomatic Asian subjects', *Journal of Gastroenterology and Hepatology*, vol. 27, pp. 286–287.

Bradley, AG, Crowell, MD, DiBaise, JK, Kim, HJ, Sharma, VK & Fleischer, DE 2009, 'Patient characteristics predict variation in tolerance of wireless pH-metry', *Gastroenterology*, vol. 136, no. 5, pp. A525–A526.

Chander, B, Hanley-Williams, N, Deng, Y & Sheth, AA 2011, '48-hour BRAVO (wireless) esophageal pH monitoring improves symptom correlation as compared to 24-hour studies alone', *Gastroenterology*, vol. 140, no. 5, p. S251.

Chehade, RJ, Lacy, BE & Crowell, MD 2010, 'Gastroesophageal reflux disease: a prospective study comparing a symptom-based reflux disease questionnaire to BRAVO pH capsule monitoring', *Gastroenterology*, vol. 138, no. 5, p. S493.

Crowell, M, Lacy, B & DiBaise, J 2010, 'Intrasession variability and yield of wireless pH-metry following sedated and unsedated capsule placement', *American Journal of Gastroenterology*, vol. 105, p. S17.

Daboul, I & Earla, JRP 2005, 'Comparison of the wireless (Bravo) pH-monitoring device monitoring parameters with the traditional 24 hr pH monitoring in the evaluation of gastroesophageal reflux disease', *American Journal of Gastroenterology*, vol. 100, no. 9, pp. S42.

Daboul, I & Ghai, V 2006, 'Validity of the DeMeester's score for the evaluation of GERD with 24 hr conventional pH monitor and the wireless Bravo capsule', *American Journal of Gastroenterology*, vol. 101, no. 9, pp. S72.

Dutton, TJ, Owais, D, Whybrow, J, McGhee, R & Wajed, S 2011, 'Six year experience of the Bravo (TM) pH-monitoring device for detecting gastro-oesophageal reflux requiring surgery', *British Journal of Surgery*, vol. 98, pp. 182.

Finn, R, Lacy, B & Crowell, M 2011, 'Wireless pH capsule testing: the effects of diabetes and narcotic use and the safety of clopidogrel and warfarin during testing', *American Journal of Gastroenterology*, vol. 106, pp. S30–S31.

Fox, M 2012, 'Prolonged wireless pH-monitoring', *Diseases of the Esophagus*, vol. 25, p. 7A.

Fox, MR, Sweis, R, Menne, D, Anggiansah, A & Wong, T 2011, 'Increasing reflux study duration progressively improves diagnostic consistency: a prospective study with 96-hr wireless pH recordings', *Gastroenterology*, vol. 140, no. 5, pp. S246–S247.

Gawron, A, Fareeduddin, A, Fought, A, Toto, E, Boris, L, Kahrilas, PJ & Pandolfino, JE 2011, 'Proton pump inhibitor (PPI) utilization in patients after negative BRAVO (TM) pH testing for GERD; do they discontinue therapy?', *Gastroenterology*, vol. 140, no. 5, p. S390.

Gillies, RS, Stratford, JM, Jones, L, Sohanpal, J, Booth, MI & Dehn, TCB 2004, 'Does a 48-hour Bravo pH study facilitate the diagnosis of gastro-oesophageal reflux disease', *Gut*, vol. 53, pp. A30.

Gomez, J, Bohm, M & Richter, J 2009, 'Refractory gastroesophageal reflux disease (RGERD) patients with normal upper endoscopy (EGD) should have pH testing off proton pump inhibitors (PPIs)', *Gastroenterology*, vol. 136, no. 5, p. A85.

Grigolon, A, Consonni, D, Bravi, I, Tenca, A & Penagini, R 2010, '96 h wireless vs 24 h traditional pH monitoring: usefulness in patients' management', *Digestive and Liver Disease*, vol. 42, pp. S103–S104.

Guardino, JM, Ours, T & Richter, J 2002, 'Utilizing the Bravo (TM) esophageal pH-monitoring system to compare pH profiles in GERD patients and healthy adult volunteers', *Gastroenterology*, vol. 122, no. 4, pp. A338.

Guerriero, C, Cuofano, S & Spinelli, P 2012, 'Impact of wireless prolonged ph studies on diagnosis of gastroesophageal reflux: retrospective study', *Digestive and Liver Disease*, vol. 44, p. S219.

Jung, KW, Francis, DL, Romero, Y & Murray, JA 2010, 'The comparison of the discomfort between catheter and catheterless pH test', *Gastroenterology*, vol. 1, p. S492.

Korrapati, V, Bonasera, RJ, Balani, AR, Grendell, JH & Kongara, KR 2009, 'Does deep sedation impact the results of 48-hour catheter-less pH testing?', *Gastroenterology*, vol. 136, no. 5, p. A291.

Kushnir, V, Sayuk, GS & Gyawali, CP 2009, 'Improving diagnostic utility of ambulatory pH testing', *Gastroenterology*, vol. 136, no. 5, p. A84.

Lee, SY, Park, YD, Yoon, SJ, Chung, YJ, Jeon, SW, Cho, CM, Tak, WY, Kweon, YO, Kim, SR, Choi, YH & Kim, JY 2007, 'Wireless esophageal pH monitoring: is it really tolerable?', *Neurogastroenterology and Motility*, vol. 19, pp. 65–66.

Lin, E, Waring, JP, Ramaswamy, A, Swafford, V, Tucker, A, Khaitan, L, Ramshaw, BJ & Smith, CD 2003, 'Analysis of 245 consecutive studies using 48-hour wireless pH (BRAVO) probes for GERD evaluation', *Gastrointestinal Endoscopy*, vol. 57, no. 5, pp. AB100.

Macdonald, J, Porter-Ritchie, V, Mackie, B & McNamara, D 2011, 'Assessment of tolerance, safety and efficacy of ambulatory oesophageal pH monitoring using the catheter-free Bravo system in a scottish population', *Gut*, vol. 60, pp. A182-A183.

McCann, V & Nompleggi, D 2012, 'Chest discomfort with catheter-free ambulatory pH monitoring can be predicted by response to PPIs', *American Journal of Gastroenterology*, vol. 107, p. S27.

Morley-Fletcher, A, Souza, ALL, Nurko, S & Rodriguez, L 2009, 'Experience with BRAVO wireless pH study in children', *Gastroenterology*, vol. 1, p. A513.

Rao, N, Campbell, DI, Narula, P, Taylor, C, Thomson, M & Rao, P 2012, 'Two year experience of a catheter-free, radio telemetric, oesophageal pH-monitoring (Bravo (TM)) system in children at a single UK tertiary centre', *Archives of Disease in Childhood*, vol. 97, p. A56.

Rodriguez, L, Rosengaus, L, Morley-Fletcher, A & Nurko, S 2011, 'Does general anesthesia change the amount of gastroesophageal reflux in children? A study using BRAVO wireless ph capsule', *Gastroenterology*, vol. 140, no. 5, p. S748.

Seren, S, Bayraktar, UD & Ehrinpreis, MN 2007, 'Esophageal pH monitoring with Bravo (TM) capsule in patients with suspected gastroesophageal reflux disease', *American Journal of Gastroenterology*, vol. 102, pp. S130.

Seza, A, Odaka, T, Kobayashi, M, Yamaguchi, T, Genn, S, Suzuki, T & Saisho, H 2006, 'Esophageal and intragastric pH monitoring using a wireless system (BRAVO) to evaluate esophageal acid exposure and gastric acid secretion', *Gastroenterology*, vol. 130, no. 4, pp. A166.

Sifrim, D 2012, '24-h esophageal pH-monitoring: why do we need more?', *Diseases of the Esophagus*, vol. 25, p. 7A.

Souza, ALL, Morley-Fletcher, A, Nurko, S & Rodriguez, L 2009, 'BRAVO wireless pH in children: is there an effect of anesthesia?', *Gastroenterology*, vol. 136, no. 5, p. A510.

Strohl, A, Khatri, N, Palese, C, Evans, S & Benjamin, S 2010, 'Is 48-hr Bravo pH monitoring more efficacious than 24-hr ph monitoring for patients with refractory GERD undergoing evaluation for Nissen fundoplication?', *American Journal of Gastroenterology*, vol. 105, p. S16.

Sweis, R, Anggiansah, A, Anggiansah, R, Fox, MR & Wong, T 2009, '96 hour BRAVO increases yield in patients with negative 24-hour esophageal catheter pH test results', *Gastroenterology*, vol. 1, p. A286.

Sweis, RR, Anggiansah, A, Anggiansah, R, Fox, M & Wong, T 2009, '96 hour Bravo increases yield in patients with negative 24-hour oesophageal catheter ph tests results', *Gut*, vol. 58, p. A33.

Xu, B, Yuan, Y & Zhu, Q 2009, '48-hour wireless esophageal pH monitoring increases sensitivity in detecting abnormal esophageal acid exposure', *Digestion*, vol. 1, p. 71.

Incorrect/irrelevant intervention

Caparello, C, Bravi, I, Cantu, P, Grigolon, A, Tenca, A, Mauro, A & Penagini, R 2012, 'Traditional vs wireless intragastric pH monitoring: are the two techniques comparable?', *Neurogastroenterology and Motility*, vol. 24, no. 10, pp. 44– 49.

Li, JN, Liu, CL & Tao, XH 2013, 'Clinical utility and tolerability of JSPH-1 wireless esophageal pH-monitoring system', *BMC Gastroenterology*, vol. 13, p. 10.

Yamaguchi, T, Seza, A, Odaka, T, Shishido, T, Ai, M, Gen, S, Kouzu, T & Saisho, H 2006, 'Placement of the Bravo wireless pH-monitoring capsule onto the gastric wall under endoscopic guidance', *Gastrointestinal Endoscopy*, vol. 63, no. 7, pp. 1046–1050.

Irrelevant outcomes

Gonzalez, JO & Barkin, JS 2005, 'Endoscopic visualization of deployment of the Bravo pH System to prevent malplacement', *Gastrointestinal Endoscopy*, vol. 62, no. 1, pp. 178–180.

Unable to retrieve full text

Armas, H & Cabrera, R 2005, 'Esophageal pH monitoring: indications and techniques', *PHmetría esofágica: indicaciones y técnica*, *An Pediatr Contin*, vol. 3, no. 4, pp. 253–256.

Becmeur, F, Rebeuh, J & Molinaro, F 2009, 'Gastroesophageal reflux in children', [*Le reflux gastro-œsophagien chez l'enfant*], *Médecine et Thérapeutique Pédiatrie*, vol. 12, no. 1, pp. 45–51.

Hagemann, D & Schmitz, F 2006, 'Esophageal pH monitoring with the BRAVO capsule: safety, tolerability and variability of the reflux pattern', [*Ösophagus-pH-metrie mit der BRAVO-Kapsel: Sicherheit, verträglichkeit und variabilität der refluxmuster*], *Der Gastroenterologe*, vol. 1, no. 1, pp. 56–59.

Hellmig, S & Fölsch, UR 2007, 'Gastroesophageal reflux disease: diagnosis and treatment of an endemic disease', [*Gastroösophageale refluxerkrankung: diagnostik und therapie einer volkskrankheit*], *Medizinische Klinik*, vol. 102, no. 5, pp. 373–380.

Labenz, J & Borkenstein, DP 2003, 'Pathophysiology and diagnosis of gastroesophageal reflux disease', [*Pathophysiologie und diagnostik der refluxkrankheit*], *Internist*, vol. 44, no. 1, pp. 11–19.

Magistà, AM, Lionetti, E, Boscarelli, G, Chiarelli, T, Latela, S, Rigillo, N, Miniello, VL & Francavilla, R 2003, 'Instrumental diagnosis of gastroesophageal reflux', [*Diagnostica*

strumentale del reflusso gastro-esofageo], *Otorinolaringologia Pediatrica*, vol. 13-14, pp. 26–33.

Moreno, JP, Martínez, C, Del Mar Tolín Hernani, M & Calatayud, GÁ 2013, 'Diagnosis and treatment of gastroesophageal reflux', [*Diagnóstico y tratamiento del reflujo gastroesofágico*], *Publicado en An Pediatr Contin*, vol. 11, no. 1, pp. 1–10.

Not in English, and not of a higher level of evidence than the English language literature

des Varannes, BS, Roman, S, Mion, F, Letard, JC & Zerbib, F 2012, 'Wireless pH-monitoring: technical aspects and clinical contribution', *Hepato-Gastro*, vol. 19, no. 1, pp. 21–27.

Carmona-Sánchez, R & Solana-Sentíes, S 2004, '[Efficacy, diagnostic utility and tolerance of intraesophageal pH ambulatory determination with wireless pH-testing monitoring system]', *Revista de Gastroenterología de México*, vol. 69, no. 2, pp. 69–75, <http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/916/CN-00527916/frame.html>.

[No authors listed] 2007, '[Clinical guidelines of diagnosis and treatment of gastroesophageal reflux disease]', *Revista de Gastroenterología de México*, vol. 72, no. 2, pp. 163–176.

Dantas, RO 2008, '[Electrode placement for esophageal pH recording]', *Arquivos de Gastroenterologia*, vol. 45, no. 4, pp. 259–260.

Gonzalez Martinez, MA 2004, '[Ambulatory esophageal pH-monitoring with wireless system]', *Revista de Gastroenterología de México*, vol. 69, no. 2, pp. 111–113.

Marchese, M, Spera, G, Spada, C & Costamagna, G 2004, 'Wireless pH-monitoring. [Italian] Capsula per studio pH-metrico', *Giornale Italiano di Endoscopia Digestiva*, vol. 27, no. 2, pp. 129–132.

Mönkemüller, K, Neumann, H, Fry, LC, Kolfenbach, S & Malfertheiner, P 2009, '[Catheter-free pH-metry using the Bravo capsule versus standard pH-metry in patients with non-erosive reflux disease (NERD)]' *Zeitschrift für Gastroenterologie*, vol. 47, no. 4, pp. 351–356, DOI 10.1055/s-2008-1027945, <http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/709/CN-00699709/frame.html>.

Valdovinos Diaz, MA, Remes Troche, JM, Ruiz Aguilar, JC, Schmulson, MJ & Valdovinos-Andraca, F 2004, 'Successful esophageal pH monitoring with Bravo capsule in patients with gastroesophageal reflux disease. [Spanish] ['*Medicion exitosa del pH esofagico con la capsula Bravo en pacientes con enfermedad por reflujo gastroesofagico*'], *Revista de Gastroenterología de México*, vol. 69, no. 2, pp. 62–68.

Irrelevant or incorrect study design

Bittinger, M & Messmann, H 2007, 'Gastroesophageal reflux disease and Barrett's esophagus', *Endoscopy*, vol. 39, no. 2, pp. 118–123.

Chotiprashidi, P, Liu, J, Carpenter, S, Chuttani, R, DiSario, J, Hussain, N, Somogyi, L & Petersen, BT 2005, 'ASGE Technology Status Evaluation Report: wireless esophageal pH-monitoring system', *Gastrointestinal Endoscopy*, vol. 62, no. 4, pp. 485–487.

Dranove, JE 2008, 'Focus on diagnosis: new technologies for the diagnosis of gastroesophageal reflux disease', *Pediatrics in Review*, vol. 29, no. 9, pp. 317–320.

Fox, M 2006, 'Bravo wireless versus catheter pH-monitoring systems', *Gut*, vol. 55, no. 3, pp. 434–435.

Federação Brasileira de Gastroenterologia, Sociedade Brasileira de Endoscopia Digestiva, Colégio Brasileiro de Cirurgia Digestiva & Sociedade Brasileira de Pneumologia e Tisiologia. 'Gastroesophageal reflux disease: diagnosis', 2011, *Revista da Associacao Medica Brasileira*, vol. 57, no. 5, pp. 489–497.

Grigolon, A, Bravi, I, Duca, P, Pugliese, D & Penagini, R 2010, 'Prolonged wireless pH monitoring: importance of how to analyse oesophageal acid exposure', *Scandinavian Journal of Gastroenterology*, vol. 45, no. 9, pp. 1133–1134.

Herbella, FA, Nipominick, I & Patti, MG 2009, 'From sponges to capsules: the history of esophageal pH monitoring', *Diseases of the Esophagus*, vol. 22, no. 2, pp. 99–103.

Hess, T, Hahn, EG & Konturek, PC 2009, 'BRAVO wireless 48-hour pH-metric system', *Medical Science Monitor*, vol. 15, no. 1, pp. RA17–RA21.

Holloway, RH 2005, 'Capsule pH monitoring: is wireless more?', *Gut*, vol. 54, no. 12, pp. 1672–1673.

Horton, BE, Schweitzer, S, DeRouin, AJ & Ong, KG 2011, 'A varactor-based, inductively coupled wireless pH sensor', *IEEE Sensors Journal*, vol. 11, no. 4, pp. 1061–1066.

Javors, BR & Bukman, R 2011, 'Radiographic demonstration of a wireless capsule pH monitor in the esophagus', *Dysphagia*, vol. 26, no. 2, pp. 193–195.

Kahrilas, PJ & Pandolfino, JE 2005, 'Review article: oesophageal pH monitoring--technologies, interpretation and correlation with clinical outcomes', *Alimentary Pharmacology & Therapeutics*, vol. 22, Suppl 3, pp. 2–9.

Lee, JS 2012, 'Is wireless capsule pH monitoring better than catheter systems?', *Journal of Neurogastroenterology and Motility*, vol. 18, no. 2, pp. 117–119.

Malfertheiner, P, Asaka, M, Freston, JW 2006, 'Looking to the future: unmet needs in understanding and managing upper GI diseases', *Drugs*, vol. 66, pp. 1–33.

Mainie, I, Tutuian, R & Castell, DO 2006, 'The limitations of pH monitoring for detecting gastroesophageal reflux', *Clinical Gastroenterology and Hepatology*, vol. 4, no. 9, p. 1184.

Perez-de-la-Serna, J, Ruiz-de-Leon-San-Juan, A & Sevilla-Mantilla, C 2010, 'Proximal esophageal pH-monitoring', *Revista Española de Enfermedades Digestivas*, vol. 102, no. 7, pp. 403–405.

Richter, JE 2009, 'Optimal utilization of esophageal pH monitoring in GERD patients', *Gastroenterology and Hepatology*, vol. 5, no. 4, pp. 249–251.

Blue Cross and Blue Shield Association Technology Evaluation Center (TEC) 2006, 'Special report: wireless esophageal pH monitoring', Executive summary, *TEC Assessments*, vol. 21, no. 2, pp. 1–2.

Vaezi, MF 2006, 'Should we Bravo?', *Gastroenterology*, vol. 130, no. 7, pp. 2238–2239.

Williams, C, Neshev, E, Heitman, S, Storr, M, Cole, M & Andrews, C 2009, 'A prospective, randomized trial comparing non-endoscopic, unsedated Bravo (TM) wireless pH capsule insertion versus conventional catheter-based pH monitoring', *Canadian Journal of Gastroenterology*, vol. 23, Supplement SA, Poster 137.

Did not meet inclusion criteria regarding population

Ang, D, Xu, Y, Ang, TL, Law, NM, Poh, CH, Teo, EK & Fock, KM 2013, 'Wireless oesophageal pH monitoring: Establishing values in a multiracial cohort of asymptomatic Asian subjects', *Digestive and Liver Disease*, vol. 45, no. 5, pp. 371–376.

Friedman, M & Schalch, P 2007, 'Upper esophageal pH monitoring: transnasal placement of the Bravo wireless capsule system', *Operative Techniques in Otolaryngology–Head & Neck Surgery*, vol. 18, no. 4, pp. 324–328.

Pandolfino, JE, Schreiner, MA, Lee, TJ, Zhang, Q, Boniquit, C & Kahrilas, PJ 2005, 'Comparison of the Bravo (TM) wireless and Digitrapper (TM) catheter-based pH-monitoring systems for measuring esophageal acid exposure', *American Journal of Gastroenterology*, vol. 100, no. 7, pp. 1466–1476.

Pandolfino, JE, Zhang, Q, Ghosh, SK, Roth, MP, Boniquit, C & Kahrilas, PJ 2007, 'Does the Bravo (TM) pH capsule affect esophageal motor function?', *Diseases of the Esophagus*, vol. 20, no. 5, pp. 406–410.

Pandolfino, JE, Zhang, Q, Schreiner, MA, Ghosh, S, Roth, MP & Kahrilas, PJ 2005, 'Acid reflux event detection using the Bravo wireless versus the Slimline catheter pH systems: why are the numbers so different?', *Gut*, vol. 54, no. 12, pp. 1687–1692.

Appendix E   Safety data

Table 70   Complications and adverse events of catheter-free monitoring compared with catheter-based monitoring in adults

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	p-value	Additional notes
Wenner, Johnsson et al. (2007)	Level II Quality: 20/26 CX P2 Randomised crossover trial	31 participants catheter group 31 participants wireless group	Patients' experience, median cm VAS (IQR) Nose/throat symptoms: 0.2 (0.0–1.9) Swallow-induced symptoms: 2.1 (0.5–5.0) Chest symptoms: 2.4 (0.3–5.9) All adverse symptoms: 2.1 (0.5–4.6) Interference work activities: 0.3 (0.0–1.4) Interference physical activities: 0.6 (0.2–2.7) Interference food intake: 1.4 (0.2–4.9) Interference normal daily life: 0.7 (0.2–3.4)	Patients' experience, median cm VAS (IQR) Nose/throat symptoms: 6.5 (1.5–8.0) Swallow-induced symptoms: 5.2 (2.1–6.7) Chest symptoms: 1.1 (0.3–2.9) All adverse symptoms: 5.1 (2.0–6.6) Interference work activities: 6.8 (1.7–9.0) Interference physical activities: 5.0 (2.6–8.5) Interference food intake: 4.0 (1.8–7.0) Interference normal daily life: 5.7 (2.3–8.0)	- <0.0001 0.033 0.084 <0.001 0.005 <0.0001 0.056 <0.0001	-
Andrews et al. (2012)	Level II Quality: 19/26 CX P2 Dual centre, randomised, non-blinded trial	43 participants catheter group 43 participants wireless group	pH placement discomfort (mm VAS ± SE) Nasal: 6 ± 2 Throat: 32 ± 4 Chest: 14 ± 3 Overall: 29 ± 4 - pH-test discomfort (mm VAS ± SE) Nasal: 10 ± 3 Throat: 19 ± 4 Chest: 29 ± 4 Overall: 26 ± 4 Eating/drinking: 75 ± 5 Ability usual activities: 92 ± 2	pH placement discomfort (mm VAS ± SE) Nasal: 36 ± 4 Throat: 37 ± 3 Chest: 13 ± 3 Overall: 33 ± 4 - pH-test discomfort (mm VAS ± SE) Nasal: 39 ± 3 Throat: 43 ± 4 Chest: 14 ± 3 Overall: 39 ± 4 Eating/drinking: 51 ± 4 Ability usual activities: 75 ± 5	- <0.001 0.317 0.968 0.406 - - <0.001 <0.001 0.001 0.012 <0.001 <0.001	2 patients were unable to tolerate the nasal (standard) pH catheter and 1 patient did not tolerate peroral insertion of the capsule due to excessive gagging and anxiety
Wong et al. (2005)	Level II Quality: 16.5/26 CX P2 Randomised	25 participants catheter group 25 participants wireless group	Nose pain: 8 (32%) Runny nose: 13 (52%) Throat pain: 4 (16%) Throat discomfort: 12 (48%) Cough: 5 (20%)	Nose pain: 15 (60%) Runny nose: 24 (96%) Throat pain: 12 (48%) Throat discomfort: 23 (92%) Cough: 7 (28%)	0.047 0.001 0.032 0.001 0.508	-

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	p-value	Additional notes
	controlled trial		Chest discomfort: 9 (36%) Headache: 5 (20%) Nose bleeding: 1 (4%)	Chest discomfort: 2 (8%) Headache: 14 (56%) Nose bleeding: 0	0.037 0.009 1.0	
Grigolon et al. (2007)	Level III-1 Quality: 13.5/26 CX P2 Pseudo-randomised controlled trial	78 participants catheter group 55 participants wireless group	Placement discomfort (mm VAS ± SE) 29 ± 4 Discomfort during whole test (mm VAS ± SE) 22 ± 3 Limitation of food intake (score 0–3) 0.4 ± 0.1 Limitation of daily activities (score 0–3) 0.2 ± 0.1 None to mild discomfort: 41/55 (74.5%) Chest discomfort: 17/55 (30.9%)	Placement discomfort (mm VAS ± SE) 32 ± 3 Discomfort during whole test (mm VAS ± SE) 37 ± 3 Limitation of food intake (score 0–3) 0.9 ± 0.1 Limitation of daily activities (score 0–3) 1.2 ± 0.1 None to mild discomfort: 40/78 (51.3%)	NS <0.001 <0.05 <0.001 <0.05	17/55 catheter-free monitoring: chest discomfort
Sweis et al. (2009)	Level III-2 Quality: 15/26 CX P1 Cohort study	110 participants catheter group 134 participants wireless group	Overall satisfaction: 4.4/5	Overall satisfaction: 3.5/5	<0.001	
Bradley et al. (2011)	Level III-2 Quality: 15/26 CX P2 Cohort study	106 participants catheter group 234 participants wireless group	Foreign body chest: 62% Foreign body throat: 21% Chest discomfort: 51% Chest pain: 33% Difficulty swallowing: 42% Painful swallowing: 44% Early fullness: 29% Diminished appetite: 24% Abdominal pain: 19% Nausea: 18% Vomiting: 6%	Foreign body chest: 34% Foreign body throat: 22% Chest discomfort: 26% Chest pain: 19% Difficulty swallowing: 91% Painful swallowing: 62% Early fullness: 22% Diminished appetite: 32% Abdominal pain: 14% Nausea: 20% Vomiting: 6%	0.000 0.887 0.000 0.009 0.000 0.002 0.188 0.144 0.354 0.766 1.000	
Ang et al. (2010)	Level III-2 Quality: 14/26	59 participants catheter group	0 intolerance for device (0%)	4 intolerance for device (6.8%)	?	

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	p-value	Additional notes
	CX P2 Cohort study	66 participants wireless group				
Hakanson et al. (2009)	Level III-2 Quality: 14/26 CX P2 Case-control study	45 volunteers and 47 patients (simultaneous recording)	0 complications	0 complications	NS	
Gillies et al. (2007)	Level III-2 Quality: 13.5/26 CX P2 Non-randomised controlled trial	100 participants catheter group 85 participants wireless group (but 100 studies)	27% no overall discomfort 23% discomfort score >5 86% no impact on work attendance	3% no overall discomfort 45% discomfort score >5 18% no impact on work attendance	<0.0001 NS <0.0001	
Pandolfino et al. (2003)	Level III-2 Quality: 11.5/26 CX P3 Case-control study	30 participants catheter group 29 participants wireless group	Satisfaction (the lower the better): 0.8 ± 0.1 Throat discomfort: 4/29 Oesophageal discomfort: 10/29 Change daily routine: 0/29 Affected diet: 3% Affected activity: 0% Disrupted sleep: 2/22	Satisfaction (the lower the better): 1.9 ± 0.2 Throat discomfort: 22/30 Oesophageal discomfort: 5/30 Change daily routine: 11/30 Affected diet: 47% Affected activity: 60% Disrupted sleep: 9/30	<0.001 <0.001 <0.05 <0.001 <0.001 NS	Wireless group: 2 subjects had discomfort to the extent that the capsule had to be removed; 1 non-detachment; 4 had moderate chest pain
Schneider et al. (2007)	Level III-2 Quality: 9/26 CX P3 Non-randomised controlled trial	55 participants catheter group 78 participants wireless group	Throat discomfort: 3 (3.8%) Oesophagus discomfort: 14 (17.9%) Reduced activities: 5 (6.4%) Reliability: 75 (96.1%)	Throat discomfort: 52 (94.5%) Oesophagus discomfort: 18 (32.7%) Reduced activities: 55 (100%) Reliability: 54 (98.1%)	0.001 NS 0.0001 NS	

Table 71   Complications and adverse events of catheter-free monitoring compared with catheter-based monitoring in children

Study	Level and quality	N	Wireless pH monitoring	Catheter-based pH monitoring	p-value
Croffie et al. (2007)	Level II Quality: 14.5/25 CX P2 Randomised controlled trial	16 participants catheter group 18 participants wireless group	Vomiting more than usual 6/18 (33%) Chest pain 7/18 (39%) Throat pain 8/18 (44%) Oesophageal discomfort 7/18 (39%) Appetite on scale from 1–5: 3.54 Activity on scale from 1–5: 3.66 Overall satisfaction on scale from 1–5: 4.31	Vomiting more than usual 3/16 (19%) Chest pain 1/16 (6%) Throat pain 15/16 (94%) Oesophageal discomfort 11/16 (69%) Appetite on scale from 1–5: 2.72 Activity on scale from 1–5: 2.33 Overall satisfaction on scale from 1–5: 3.11	0.25 0.02 0.001 0.06 0.029 0.001 0.003

Table 72   Complications and adverse events resulting from catheter-free (wireless) oesophageal pH monitoring in adults

Study	Level and quality	Population	Complications per procedure
Lacy et al. (2011)	Level IV Q1 Post-test case series	375 patients	1/375 procedures 1 transient chest discomfort
Nusrat, Roy & Bielefeldt (2012)	Level IV Q1 Retrospective case series	356 patients	0/356 procedures All patients tolerated the study without significant side effects
Crowell et al. (2008)	Level IV Q1 Case series	294 patients	?/294 procedures 147 foreign-body sensation 121 chest discomfort 82 chest pain 82 diminished appetite
Tseng et al. (2005)	Level IV Q1 Retrospective case series	190 procedures, 186 patients	?/190 procedures 1 extreme gagging 2 sufficient chest pain Most patients noted a foreign-body sensation
Ahlawat et al. (2006)	Level IV Q1 Retrospective case series	82 patients	4/82 procedures 4 severe chest pain 53 patients were physically aware of the presence of the pH-monitoring capsule in the oesophagus, symptoms ranging from a foreign-body sensation to a mild discomfort, to chest pain
Tharavej et al. (2006)	Level IV Q1 Case series	80 patients	10/80 procedures 10 new onset chest pain In 6 of these patients the chest pain was associated with hypertensive oesophageal contractions
Remes-Troche et al. (2005)	Level IV Q1 Prospective case series	77 patients	51/77 procedures 26 chest pain 11 foreign-body sensation 5 nausea 9 patients had more than one symptom
des Varannes et al. (2005)[1]	Level IV Q1 In-subject simultaneous recording study	40 patients	4/40 procedures 2 placement failures due to poor tolerance with vomiting 1 dizzy spell during insertion 1 epistaxis
Marchese et al. (2006)	Level IV Q1 Prospective case series	39 patients	7/39 procedures 2 severe retrosternal chest pain 2 mild epistaxis 3 pharyngeal irritation
Karamanolis et al. (2012)	Level IV Q1 Post-test case series	32 patients	4/32 procedures 3 throat pain 2 back pain 1 globus sensation 1 patient with 2 symptoms
Tu et al. (2004)	Level IV Q1 Prospective case series	30 patients	1/30 procedures 1 mucosal abrasion and minor haemorrhage

Study	Level and quality	Population	Complications per procedure
Azzam et al. (2012)[a]	Level IV Q1 In-subject simultaneous recording study	25 patients	0/25 procedures No patient experienced severe chest pain or any other symptom requiring endoscopic removal of the capsule
Prakash et al. (2006)	Level IV Q2 Retrospective case series	452 patients	8/452 procedures (but this only gives the major side effects of the patients who requested removal of the capsule) 7 severe chest pain and/or odynophagia 1 severe foreign-body sensation
Bhat, McGrath & Bielefeldt (2006)	Level IV Q2 Prospective case series	203 patients	19/203 procedures 18 significant chest discomfort 1 rash
Turner et al. (2007)	Level IV Q2 Case series	191 patients	1/191 procedures 1 intolerable chest pain requiring capsule removal
Domingues, Moraes-Filho & Domingues (2011)	Level IV Q2 Retrospective case series	74 patients	19/74 procedures 17 mild foreign-body sensation 2 chest pain during meal periods
Scarpulla et al. (2007)	Level IV Q2 Retrospective case series	83 patients	21/83 procedures 21 occasional retrosternal discomfort on swallowing
Francis (2008)	Level IV Q2 Retrospective case series	76 patients	4/76 procedures 1 severe chest pain 1 capsule dislodgement in the mouth 1 capsule dislodgement in the pyriform sinus 1 capsule dislodgement in the stomach
de Hoyos & Esparza (2010); de Hoyos, Esparza & Loredo (2010)	Level IV Q2 Prospective case series	66 patients	2/66 procedures 1 intolerable chest pain 1 detachment failure
Tankurt et al. (2011)	Level IV Q2 Case series	64 patients	11/64 procedures 11 mild to moderate retrosternal pain or discomfort
Garrean et al. (2008)	Level IV Q2 Post-test case series	60 patients	8/60 procedures 3 back pain 3 throat pain 1 globus sensation 1 chest pain
Belafsky et al. (2004)	Level IV Q2 Prospective case series	46 patients	9/46 procedures 2 laryngospasm 2 epistaxis 1 vasovagal reaction 4 foreign-body sensation
Sweis et al. (2011)	Level IV Q2 Case series	38 patients	0/38 procedures No patient complained of severe symptoms during the study and no capsule needed to be removed
Calabrese et al. (2008)	Level IV Q2 Post-test case series	24 patients	5/24 procedures 5 occasional retrosternal discomfort on swallowing

Study	Level and quality	Population	Complications per procedure
Hirano et al. (2005)	Level IV Q2 Post-test case series	18 patients	0/18 procedures No patient reported any adverse events of the Bravo procedure such as chest pain, dysphagia or bleeding
Pandolfino et al. (2006)[a]	Level IV Q2 Case-control study	10 control subjects and 10 patients with GORD	7/20 procedures 7 mild chest discomfort or foreign-body sensation
Sofi et al. (2011)	Level IV Q3 Retrospective case series	58 patients	?/58 procedures 19 sore throat 17 dysphagia 13 chest pain (2 with severe chest pain) 6 cough 12 patients had day(s) off work
Lee et al. (2005)	Level IV Q3 Post-test case series	40 patients	18/40 procedures 1 retrosternal discomfort with protuberant mucosal lesion 17 retrosternal foreign-body discomfort

[a] These studies are not case series but are still included as level IV safety evidence because they do not report a(n) (appropriate) comparison regarding complications or adverse events.

Table 73   Complications and adverse events resulting from catheter-free (wireless) oesophageal pH monitoring in children

Study	Level and quality	Population	Complications per procedure
Cabrera et al. (2011)	Level IV Q1 Prospective case series	289 children	1/289 procedures 1 failure to release capsule with tear of oesophageal mucosa
Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008)	Level IV Q1 Case series	58 children	5/58 procedures 1 mild chest discomfort 1 period of coughing 3 dysphagia
Hochman & Favaloro-Sabatier (2005)	Level IV Q2 Retrospective case series	38 children	26/38 procedures 26 discomfort 7 significant discomfort
Bothwell, Phillips & Bauer (2004)	Level IV Q2 Retrospective case series	25 children	1/25 procedures 1 superficial mucosal tear

Table 74   Case reports reporting complications and adverse events resulting from catheter-free (wireless) oesophageal pH monitoring

Study	Study design	Population	Complications per procedure
Agrawal, Kerman & Rich (2009)	Case report	1 patient	1/1 procedure: 1 severe retrosternal chest discomfort
Fajardo et al. (2006)	Case report	1 patient	1/1 procedure: 1 oesophageal perforation
Hogan et al. (2009)	Case report	1 patient	1/1 procedure: 1 retention (>2 years) of the capsule in a colonic diverticulum
Kramer & Chokhavatia (2012)	Case report	1 patient	1/1 procedure: 1 capsule dislodgement in the left pyriform sinus
Shahid & Fisher (2011)	Case report	1 patient	1/1 procedure: 1 capsule dislodgement in the left mainstem bronchus
Triester et al. (2005)	Case report	1 patient	1/1 procedure: 1 sharp substernal chest pain and dyspnoea
Renteln et al. (2008)	Case report	1 patient	1/1 procedure: 1 capsule aspiration and decreased oxygen saturation

Table 75   Chest discomfort / complications resulting from catheter-free (wireless) oesophageal pH monitoring in adults

Study	Level and quality	Population	Chest discomfort / complications per procedure
Lacy et al. (2011)	Level IV Q1 Post-test case series	375 patients	1/375 procedures 1 transient chest discomfort
Crowell et al. (2008)	Level IV Q1 Case series	294 patients	203/294 procedures 121 chest discomfort 82 chest pain
Tseng et al. (2005)	Level IV Q1 Retrospective case series	190 procedures, 186 patients	2/190 procedures 2 sufficient chest pain
Ahlawat et al. (2006)	Level IV Q1 Retrospective case series	82 patients	4/82 procedures 4 severe chest pain 53 patients were physically aware of the presence of the pH-monitoring capsule in the oesophagus, with symptoms ranging from a foreign-body sensation to a mild discomfort, to chest pain
Tharavej et al. (2006)	Level IV Q1 Case series	80 patients	10/80 procedures 10 new onset chest pain In 6 of these patients the chest pain was associated with hypertensive oesophageal contractions
Remes-Troche et al. (2005)	Level IV Q1 Prospective case series	77 patients	26/77 procedures 26 chest pain
Marchese et al. (2006)	Level IV Q1 Prospective case series	39 patients	2/39 procedures 2 severe retrosternal chest pain
Azzam et al. (2012)[a]	Level IV Q1 In-subject simultaneous recording study	25 patients	0/25 procedures No patient experienced severe chest pain or any other symptom requiring endoscopic removal of the capsule
Prakash et al. (2006)	Level IV Q2 Retrospective case series	452 patients	7/452 procedures (but this only gives the major side effects of the patients who requested removal of the capsule) 7 severe chest pain and/or odynophagia
Bhat, McGrath & Bielefeldt (2006)	Level IV Q2 Prospective case series	203 patients	18/203 procedures 18 significant chest discomfort
Turner et al. (2007)	Level IV Q2 Case series	191 patients	1/191 procedures 1 intolerable chest pain requiring capsule removal
Domingues, Moraes-Filho & Domingues (2011)	Level IV Q2 Retrospective case series	74 patients	2/74 procedures 2 chest pain during meal periods
Scarpulla et al. (2007)	Level IV Q2 Retrospective case series	83 patients	21/83 procedures 21 occasional retrosternal discomfort on swallowing

Study	Level and quality	Population	Chest discomfort / complications per procedure
Francis (2008)	Level IV Q2 Retrospective case series	76 patients	1/76 procedures 1 severe chest pain
de Hoyos & Esparza (2010); de Hoyos, Esparza & Loredo (2009)	Level IV Q2 Prospective case series	66 patients	1/66 procedures 1 intolerable chest pain
Tankurt et al. (2011)	Level IV Q2 Case series	64 patients	11/64 procedures 11 mild to moderate retrosternal pain or discomfort
Garrean et al. (2008)	Level IV Q2 Post-test case series	60 patients	1/60 procedures 1 chest pain
Calabrese et al. (2008)	Level IV Q2 Post-test case series	24 patients	5/24 procedures 5 occasional retrosternal discomfort on swallowing
Hirano et al. (2005)	Level IV Q2 Post-test case series	18 patients	0/18 procedures No patient reported any adverse events of the Bravo procedure such as chest pain, dysphagia or bleeding
Pandolfino et al. (2006)[a]	Level IV Q2 Case-control study	10 control subjects and 10 patients with GORD	7/20 7 mild chest discomfort or foreign-body sensation
Sofi et al. (2011)	Level IV Q3 Retrospective case series	58 patients	13/58 11 chest pain 2 severe chest pain
Lee et al. (2005)	Level IV Q3 Post-test case series	40 patients	18/40 procedures 1 retrosternal discomfort with protuberant mucosal lesion 17 retrosternal foreign-body discomfort

[a] These studies are not case series but are still included as level IV safety evidence because they do not report a(n) (appropriate) comparison regarding complications or adverse events.

Appendix F Non-comparative technical efficacy data

Table 76   Technical efficacy of catheter-free (wireless) oesophageal pH monitoring in adults

Study	Study design	N	Technical/operative/equipment/recording failure
Lacy et al. (2011)	Case series Level: IV Quality: Q1 CX P2	358 patients	6/358 procedures: (1.7%) 6 early detachments or loss of signal
Nusrat, Roy & Bielefeldt (2012)	Retrospective case series Level: IV Quality: Q1 CX P2	356 patients	6/356 procedures: (1.7%) 6 early detachments
Park et al. 2013 (2013)	Case series Level: IV Quality: Q1 CX P2	230 patients	32/230 procedures: (13.9%) 17 early detachments 10 attachment failures or misplacements 5 recording failures
Tseng et al. (2005)	Retrospective case series Level: IV Quality: Q1 CX P2	209 patients (and 213 procedures)	23/213 procedures (10.8%): 18 incomplete data captures (<18 hours on at least 1 day) 5 early detachment
Crowell et al. (2009)	Case series Level: IV Quality: Q1 CX P2	180 patients	15/180 procedures (8.3%) 15 early detachments
Doma et al. (2010)	Post-test case series Level: IV Quality: Q1 CX P2	161 patients	56/161 procedures: (38.0%) 56 errors in placement (inaccurate placement of the capsule)
Iqbal et al. (2007)	Retrospective case series Level: IV Quality: Q1 CX P2	100 patients	11/100 procedures (11.0%) 11 early detachments
Ahlawat et al. (2006)	Retrospective case series Level: IV Quality: Q1 CX P2	90 patients	9/90 procedures: (10.0%) 3 failure to transmit data to receiver 1 capsule failed to deploy from catheter 4 incomplete data captures 1 failure of attachment to the oesophageal mucosa

Study	Study design	N	Technical/operative/equipment/recording failure
Pandolfino et al. (2003)[a]	Case-control study Level: IV Quality: Q1 CX (clinical diagnosis) P3	41 patients and 44 volunteers	3/85 procedures (3.5%): 2 capsules failed to deploy from catheter 1 detachment failure
Remes-Troche et al. (2005)	Prospective case series Level: IV Quality: Q1 CX P2	84 patients	7/84 procedures (8.3%): 4 attachment failures 3 early detachments
Bansal et al. (2009)	Case series Level: IV Quality: Q1 CX P2	48 patients	8/48 procedures: (16.7%) 4 premature detachments 2 cardia placements 2 incomplete data captures
Marchese et al. (2006)	Prospective case series Level: IV Quality: Q1 CX P2	38 patients	2/38 procedures (5.3%): 1 capsule displacement (in the stomach) 1 failure because of previous nasal surgery (tight nasal vault?)
Tu et al. (2004)	Prospective case series Level: IV Quality: Q1 CX P2	30 patients	1/30 procedures (3.3%) 1 attachment failure
Bhat, McGrath & Bielefeldt (2006)	Prospective case series Level: IV Quality: Q2 CX P2	217 patients	10/217 procedures (4.6%): 1 capsule failed to deploy from catheter 2 receiver malfunction 7 early detachments
Turner et al. (2007)	Case series Level: IV Quality: Q2 CX P2	191 patients	42/191 procedures (22.0%): 27 incomplete data captures 15 early detachment
Gillies et al. (2007)[a]	Non-randomised controlled trial Level IV Quality: Q2 CX P2	100 procedures (85 patients) in catheter-free group	17/100 procedures (17%) 7 early detachments 4 incomplete or no data capture 6 failures of attachment to the oesophageal mucosa

Study	Study design	N	Technical/operative/equipment/recording failure
Domingues, Moraes-Filho & Domingues (2011)	Retrospective case series Level: IV Quality: Q2 CX P2	100 patients	7/100 procedures (7.0%): 1 failure of attachment to the oesophageal mucosa 2 early detachments 2 recording failure 2 downloaded data lost
Scarpulla et al. (2007)	Retrospective case series Level: IV Quality: Q2 CX P2	83 patients	9/83 procedures (10.8%): 4 failures to deploy or transmission failures 5 early detachments
Francis (2008)	Retrospective case series Level: IV Quality: Q2 CX P2	76 patients	9/76 procedures (11.8%): 5 technical failures 2 early detachments 4 failures of attachment to the oesophageal mucosa
Wenner et al. (2007 & 2008)[a]	Diagnostic case-control study Level: IV Quality: Q2 CX P3	70 patients	4/70 procedures (5.7%): 1 failure to pass through the upper oesophagus 1 early detachment 1 technical failure 1 placement failure
de Hoyos & Esparza (2010); de Hoyos, Esparza & Loredo (2009)	Prospective case series Level: IV Quality: Q2 CX P2	66 patients	9/66 procedures (13.6%): 3 capsules had poor data reception 3 early detachments 1 transmission failure 1 detachment failure 1 error in placement
Tankurt et al. (2011)	Case series Level: IV Quality: Q2 CX P2	64 patients	2/64 procedures (3.1%): 2 early detachments
Ward et al. (2004)	Case series Level: IV Quality: Q2 CX P2	60 patients	9/60 procedures (15%): 7 failures of attachment to the oesophageal mucosa 2 failures to retrieve data from recorder
Garrean et al. (2008)	Post-test case series Level: IV Quality: Q2 CX P2	60 patients	1/60 procedures (1.7%): 1 capsule malfunction

Study	Study design	N	Technical/operative/equipment/recording failure
Grigolon et al. (2011)[a]	One arm of matched pairs retrospective cohort study Level: IV Quality: Q2 CX P2	57 patients in the catheter-free group	6/57 procedures (10.5%): 3 incomplete data capture 3 early detachment
Belafsky et al. (2004)	Prospective case series Level: IV Quality: Q2 CX P2	46 patients	6/46 procedures (13.0%): 2 failures because of tight nasal vault 2 early detachments 1 lost data recorder 1 capsule failed to deploy from catheter
Sweis et al. (2011)	Case series Level: IV Quality: Q2 CX P2	38 patients	0/38 procedures (0%): 0 early detachments (detachment occurred between 48 and 72 hours in 5 patients and between 72 and 96 hours in 12 patients)
Martinez de Haro et al. (2008)[a]	One arm of non-randomised controlled trial Level: IV Quality: Q2 CX P2	25 patients and 10 volunteers	6/35 procedures (17.1%) 3 incorrect placements 2 early detachments 1 immediate detachment with aspiration
Bechtold et al. (2007)	Retrospective case series Level: IV Quality: Q2 CX P2	27 patients	1/27 procedures (3.7%): 1 early detachment
Francis et al. (2012)[a]	Randomised blinded controlled trial Level: IV Quality: Q2 C1 P2	22 patients	0/22 procedures (0%): All capsules were placed successfully on the first attempt
Pandolfino et al. (2006)[a]	Case-control study Level: IV Quality: Q2 CX (clinical diagnosis) P3	10 patients and 10 volunteers	2/20 procedures (10%): 2 early detachments
Sofi et al. (2011)	Retrospective case series Level: IV Quality: Q3 CX P2	58 patients	5/58 procedures (8.6%): 2 malfunctions of the device during insertion 3 transmission failures

Study	Study design	N	Technical/operative/equipment/recording failure
Lee et al. (2005)	Prospective case series Level: IV Quality: Q3 CX P2	40 patients	4/40 procedures (10%): 3 attachment failures 1 early detachment

a These studies are not case series but are still included as level IV safety evidence because they do not report a(n) (appropriate) comparison regarding technical efficacy.

Table 77   Technical efficacy of catheter-free (wireless) oesophageal pH monitoring in children

Study	Study design	N	Technical/operative/equipment/recording failure
Cabrera et al. (2011)	Prospective case series Level: IV Quality: Q1 CX P2	289 children	41/289 procedures (14.2%) 3 capsules failed to calibrate 8 failures of attachment to the oesophageal mucosa 4 capsules failed to deploy from catheter 1 recording failure 4 early detachments 21 incomplete data captures
Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008)	Case series Level: IV Quality: Q1 CX P2	58 children	3/58 procedures (5.2%): 1 transmission failure 2 attachment failures
Hochman & Favaloro-Sabatier (2005)	Retrospective case series Level: IV Quality: Q2 CX P2	50 children	5/50 procedures (10%): 1 failure of attachment to the oesophageal mucosa 4 shorter monitoring periods

Appendix G Day-to-day variability in oesophageal acid exposure

Table 78 Day-to-day variability in oesophageal acid exposure with catheter-free monitoring in adults

Study	Level and quality	N	Outcomes	Day 1	Day 2	p-value	Correlation co-efficient
des Varannes et al. (2005)	In-subject simultaneous recording study Level: II Quality: 20/26 CX P2	30 patients	Median % total time pH <4 (25%–75%) Number of reflux episodes >5 minutes Number of reflux episodes 24-hour acid exposure values	2.4 (0.7–3.9) 2 (1–5) 23 (10–32)	3.3 (1.4–7.5) 4 (2–8) 23 (13–36)	<0.04 0.07 NS	0.79
Grigolon et al. (2007)	Pseudo-randomised controlled trial Level: III-1 Quality: 13.5/26 CX P2	55 patients	% total time pH <4	3.4 (0.6–9.8)	4.9 (1.7–8.3)	NS	
Gillies et al. (2007)	Non-randomised Controlled trial Level: III-2 Quality: 13.5/26 CX P2	89 patients	Median % total time pH <4 (IQR) Median % upright time pH <4 (IQR) Median % supine time pH <4 (IQR) DeMeester score	4.0 (1.2–10.3) 5.0 (1.4–12.0) 0.5 (0–5.0) 15.9 (5.1–34.6)	4.3 (1.0–10.0) 5.0 (1.5–11.3) 0.5 (0–3.9) 16.2 (4.5–34.8)	0.64 0.56 0.23 0.90	
Hakanson et al. (2009)	Case-control study with in-subject simultaneous recording Level: III-2 Quality: 14/26 C1 (for accuracy) P2	55 patients	Median % total time pH <4 in 1st series of patients (5th–95th percentile) Median % total time pH <4 in 2nd series of patients (5th–95th percentile)	3.2 (0.1–7.9) 2.4 (0–9.6)	6.9 (0.7–16.3) 5.2 (0.2–14.8)		
Pandolfino et al. (2003)	Case series Level: Q2 Quality: 11.5/26	37 patients	Mean % total time pH <4 (SEM) Mean % upright time pH <4 (SEM) Mean % supine time pH <4 (SEM)	6.6 (0.8–27.6) 7.9 (1.40–29.1) 1.2 (0.0–23.9)	7.7 (0.9–28.6) 7.2 (1.1–31.1) 1.5 (0.0–45.5)	NS NS NS	

Study	Level and quality	N	Outcomes	Day 1	Day 2	p-value	Correlation co-efficient
	CX (clinical diagnosis) P3						
Ayazi et al. (2011)	Case series Level: IV Quality: Q1 CX (no comparator) P2	310 patients	% total time pH <4 % upright time pH <4 % supine time pH <4 Number of reflux episodes Number of episodes >5 minutes Longest reflux episode (minutes) Composite pH score Abnormal composite pH score on 1 of the 2 days Only abnormal pH score on day 1 Only abnormal pH score on day 2 % total time pH <4 (day 1 vs day 2) Composite pH score (day 1 vs day 2)	3.7 (1.0–9.9) 4.5 (1.2–12.1) 0.3 (0.0–11.5) 39 (13–74) 2 (0–7) 10 (3–25) 13.1 (4.7–32.0) 19% (60/310) 45% (27/60) 55% (33/60)	4.5 (0.9–9.7) 4.8 (1.0–11.5) 0.3 (0.0–6.3) 36 (14.5–76) 2 (0–7) 9 (3.0–21.5) 14.4 (3.9–30.5)	0.1679 0.5238 0.4963 0.5746 0.3869 0.3006 0.6552	0.94 (0.93-0.95) 0.83 (0.79-0.86)
Tseng et al. (2005)	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	190 patients	% total time pH <4 more than 5.3% (N) DeMeester score >14.7 (N)	96 106	87 78		
Crowell et al. (2009)	Case series Level: IV Quality: Q1 CX (no comparator) P2	165 patients	% total time pH <4	6.9	7.6	0.20	
Chander et al. (2012)	Retrospective case series Level: IV Quality: Q1 CX (no comparator) P2	124 patients	Median % total time pH <4 (range) Median % upright time pH <4 (range) Median % supine time pH <4 (range) Median number of reflux episodes (range) DeMeester scores	3.55 (0.00–29.90) 4.50 (0.00–35.30) 0.80 (0.00–46.40) 35.50 (0.00-253.00) 13.25 (0.30-104.00)	3.75 (0.00–29.10) 5.50 (0.00–32.10) 0.20 (0.00–28.80) 38.5 (0.00–299.00) 15.15 (0.00–96.20)	0.8092 0.5464 0.1778 <0.0001 NA	
Ahlawat et al. (2006)	Retrospective case series Level: IV	82 patients	Significant correlation in reflux patterns in patients from day 1 to day 2	74.4%			

Study	Level and quality	N	Outcomes	Day 1	Day 2	p-value	Correlation co-efficient
	Quality: Q1 CX (no comparator) P2		No predictable pattern of reflux seen	28% (23/82)			
Remes-Troche et al. (2005)	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	77 patients	Median % total time pH <4 (range) Median % upright time pH <4 (range) Median % supine time pH <4 (range) Mean number of reflux episodes ± SE Mean number of long reflux episodes ± SE	5.5 (0–35) 5.9 (0–22) 1.4 (0–31.3) 45.3 ± 3 4.2 ± 0.5	5.7 (0–22) 6.0 (0–20) 0.66 (0–24) 65 ± 5 3.05 ± 0.9	NS NS NS 0.004 NS	
Prakash & Clouse (2006)	Case series Level: IV Quality: Q1 CX (no comparator) P2	62 patients	% total time pH <4 off therapy % total time pH <4 on therapy	5.2 ± 0.8 1.6 ± 0.4	5.6 ± 0.9 0.8 ± 0.4		
Lacy et al. (2009)	Prospective case series Level: IV Quality: Q1 CX (no comparator) P2	50 patients	Mean % total time pH <4 (SD) Mean total time pH <4 (SD) Mean total number of reflux episodes (SD) Mean total number of long reflux episodes (SD) Mean DeMeester score (SD) Abnormal study (% time pH <4 is more than 6%)	6.6 ± 5.9 90.2 ± 82.7 80 ± 52 3 ± 4 24.7 ± 21.6 28 (56%)	5.4 ± 4.6 74.6 ± 64.2 67 ± 44 3 ± 3 19.4 ± 15.9 21 (42%)		
Bhat, McGrath & Bielefeldt (2006)	Prospective case series Level: IV Quality: Q2 CX (no comparator) P2	203 patients	% time abnormal pH	5.5 (CI = 2.1–10.4)	4.2 (CI = 1.2–9.2)	<0.001	
Turner et al. (2007)	Case series Level: IV Quality: Q2 CX (no comparator	148 patients	% total time pH <4 is more than 4% (N) % upright time pH <4 is more than 6.3% (N) % supine time pH <4 is more than 1.2% (N)	76 67 54	65 51 53	0.03	
Korrapati et al. (2011)	Retrospective case series Level: IV Quality: Q2	108 patients	% total time pH <4: mean, median, SD (range) % upright time pH <4: mean, median, SD (range)	6.36, 4.85, 7.60 (0.0–48.1) 7.31, 5.70, 7.55	6.02, 4.25, 8.02 (0.0–67.1) 6.52, 4.85, 7.89	0.73 0.0576	

Study	Level and quality	N	Outcomes	Day 1	Day 2	p-value	Correlation co-efficient
	CX (no comparator) P2		% supine time pH <4: mean, median, SD (range) DeMeester score: mean, median, SD (range)	(0.0–41.4) 4.01, 0.25, 8.94 (0.0–60.2) 22.0, 16.3, 25.7 (0.0–163.7)	(0.0–59.6) 4.66, 0.00, 12.20 (0.0–87.4) 21.1, 15.2, 25.6 (0.0–213.0)	0.56 0.90	
Domingues, Moraes-Filho & Domingues (2011)	Retrospective case series Diagnostic yield study Level: IV Quality: Q2 CX (no comparator) P2	74 patients	Abnormal pH data (total time pH <4 is >4.4%) Positive SAP	51 (68.9%) 19 (44.1%)	49 (70.2%) 17 (40.4%)		
Belafsky et al. (2005)	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	78 sedated placements and 128 unsedated placements	Mean reflux episodes (sedated group) Mean reflux episodes (unsedated group) % total time pH <4 (sedated group) % total time pH <4 (unsedated group) Mean composite score (sedated group) Mean composite score (unsedated group) Number of reflux episodes day 1 : day 2 % total time pH <3 day 1 : day 2 Composite score day 1 : day 2 Number of long reflux episodes day 1 : day 2	46 56 5.5 6.7 18.3 23.3	46 52 5.1 6.8 18.2 22.2		0.715 0.650 0.694 0.773
Garrean et al. (2008)	Post-test case series Level: IV Quality: Q2 CX (no comparator) P2	60 patients	Number of reflux episodes (median)	24	20	NS	
Bechtold et al. (2007)	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	26 patients	% total time pH <4 % total upright time pH <4 % total supine time pH <4 Number of reflux episodes (mean) Number of long reflux episodes (mean)	10.4 11.7 9.4 65.3 6.8	7.1 6.3 10.0 55.6 3.8	0.0049 0.0051 0.8596 0.0257 0.0077	

Study	Level and quality	N	Outcomes	Day 1	Day 2	p-value	Correlation co-efficient
			Mean duration of long refluxes (minutes)	26.5	19.7	0.0617	
			Total heartburn episodes	6.2	8.0	0.2752	
			Heartburn episodes with pH <4	2.5	2.9	0.686	
			Symptom score	0.323	0.254	0.1353	

NS = not significant

Table 79   Day-to-day variability in oesophageal acid exposure with catheter-free monitoring in children

Study	Level and quality	N	Outcomes	Day 1	Day 2	p-value
Croffie et al. (2007)	Level: II Quality:14.5/26 Randomised controlled trial	27 children	Median reflux index (range)	2.3 (0.4–16.6)	2.8 (0.2–15.6)	0.345
Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008)	Case series Level: IV Quality: Q1 CX (no comparator) P2	58 children	Median % total time pH <4, SD (range) Median number of reflux episodes, SD (range) Median number of long reflux episodes, SD (range) Median longest reflux episode, SD (range) DeMeester score, SD (range)	5.4 ± 6.8 (0–32.2) 60 ± 47 (0–203) 3 ± 5 (0–70) 13 ± 15 (0–70) 20.5 ± 23.7 (0.3–112.4)	6.2 ± 8.4 (0–41.2) 59 ± 51 (0–258) 3 ± 5 (0–26) 17 ± 20 (0–83) 22.9 ± 28.7 (0.3–140.8)	NS NS NS NS NS
Hochman & Favaloro-Sabatier (2005)	Retrospective case series Level: IV Quality: Q2 CX (no comparator) P2	44 children	Median number of reflux episodes (range) Median number of long reflux episodes (range) Median % total time pH <4 (range) Median reflux time (minutes) (range) Longest reflux time (minutes) (range)	28 (0, 83) 2 (0, 13) 4.2 (0, 23.6) 54.5 (0, 309) 10.5 (0, 53)	22 (2, 96) 1 (0, 14) 2.6 (0.1–18.6) 31 (1, 257) 5 (1, 46)	0.99 0.11 0.01 0.01 0.0007

No correlation co-efficients calculated.
NS = not significant

Appendix H   Non-comparative diagnostic yield

Table 80   Non-comparative diagnostic yield with catheter-free monitoring in adults

Study	N	Diagnostic yield (%) wireless pH monitoring
Ayazi et al. (2011)	310 patients suspected of GORD	Based on composite pH score Both days: 123 (39.7) On 1 day only: 60 (19.4)
Ayazi et al. (2009)	28 patients with strong clinical GORD evidence (1) and 77 patients with indeterminate clinical evidence of GORD (2)	Based on composite pH score 1: Both days: 19 (67.9) 1: Overall: 26 (92.9) 2: Day 1: 40 (52) 2: Day 2: 42 (55) 2: Both days: 31 (40)
Bansal et al. (2009)	20 patients with erosive oesophagitis (EO) and 20 patients with non-erosive reflux disease (NERD)	EO proximal: 15 (75) EO distal: 17 (85) NERD proximal: 11 (51) NERD distal: 6 (30)
Bhat, McGrath & Bielefeldt (2006)	209 patients suspected of GORD	95 (45.5)
Domingues, Moraes-Filho & Domingues (2011)	74 patients suspected of GORD	Overall: 58 (78.3) Day 1: 51 (68.9) Day 2: 49 (70.2)
Gillies et al. (2007)	95 patients suspected of GORD, or following anti-reflux surgery	Overall: 66 (69) Only on day 1 (normal day 2): 7 (7) Only on day 2 (normal day 1): 10 (10)
Grigolon et al. (2011)	51 patients suspected of GORD	Day 1: 24 (47.1) Overall (48h): 25 (49.0) Overall (96h): 29 (56.9)
Karamanolis et al. (2012)	32 patients with non-cardiac chest pain	17 (53.1)
Lacy et al. (2011)	175 patients on PPI therapy and 177 patients off PPI therapy	On PPI therapy: 45 (25.7) Off PPI therapy: 108 (61.0)
Pandolfino et al. (2003)	37 GORD patients	Both days: 24 (64.9) Day 1: 25 (67.6) Worst day: 31 (83.8)
Scarpulla et al. (2007)	34 patients suspected of GORD	Day 1: 12 (35.3) 48h: 14 (41.2) 72h: 16 (47.1) 96h: 19 (55.9) Worst day: 21 (61.8)
Turner et al. (2007)	115 patients suspected of GORD	Both days: 58 (50.4) Only on day 1 (normal day 2): 18 (15.7) Only on day 2 (normal day 1): 7 (6.1) On 1 day only: 25 (21.7) Day 1: 76 (66.1) Day 2: 65 (56.5)
Ward et al. (2004)	58 patients suspected of GORD	30 (52)

Table 81   Non-comparative diagnostic yield with catheter-free monitoring in children

Study	N	Diagnostic yield (%) wireless pH monitoring
Gunnarsdottir, Stenstrom & Arnbjornsson (2007, 2008)	49 children suspected of GORD	Based on DeMeester scores: 33 (67.3)

Appendix I   MBS items for anti-reflux surgery

Table 82   Anti-reflux surgery MBS item descriptors and claims from July 2012 to June 2013

| MBS item number	Description	Number of claims 2012–13
43951	GASTRO-OESOPHAGEAL REFLUX with or without hiatus hernia, laparotomy and fundoplication for, without gastrostomy, Fee: $871.30	86
43954	GASTRO-OESOPHAGEAL REFLUX with or without hiatus hernia, laparotomy and fundoplication for, with gastrostomy, Fee: $1,065.75	5
31464	ANTIREFLUX OPERATION BY FUNDOPLASTY, via abdominal or thoracic approach, with or without closure of the diaphragmatic hiatus, by laparoscopic technique - not being a service to which item 30601 applies, Fee: $871.30	1,129
30527	ANTIREFLUX OPERATION by fundoplasty, via abdominal or thoracic approach, with or without closure of the diaphragmatic hiatus not being a service to which item 30601 applies, Fee: $871.30	309
30530	ANTIREFLUX operation by cardiopexy, with or without fundoplasty, Fee: $784.20	305
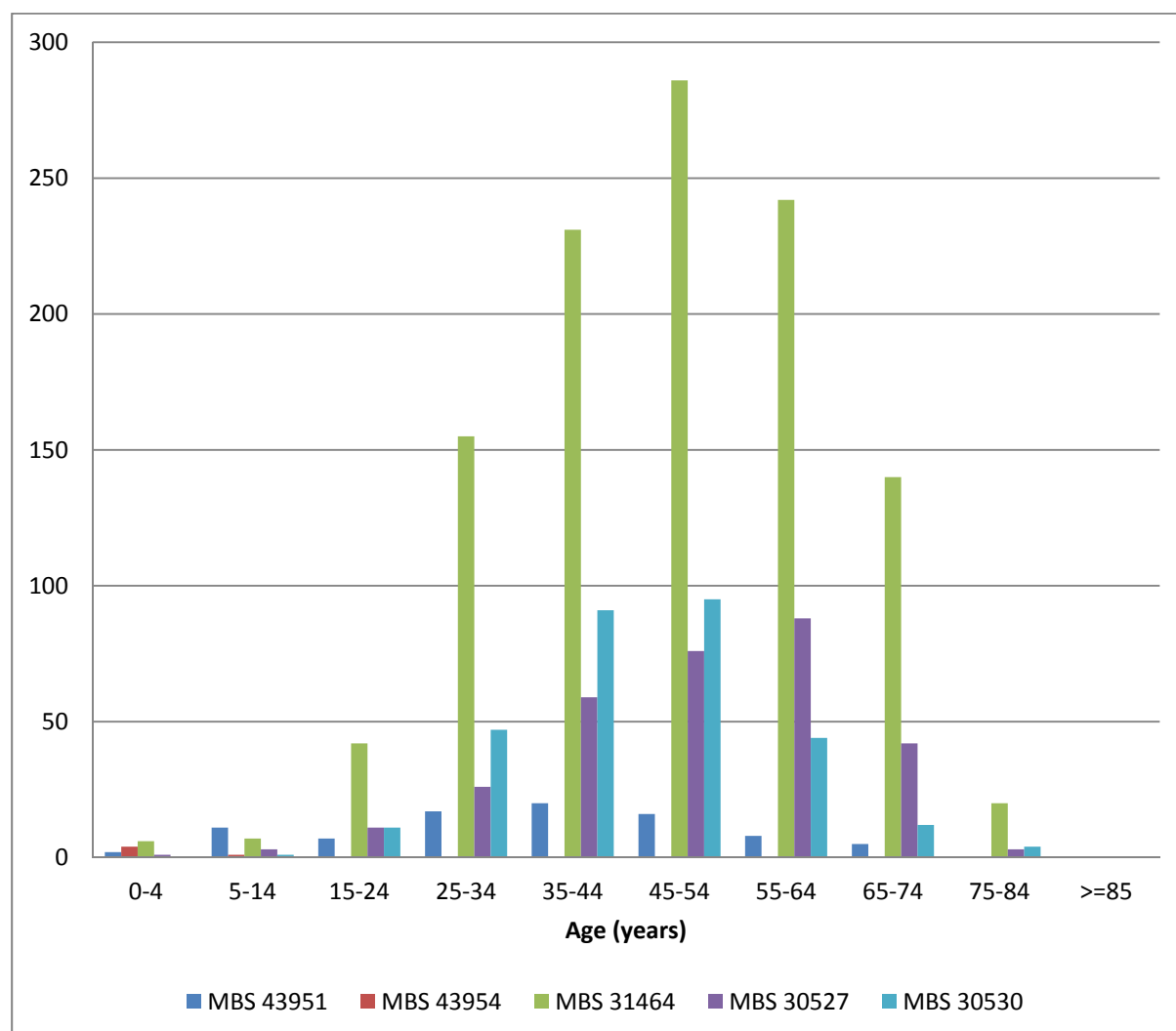
Figure 10   Claims for anti-reflux procedures for GORD, July 2012 – June 2013, per age group

Appendix J   Economic literature search

Table 83   Literature search in PubMed 21 November 2013

	Add to builder	Query	Items found
#13	Add	Search (#9 and #12)	6
#12	Add	Search pH monitoring	194301
#9	Add	Search (#4) AND model	62
#4	Add	Search (#2) AND #3	758
#3	Add	Search ((economic) OR cost-effective) OR cost-utility	693357
#2	Add	Search ((gastroesophageal reflux) OR GORD) OR GERD	26220
#1	Add	Search ((gastroesophageal reflux) OR GORD) OR GORD	25124

Table 84   Literature containing economic models relating to pH monitoring and management of GORD symptoms

Author, Journal	Title	Description, *Comment*
Kleiman DA, Beninato T, Bosworth BP, Brunaud L, Ciecierega T, Crawford CV Jr, Turner BG, Fahey TJ 3rd and Zarnegar R 2014 *Journal of Gastrointestinal Surgery*, vol. 18, no. 1, pp. 2–33; discussion pp. 33–34.	Early referral for esophageal pH monitoring is more cost-effective than prolonged empiric trials of proton-pump inhibitors for suspected gastroesophageal reflux disease	Cost analysis of routine pH monitoring (after 8-week trial of PPI) vs ongoing treatment and delayed pH monitoring. Modelled cohort-based analysis with up to 10 years of costing data (retrospective). No outcome analysis. US setting. *Some clinical inputs/assumptions are applicable.*
Lee WC, Yeh YC, Lacy BE, Pandolfino JE, Brill JV, Weinstein ML, Carlson AM, Williams MJ, Wittek MR and Pashos CL 2008 *Current Medical Research and Opinion*, vol. 24, no. 5, pp. 1317–1327.	Timely confirmation of gastro-esophageal reflux disease via pH monitoring: estimating budget impact on managed care organizations	Modelled cost analysis of wireless pH monitoring vs empiric treatment in patients with normal endoscopy. One-year time horizon. No outcome analysis. US setting. *Some clinical inputs/assumptions are applicable.*
Ofman JJ, Dorn GH, Fennerty MB and Fass R 2002 *Alimentary Pharmacology & Therapeutics*, vol. 16, no. 2, pp. 261–273.	The clinical and economic impact of competing management strategies for GORD	Cost-effectiveness analysis of diagnostic management of GORD; 'step-up management [H2RAs reg then high-dose, PPIs reg then high-dose, endoscopy, pH monitoring] vs high-dose PPI trial. Results presented as costs/symptom-free patient and cost/QALY. One-year time horizon. US setting. *Structure/inputs not directly applicable.*
Botoman VA 2002 *Journal of Clinical Gastroenterology*, vol. 34, no. 1, pp. 6–14.	Non-cardiac chest pain	Opinion piece – no economic modelling.
Borzecki AM, Pedrosa MC, Prashker MJ 2000 *Archives of Internal Medicine*, vol. 160, no. 6, pp. 844–452.	Should non-cardiac chest pain be treated empirically? A cost-effectiveness analysis	Cost-minimisation analysis of empirical medication vs series of investigations. US setting. *Not directly applicable.*
Ancona E, Zaninotto G, Costantini M, Polo R and Origbe E 1991 *Minerva Chirurgica* vol. 46, 7 Suppl, pp. 263–271.	[Esophageal manometry and pH-monitoring: cost–benefit analysis]	Article in Italian and inadequate detail in abstract to be useful. *Not reviewed further.*

H2RA = histamine-2 receptor antagonist, PPI = proton pump inhibitor

Appendix K   Additional information relating to the economic evaluation

Table 85   A summary of variables used in the model

Parameter	Value	Alternatives tested	Source and Discussion
Prevalence	40%	20% to 80%	Broeders et al 2010, for the base case. Range approximates Lacy et al 2011 to Domingues et al 2011).
Sensitivity of wireless testing	100%	86%, 90%, 95%,	Assumption, informed by the clinical report and with sensitivity analysis based on clinical report findings.
Specificity of wireless testing	100%	77%, 90%, 95%	Assumption, informed by the clinical report and with sensitivity analysis based on clinical report findings.
Sensitivity of catheter-based testing	100%	80.5%	Specified as the reference standard in the DAP specification. Sensitivity analysis informed by diagnostic yield in the report findings and expert opinion.
Specificity of catheter-based testing	100%	No tested	Specified as the reference standard in the DAP specification.
Technical failure rate of wireless testing	7.50%	22.5% (i.e. trebled)	Main body of Report. Although this technically refers to the failure rate during insertion, reports vary considerably and a learning curve may exist.
Technical failure rate of catheter-based testing	0.89%	-	Main body of Report. A conservative estimate given that it its less than one third the rate attributed to wireless monitoring, yet the best comparative evidence suggested a RR of failure of 3.0.
Technical failure re-attempt rate	50%	90%	Arbitrary estimates.
Standard GORD treatment effectiveness	89%		Effectiveness of high dose GORD medication or routine fundoplication surgery (in patients with acid-related conditions) assumed equivalent, based on Beck et al 2009, Broeders et al 2010, Grant, Boachie, 2013).
Surgery uptake in +ve patients	25%	75%	Calibration of Medicare data on catheter-based testing and surgery rates.
Surgery revision rate	10%	*	Medicare revision surgery rate. Note: Sensitivity analysis were surgery costs are increased by $10,000pp is equivalent to base case surgery cost and revision rate >100%.
Surgery Mortality rate	0.05%	nil - 0.5%	Base case from Grant et al 2008. Alternative; Epstein et al 2009.
Medication compliance	100%	-	Applies to both PPIs and alternative medications – assumption.
Investigation for alternative diagnosis	90%	50%-100%	Arbitrary estimate, consistent with expert opinion and conservative approach.
Treatment success for alternative conditions (non-acid related)	100%	-	Specified assumption in DAP

Table 86   Scenario/sensitivity tests around other inputs into the economic model (Model 1 vs empirical treatment)

	Cost (wireless pH monitoring)	Outcomes (wireless pH monitoring)	Costs (no test)	Outcomes (no test)	ICER
Base-case	$8,705	7.874	$6,927	7.751	$14,457
No discounting vs 5% per year discounting costs and outcomes	$8,924	8.137	$7,158	8.01	$13,907
Surgical mortality rate 0.1% vs base-case rate of 0.05%	$8,705	7.874	$6,927	7.751	$14,457
Surgical mortality rate 0.5% vs base-case rate of 0.05%	$8,701	7.871	$6,927	7.751	$14,785
Technical failure rate tripled to 22.5% vs base-case rate of 7.5%	$8,828	7.864	$6,927	7.751	$16,827
Re-test after failure uptake: 90% vs base-case rate of 50%	$8,756	7.878	$6,927	7.751	$14,403
Surgery costs (incl. revisions) with additional cost of $10,000 per patient ($17,294 vs base-case of $7,294)	$9,667	7.874	$6,927	7.751	$22,278
Costs of alternative treatment investigation/treatment increased by $1000/month in the initial 12 months. (i.e. up to ($1050/month vs $50/month)	$14,607	7.874	$12,723	7.751	$16,050
Prevalence of acid-related condition in refractory population; 20% (vs base case rate of 40%)	$8,308	7.917	$6,661	7.750	$9,863
Prevalence of acid-related condition in refractory population; 80% (vs base case rate of 40%)	$9,499	7.787	$7,460	7.755	$63,723

Appendix L   Additional information for the financial and costing analysis

Table 87   Copy of "Table 1. Bravo pH Monitoring System components and costs" as published in the 'ASGE Technology Status Evaluation Report: wireless esophageal pH monitoring system' (Chotiprashidi et al, 2005).

Description	Price (US$)
Bravo pH capsule with delivery system, box of 5	$1,125
AA lithium batteries (pack of 4)	$15
Bravo pH receiver	$6,900
Calibtration buffer pH 1.07 (500mL bottle)	$22
Calibtration buffer pH 7.01 (500mL bottle)	$22
Bravo Kit (includes 2 Bravo receivers, Datalink, calibration stand, vacuum pump, calibration buffers pH 7.01 and 1.07, 4 pack AA lithium batteries)	$16,649
Bravo Kit with software (includes 2 Bravo receivers, POLYGRAM NET pH Testing Application software, Datalink, calibration stand, vacuum pump, calibration buffers pH 7.01 and 1.07, 4 pack AA lithium batteries)	$22,149
Bravo Kit with softare and Workstation (includes 2 Bravo receivers, POLYGRAM NET pH Testing Application software, Datalink, calibration stand, vacuum pump, calibration buffers pH 7.01 and 1.07, 4 pack AA lithium batteries, Gastro Workstation - Standing)	$28,139
Bravo-in-a-box (includes 2 Bravo receivers, POLYGRAM NET pH Testing Application software, Datalink, calibration stand, vacuum pump, calibration buffers pH 7.01 and 1.07, 4 pack AA lithium batteries, dedicated notebook computer, HP Desk Jet color printer)	$25,704
Filter	$4.50
System tubing (canister-to-filter)	$6
Patient tubing (Bravo delivery system-to-filter)	$4.50

References

Agrawal, D, Akerman, PA & Rich, H 2009, 'Removal of a Bravo 24-hour pH capsule with endoscopic scissors', *Gastrointestinal Endoscopy*, vol. 70, no. 2, pp. 385–386.

Ahlawat, SK, Novak, DJ, Williams, DC, Maher, KA, Barton, F & Benjamin, SB 2006, 'Day-to-day variability in acid reflux patterns using the BRAVO pH-monitoring system', *Journal of Clinical Gastroenterology*, vol. 40, no. 1, pp. 20–24.

Al Talalwah, N & Woodward, S 2013, 'Gastro-oesophageal reflux. Part 3: medical and surgical treatment', *British Journal of Nursing*, vol. 22, no. 7, pp. 409–415.

Allgood, PC & Bachmann, M 2000, 'Medical or surgical treatment for chronic gastrooesophageal reflux? A systematic review of published evidence of effectiveness', *European Journal of Surgery*, vol. 166, no. 9, pp. 713–721.

Andrews, CN, Sadowski, DC, Lazarescu, A, Williams, C, Neshev, E, Storr, M, Au, F & Heitman, SJ 2012, 'Unsedated peroral wireless pH capsule placement vs standard pH testing: a randomized study and cost analysis', *BMC Gastroenterology*, vol. 12, p. 58.

Ang, D, Teo, EK, Ang, TL, Ong, J, Poh, CH, Tan, J & Fock, KM 2010, 'To Bravo or not? A comparison of wireless esophageal pH monitoring and conventional pH catheter to evaluate non-erosive gastroesophageal reflux disease in a multiracial Asian cohort', *Journal of Digestive Diseases*, vol. 11, no. 1, pp. 19–27.

Anvari, M, Allen, C, Marshall, J, Armstrong, D, Goeree, R, Ungar, W & Goldsmith, C 2011, 'A randomized controlled trial of laparoscopic Nissen fundoplication versus proton pump inhibitors for the treatment of patients with chronic gastroesophageal reflux disease (GERD): 3-year outcomes', *Surgical Endoscopy*, vol. 25, no. 8, pp. 2547–2554.

Arguedas, MR, Heudebert, GR, Klapow, JC, Centor, RM, Eloubeidi, MA, Wilcox, CM & Spechler, SJ 2004, 'Re-examination of the cost-effectiveness of surgical versus medical therapy in patients with gastroesophageal reflux disease: the value of long-term data collection', *American Journal of Gastroenterology*, vol. 99, no. 6, pp. 1023–1028.

Armitage, P, Berry, G & Matthews, JNS 2002, *Statistical methods in medical research*, 4th edn, Blackwell Science, Oxford.

Auvin, S, Hartman, AL, Desnous, B, Moreau, AC, Alberti, C, Delanoe, C, Romano, A, Terrone, G, Kossoff, EH, Del Giudice, E & Titomanlio, L 2012, 'Diagnosis delay in West syndrome: misdiagnosis and consequences', *European Journal of Pediatrics*, vol. 171, no. 11, pp. 1695–1701.

Ayazi, S, Hagen, JA, Zehetner, J, Banki, F, Augustin, F, Ayazi, A, DeMeester, SR, Oh, DS, Sohn, HJ, Lipham, JC & DeMeester, TR 2011, 'Day-to-day discrepancy in Bravo pH monitoring is related to the degree of deterioration of the lower esophageal sphincter and severity of reflux disease', *Surgical Endoscopy*, vol. 25, no. 7, pp. 2219–2223.

Ayazi, S, Lipham, JC, Portale, G, Peyre, CG, Streets, CG, Leers, JM, Demeester, SR, Banki, F, Chan, LS, Hagen, JA & Demeester, TR 2009, 'Bravo catheter-free pH monitoring: normal values, concordance, optimal diagnostic thresholds, and accuracy', *Clinical Gastroenterology and Hepatology*, vol. 7, no. 1, pp. 60–67.

Azzam, RS, Sallum, RA, Brandao, JF, Navarro-Rodriguez, T & Nasi, A 2012, 'Comparative study of two modes of gastroesophageal reflux measuring: conventional esophageal pH monitoring and wireless pH monitoring', *Arquivos de Gastroenterologia*, vol. 49, no. 2, pp. 107–112.

Bandolier 1999, Bandolier editorial: Diagnostic testing emerging from the gloom? *Bandolier*, vol. 70, http://www.medicine.ox.ac.uk/bandolier/band70/b70-5.html

Bansal, A, Wani, S, Rastogi, A, Rastogi, K, Goyal, A, Hall, S, Singh, V, Higbee, A & Sharma, P 2009, 'Impact of measurement of esophageal acid exposure close to the gastroesophageal junction on diagnostic accuracy and event-symptom correlation: a prospective study using wireless dual pH monitoring', *American Journal of Gastroenterology*, vol. 104, no. 12, pp. 2918–2925.

Bechtold, ML, Holly, JS, Thaler, K & Marshall, JB 2007, 'Bravo (wireless) ambulatory esophageal pH monitoring: how do day 1 and day 2 results compare?', *World Journal of Gastroenterology*, vol. 13, no. 30, pp. 4091–4095.

Beck, PE, Watson, DI, Devitt, PG, Game, PA & Jamieson, GG 2009, 'Impact of gender and age on the long-term outcome of laparoscopic fundoplication', *World Journal of Surgery*, vol. 33, no. 12, pp. 2620–2626.

Belafsky, PC, Allen, K, Castro-Del Rosario, L & Roseman, D 2004, 'Wireless pH testing as an adjunct to unsedated transnasal esophagoscopy: the safety and efficacy of transnasal telemetry capsule placement', *Otolaryngology–Head & Neck Surgery*, vol. 131, no. 1, pp. 26–28.

Bhat, YM, McGrath, KM & Bielefeldt, K 2006, 'Wireless esophageal pH monitoring: new technique means new questions', *Journal of Clinical Gastroenterology*, vol. 40, no. 2, pp. 116–121.

Bojke, L, Hornby, E & Sculpher, M 2007, 'A comparison of the cost effectiveness of pharmacotherapy or surgery (laparoscopic fundoplication) in the treatment of GORD', *Pharmacoeconomics*, vol. 25, no. 10, pp. 829–841.

Bothwell, M, Phillips, J & Bauer, S 2004, 'Upper esophageal pH monitoring of children with the Bravo pH capsule', *Laryngoscope*, vol. 114, no. 4, pp. 786–788.

Bradley, AG, Crowell, MD, DiBaise, JK, Kim, HJ, Burdick, GE, Fleischer, DE & Sharma, VK 2011, 'Comparison of the impact of wireless versus catheter-based pH-metry on daily activities and study-related symptoms', *Journal of Clinical Gastroenterology*, vol. 45, no. 2, pp. 100-106.

Broeders, JA, Draaisma, WA, Bredenoord, AJ, Smout, AJ, Broeders, IA & Gooszen, HG 2010, 'Long-term outcome of Nissen fundoplication in non-erosive and erosive gastro-oesophageal reflux disease', *British Journal of Surgery*, vol. 97, no. 6, pp. 845–852.

Cabrera, J, Davis, M, Horn, D, Pfefferkorn, M & Croffie, JM 2011, 'Esophageal pH monitoring with the BRAVO capsule: Experience in a single tertiary medical center', *Journal of Pediatric Gastroenterology & Nutrition*, vol. 53, no. 4, pp. 404–408.

Calabrese, C, Liguori, G, Gabusi, V, Gionchetti, P, Rizzello, F, Straforini, G, Brugnera, R & Di Febo, G 2008, 'Ninety-six-hour wireless oesophageal pH monitoring following proton pump inhibitor administration in NERD patients', *Alimentary Pharmacology & Therapeutics*, vol. 28, no. 2, pp. 250-255.

Chander, B, Hanley-Williams, N, Deng, Y & Sheth, A 2012, '24 Versus 48-hour Bravo pH monitoring', *Journal of Clinical Gastroenterology*, vol. 46, no. 3, pp. 197–200.

Chotiprashidi, P, Liu, J, Carpenter, S, Chuttani, R, DiSario, J, Hussain, N, Somogyi, L & Petersen, BT 2005, 'ASGE Technology Status Evaluation Report: wireless esophageal pH-monitoring system', *Gastrointestinal Endoscopy*, vol. 62, no. 4, pp. 485–487.

Comay, D, Adam, V, da Silveira, EB, Kennedy, W, Mayrand, S & Barkun, AN 2008, 'The Stretta procedure versus proton pump inhibitors and laparoscopic Nissen fundoplication in the management of gastroesophageal reflux disease: a cost-effectiveness analysis', *Canadian Journal of Gastroenterology*, vol. 22, no. 6, pp. 552–558.

Croffie, JM, Fitzgerald, JF, Molleston, JP, Gupta, SK, Corkins, MR, Pfefferkorn, MD, Lim, JR, Steiner, SJ & Dadzie, SK 2007, 'Accuracy and tolerability of the Bravo catheter-free pH capsule in patients between the ages of 4 and 18 years', *Journal of Pediatric Gastroenterology & Nutrition*, vol. 45, no. 5, pp. 559–563, DOI 10.1097/MPG.0b013e3180dc9349, <http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/189/CN-00621189/frame.html>.

Crowell, MD, Bradley, A, Hansel, S, Dionisio, P, Kim, HJ, Decker, GA, DiBaise, JK & Sharma, VK 2009, 'Obesity is associated with increased 48-h esophageal acid exposure in patients with symptomatic gastroesophageal reflux', *American Journal of Gastroenterology*, vol. 104, no. 3, pp. 553–559.

de Hoyos, A & Esparza, EA 2010, 'Technical problems produced by the Bravo pH test in nonerosive reflux disease patients', *World Journal of Gastroenterology*, vol. 16, no. 25, pp. 3183–3186.

de Hoyos, A, Esparza, EA & Loredo, ML 2009, 'Cold and hot snare endoscopic techniques for removal of the Bravo pH-monitoring capsule', *Digestion*, vol. 79, no. 1, pp. 14–16.

Deeks, JJ 2001, 'Systematic reviews of evaluations of diagnostic and screening tests', *British Medical Journal*, vol. 323, no. 21, pp. 157–162.

Dent, J, El-Serag, HB, Wallander, MA & Johansson, S 2005, 'Epidemiology of gastro-oesophageal reflux disease: a systematic review', *Gut*, vol. 54, no. 5, pp. 710–717.

des Varannes, SB, Mion, F, Ducrotte, P, Zerbib, F, Denis, P, Ponchon, T, Thibault, R & Galmiche, JP 2005, 'Simultaneous recordings of oesophageal acid exposure with conventional pH monitoring and a wireless system (Bravo)', *Gut*, vol. 54, no. 12, pp. 1682–1686.

Doma, S, Paladugu, S, Parkman, HP & Friedenberg, FK 2010, 'Wireless capsules for esophageal pH monitoring: are we placing them correctly?', *Digestion*, vol. 82, no. 1, pp. 54–59.

Domingues, GR, Moraes-Filho, JP & Domingues, AG 2011, 'Impact of prolonged 48-h wireless capsule esophageal pH monitoring on diagnosis of gastroesophageal reflux disease and evaluation of the relationship between symptoms and reflux episodes', *Arquivos de Gastroenterologia*, vol. 48, no. 1, pp. 24–29.

Downs, SH & Black, N 1998, 'The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions', *Journal of Epidemiology and Community Health*, vol. 52, no. 6, pp. 377–384.Eckardt, VF, Dilling, B & Bernhard, G 1999, 'The impact of open access 24-h pH-metry on the diagnosis and management of esophageal reflux disease', *American Journal of Gastroenterology*, vol. 94, no. 3, pp. 616–621.

Epstein D, Bojke L, Sculpher M & the REFLUX trial group. 2009, *Laparoscopic fundoplication compared with medical management for gastro-oesphageal reflux disease: cost effectiveness study*, BMJ, vol. 339, p. b2576.

Eriksson, S, Langstrom, G, Rikner, L, Carlsson, R & Naesdal, J 1995, 'Omeprazole and H2-receptor antagonists in the acute treatment of duodenal ulcer, gastric ulcer and reflux oesophagitis: a meta-analysis', *European Journal of Gastroenterology & Hepatoogyl*, vol. 7, no. 5, pp. 467–475.

Fajardo, NR, Wise, JL, Locke, GR, III, Murray, JA & Talley, NJ 2006, 'Esophageal perforation after placement of wireless Bravo pH probe', *Gastrointestinal Endoscopy*, vol. 63, no. 1, pp. 184–185.

Fass, R 2009, 'Proton pump inhibitor failure: what are the therapeutic options?', *American Journal of Gastroenterology*, vol. 104, no. S2, pp. S33–S38.

Fox, M 2006, 'Bravo wireless versus catheter pH-monitoring systems', *Gut*, vol. 55, no. 3, pp. 434–435.

Francis, DL 2008, 'Attachment disorder: a decline in the performance of the Bravo pH system', *American Journal of Gastroenterology*, vol. 103, no. 10, p. 2663.

Francis, DO, Sumner, E, Goutte, M, Slaughter, JC, Garrett, CG, Hagaman, D & Vaezi, MF 2012, 'Feasibility of dual wireless esophageal pH monitoring: randomized blinded sham-controlled trial', *Otolaryngology–Head and Neck Surgery*, vol. 147, no. 1, pp. 91-97.

Galmiche, JP, Hatlebakk, J, Attwood, S, Ell, C, Fiocca, R, Eklund, S, Langstrom, G, Lind, T, Lundell, L & Collaborators, LT 2011, 'Laparoscopic antireflux surgery vs esomeprazole treatment for chronic GERD: the LOTUS randomized clinical trial', *Journal of the American Medical Association*, vol. 305, no. 19, pp. 1969–1977.

Garrean, CP, Zhang, Q, Gonsalves, N & Hirano, I 2008, 'Acid reflux detection and symptom-reflux association using 4-day wireless pH recording combining 48-hour periods off and on PPI therapy', *American Journal of Gastroenterology*, vol. 103, no. 7, pp. 1631–1637.

Gastroenterological Society of Australia 2011, *Gastro-oesophageal Reflux Disease in Adults* 5th edn) Gastroenterological Society of Australia, Digestive Health Foundation.

Gastroenterological Society of Australia 2013, *Heartburn (reflux)*, Gasteroenterological Society of Australia, viewed 21 May 2013, <http://www.gesa.org.au/consumer.asp?cid=7&id=79>

Gawron, AJ, Rothe, J, Fought, AJ, Fareeduddin, A, Toto, E, Boris, L, Kahrilas, PJ & Pandolfino, JE 2012, 'Many patients continue using proton pump inhibitors after negative results from tests for reflux disease', *Clinical Gastroenterology and Hepatol*,ogy vol. 10, no. 6, pp. 620–625; quiz e657.

Gillies, RS, Stratford, JM, Booth, MI & Dehn, TC 2007, 'Oesophageal pH monitoring using the Bravo catheter-free radio capsule', *European Journal of Gastroenterology & Hepatology*, vol. 19, no. 1, pp. 57–63.

Grant, AM, Boachie, C, Cotton, SC, Faria, R, Bojke, L, Epstein, DM, Ramsay, CR, Corbacho, B, Sculpher, M, Krukowski, ZH, Heading, RC & Campbell, MK 2013, 'Clinical and economic evaluation of laparoscopic surgery compared with medical management for gastro-oesophageal reflux disease: 5-year follow-up of multicentre randomised trial (the REFLUX trial)', *Health Technology Assessment*, vol. 17, no. 22, pp. 1–167.

Grant, AM, Cotton, SC, Boachie, C, Ramsay, CR, Krukowski, ZH, Heading, RC, Campbell, MK & Group, RT 2013, 'Minimal access surgery compared with medical management for gastro-oesophageal reflux disease: five year follow-up of a randomised controlled trial (REFLUX)', *BMJ*, vol. 346, p. f1908.

Grant, AM, Wileman, S, Ramsay, C, Bojke, L, Epstein, D, Sculpher, M, Macran, S, Kilonzo, M, Vale, L, Francis, J, Mowat, A, Krukowski, Z, Heading, R, Thursz, M, Russell, I & Campbell, M 2008, 'The effectiveness and cost-effectiveness of minimal access surgery amongst people with gastro-oesophageal reflux disease - a UK collaborative study. The REFLUX trial', *Health Technology Assessment*, vol. 12, no. 31, pp. 1–181, iii–iv.

Grigolon, A, Bravi, I, Cantu, P, Conte, D & Penagini, R 2007, 'Wireless pH monitoring: better tolerability and lower impact on daily habits', *Digestive & Liver Disease*, vol. 39, no. 8, pp. 720–724.

Grigolon, A, Consonni, D, Bravi, I, Tenca, A & Penagini, R 2011, 'Diagnostic yield of 96-h wireless pH monitoring and usefulness in patients' management', *Scandinavian Journal of Gastroenterology*, vol. 46, no. 5, pp. 522–530.

Gunaratnam, NT, Jessup, TP, Inadomi, J & Lascewski, DP 2006, 'Sub-optimal proton pump inhibitor dosing is prevalent in patients with poorly controlled gastro-oesophageal reflux disease', *Alimentary Pharmacology & Therapeutics*, vol. 23, no. 10, pp. 1473–1477.

Gunnarsdottir, A, Stenstrom, P & Arnbjornsson, E 2007, '48-hour wireless oesophageal pH-monitoring in children: are two days better than one?', *European Journal of Pediatric Surgery*, vol. 17, no. 6, pp. 378–381.

—— 2008, 'Wireless esophageal pH monitoring in children', *Journal of Laparoendoscopic & Advanced Surgical Techniques A*, vol. 18, no. 3, pp. 443–447.

Hakanson, BS, Berggren, P, Granqvist, S, Ljungqvist, O & Thorell, A 2009, 'Comparison of wireless 48-h (Bravo) versus traditional ambulatory 24-h esophageal pH monitoring', *Scandinavian Journal of Gastroenterol*,ogy vol. 44, no. 3, pp. 276–283.

Heidelbaugh, JJ, Goldberg, KL & Inadomi, JM 2010, 'Magnitude and economic effect of overuse of antisecretory therapy in the ambulatory care setting', *American Journal of Managed Care*, vol. 16, no. 9, pp. e228–234.

Heidelbaugh, JJ, Kim, AH, Chang, R & Walker, PC 2012, 'Overutilization of proton-pump inhibitors: what the clinician needs to know', *Therapeutic Advances in Gastroenterol*,ogy vol. 5, no. 4, pp. 219–232.

Heudebert, GR, Marks, R, Wilcox, CM & Centor, RM 1997, 'Choice of long-term strategy for the management of patients with severe esophagitis: a cost-utility analysis', *Gastroenterology*, vol. 112, no. 4, pp. 1078–1086.

Hirano, I 2006, 'Review article: modern technology in the diagnosis of gastro-oesophageal reflux disease—Bilitec, intraluminal impedance and Bravo capsule pH monitoring', *Alimentary Pharmacology & Therapeutics*, vol. 23 Suppl 1, pp. 12–24.

Hirano, I, Richter, JE & Practice Parameters Committee of the American College of Gastroenterology 2007, 'ACG practice guidelines: esophageal reflux testing', *American Journal of Gastroenterology*, vol. 102, no. 3, pp. 668–685.

Hirano, I, Zhang, Q, Pandolfino, JE & Kahrilas, PJ 2005, 'Four-day Bravo pH capsule monitoring with and without proton pump inhibitor therapy', *Clinical Gastroenterology and Hepatology*, vol. 3, no. 11, pp. 1083–1088.

Hochman, JA & Favaloro-Sabatier, J 2005, 'Tolerance and reliability of wireless pH monitoring in children', *Journal of Pediatric Gastroenterology & Nutrition*, vol. 41, no. 4, pp. 411–415.

Hogan, RB 2nd, Phillips, P, Boyd, SA & Williams, JC 2009, 'Two-year retention of Bravo capsule in a giant colonic diverticulum', *American Journal of Gastroenterology*, vol. 104, no. 4, p. 1062.

IFFGD 2014, *About GERD*, International Foundation for Functional Gastrointestinal Disorders, viewed 2 August 2013, <http://www.aboutgerd.org/>.

Iqbal, A, Lee, YK, Vitamvas, M & Oleynikov, D 2007, '48-hour pH monitoring increases the risk of false positive studies when the capsule is prematurely passed', *Journal of Gastrointestinal Surgery*, vol. 11, no. 5, pp. 638–641.

Kahrilas, PJ 2008, 'Clinical practice. Gastroesophageal reflux disease', *New England Journal of Medicine*, vol. 359, no. 16, pp. 1700–1707.

Kahrilas, PJ, Shaheen, NJ & Vaezi, MF 2008, 'American Gastroenterological Association Institute technical review on the management of gastroesophageal reflux disease', *Gastroenterology*, vol. 135, no. 4, pp. 1392–1413, e1395.

Karamanolis, G, Triantafyllou, K, Psatha, P, Vlachogiannakos, I, Gaglia, A, Polymeros, D, Fessatou, S, Triantafyllou, M, Papanikolaou, IS & Ladas, SD 2012, 'Bravo 48-hour wireless pH monitoring in patients with non-cardiac chest pain: objective gastroesophageal reflux disease parameters predict the responses to proton pump inhibitors', *Journal of Neurogastroenterology & Motility*, vol. 18, no. 2, pp. 169–173.

Katz, PO, Gerson, LB & Vela, MF 2012, 'Guidelines for the diagnosis and management of gastroesophageal reflux disease', *American Journal of Gastroenterology*, vol. 108, no. 3, pp. 308–328.

Kessing, BF, Bredenoord, AJ & Smout, AJPM 2011, 'Erroneous diagnosis of gastroesophageal reflux disease in achalasia', *Clinical Gastroenterology and Hepatology*, vol. 9, no. 12, pp. 1020–1024.

Khan, KS, Ter Riet, G, Glanville, JM, Sowden, AJ & Kleijnen, J 2001, *Undertaking systematic reviews of research on effectiveness. CRD's guidance for those carrying out or commissioning reviews*, CRD Report no. 4 (2nd edn), NHS Centre for Reviews and Dissemination, University of York, York.

Kleiman, DA, Beninato, T, Bosworth, BP, Brunaud, L, Ciecierega, T, Crawford, CV, Jr., Turner, BG, Fahey, TJ, 3rd & Zarnegar, R 2013, 'Early referral for esophageal pH monitoring is more cost-effective than prolonged empiric trials of proton-pump inhibitors for suspected gastroesophageal reflux disease', *Journal of Gastrointestinal Surgery*, vol. 18, no. 1, pp. 26–33; discussion 33–34.

Kleiman, DA, Sporn, MJ, Beninato, T, Metz, Y, Crawford, C, Fahey, TJ, 3rd & Zarnegar, R 2012, 'Early referral for 24-h esophageal pH monitoring may prevent unnecessary treatment with acid-reducing medications', *Surgical Endoscopy*, vol. 27, no. 4, pp.1302–1309.

Knox, SA, Harrison, CM, Britt, HC & Henderson, JV 2008, 'Estimating prevalence of common chronic morbidities in Australia', *Medical Journal of Australia*, vol. 189, no. 2, pp. 66–70.

Korrapati, V, Babich, JP, Balani, A, Grendell, JH & Kongara, KR 2011, 'Does deep sedation impact the results of 48 hours catheterless pH testing?', *World Journal of Gastroenterology*, vol. 17, no. 10, pp. 1304–1307.

Kramer, E & Chokhavatia, S 2012, 'An unusual complication of Bravo pH capsule dislodged into the pyriform sinus', *American Journal of Gastroenterology*, vol. 107, p. S497.

Kushnir, VM, Sayuk, GS & Gyawali, CP 2011, 'The effect of antisecretory therapy and study duration on ambulatory esophageal pH monitoring', *Digestive Diseases & Sciences*, vol. 56, no. 5, pp. 1412–1419.

Lacy, BE, Chehade, R & Crowell, MD 2011, 'A prospective study to compare a symptom-based reflux disease questionnaire to 48-h wireless pH monitoring for the identification of gastroesophageal reflux (revised 2-26-11)', *American Journal of Gastroenterology*, vol. 106, no. 9, pp. 1604–1611.

Lacy, BE, Dukowicz, AC, Robertson, DJ, Weiss, JE, Teixeira, P & Kelley, ML, Jr. 2011, 'Clinical utility of the wireless pH capsule', *Journal of Clinical Gastroenterology*, vol. 45, no. 5, pp. 429–435.

Lacy, BE, Edwards, S, Paquette, L, Weiss, J, Kelley, ML, Jr. & Ornvold, K 2009, 'Tolerability and clinical utility of the Bravo pH capsule in children', *Journal of Clinical Gastroenterology*, vol. 43, no. 6, pp. 514–519.

Lagergren, J, Bergstrom, R, Lindgren, A & Nyren, O 1999, 'Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma', *New England Journal of Medicine*, vol. 340, no. 11, pp. 825–831.

Lee, JS 2012, 'Is wireless capsule pH monitoring better than catheter systems?', *Journal of Neurogastroenterology & Motility*, vol. 18, no. 2, pp. 117–119.

Lee, WC, Yeh, YC, Lacy, BE, Pandolfino, JE, Brill, JV, Weinstein, ML, Carlson, AM, Williams, MJ, Wittek, MR & Pashos, CL 2008, 'Timely confirmation of gastro-esophageal reflux disease via pH monitoring: estimating budget impact on managed care organizations', *Current Medical Research & Opinion*, vol. 24, no. 5, pp. 1317–1327.

Lee, YC, Wang, HP, Chiu, HM, Huang, SP, Tu, CH, Wu, MS & Lin, JT 2005, 'Patients with functional heartburn are more likely to report retrosternal discomfort during wireless pH monitoring', *Gastrointestinal Endoscopy*, vol. 62, no. 6, pp. 834–841.

Liberati, A, Altman, DG, Tetzlaff, J, Mulrow, C, Gotzsche, PC, Ioannidis, JP, Clarke, M, Devereaux, PJ, Kleijnen, J & Moher, D 2009, 'The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration', *BMJ*, vol. 339, p. b2700.

Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JHP & Bossuyt PMM 1999, Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*, vol. 282, no. 11, pp. 1061–1066.

Maerten, P, Ortner, M, Michetti, P & Dorta, G 2007, 'Wireless capsule pH monitoring: does it fulfil all expectations?', *Digestion*, vol. 76, no. 3–4, pp. 235–240.

Mahon, D, Rhodes, M, Decadt, B, Hindmarsh, A, Lowndes, R, Beckingham, I, Koo, B & Newcombe, RG 2005, 'Randomized clinical trial of laparoscopic Nissen fundoplication compared with proton-pump inhibitors for treatment of chronic gastro-oesophageal reflux', *British Journal of Surgery*, vol. 92, no. 6, pp. 695–699.

Marchese, M, Spada, C, Iacopini, F, Familiari, P, Shah, SG, Tringali, A & Costamagna, G 2006, 'Nonendoscopic transnasal placement of a wireless capsule for esophageal pH monitoring: feasibility, safety, and efficacy of a manometry-guided procedure', *Endoscopy*, vol. 38, no. 8, pp. 813–818.

Martinez de Haro, LF, Munitiz, V, Ortiz, A, Ruiz de Angulo, D, Navarro, MD & Parrilla, P 2008, '[Outpatient monitoring of oesophageal pH with a catheter-free pH-meter (Bravo system) a study of tolerance, safety and efficacy]', Cirugía Española, vol. 84, no. 4, pp. 201–209.

Mearin, F, Balboa, A, Dot, J, Maldonado, O & Malagelada, JR 1998, 'How standard is a standard day during a standard ambulatory 24-hour esophageal pH monitoring?', *Scandinavian Journal of Gastroenterology*, vol. 33, no. 6, pp. 583–585.

Medicare Australia 2013, *Medicare statistics, Medicare Locals Statistics Report*, Australian Government, Department of Human Services, viewed 4 February 2014, https://www.medicareaustralia.gov.au/statistics/med_locals.shtml.

Miller, G & Pan, Y 2009, 'Gastro-oesophageal reflux disease', in *Britt et al.* (eds), *General practice in Australia, health priorities and policies 1998 to 2008*, Cat. no. GEP 24, Australian Institute of Health and Welfare, Canberra.

MSAC 2005, *Guidelines for the assessment of diagnostic technologies*, Medical Services Advisory Committee, Commonwealth of Australia, Canberra, ACT.

Mulherin, SA & Miller, WC 2002. 'Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation', *Annals of Internal Medicine,* vol. 137, no. 7, pp. 598–602.

Napuri, S, Le Gall, E, Dulac, O, Chaperon, J & Riou, F 2010, 'Factors associated with treatment lag in infantile spasms', *Developmental Medicine and Child Neurology*, vol. 52, no. 12, pp. 1164–1166.

Nelson, SP, Chen, EH, Syniar, GM & Christoffel, KK 2000, 'Prevalence of symptoms of gastroesophageal reflux during childhood: a pediatric practice-based survey. Pediatric Practice Research Group', *Archives of Pediatric & Adolescent Medicine*, vol. 154, no. 2, pp. 150–154.

Netzer, P, Gut, A, Heer, R, Gries, N, Pfister, M, Halter, F & Inauen, W 1999, 'Five-year audit of ambulatory 24-hour esophageal pH-manometry in clinical practice', *Scandinavian Journal of Gastroenterology*, vol. 34, no. 7, pp. 676–682.

National Health and Medical Research Council (NHMRC) 1999, *A guide to the development, implementation and evaluation of clinical practice guidelines*, NHMRC, Canberra, Australia.

NHMRC 2000, *How to use the evidence: assessment and application of scientific evidence*, Handbook series on preparing clinical practice guidelines, National Health and Medical Research Council, Canberra.

—— 2008, *NHMRC additional levels of evidence and grades for recommendations for developers of guidelines. Stage 2 consultation*, National Health and Medical Research Council, Australian Government, viewed 11 March 2008, <www.nhmrc.gov.au/consult/index.htm>.

Nusrat, S, Roy, PM & Bielefeldt, K 2012, 'Wireless ambulatory pH studies: manometric or endoscopic guidance?', *Diseases of the Esophagus*, vol. 25, no. 1, pp. 26–32.

Pandolfino, JE, Lee, TJ, Schreiner, MA, Zhang, Q, Roth, MP & Kahrilas, PJ 2006, 'Comparison of esophageal acid exposure at 1 cm and 6 cm above the squamocolumnar junction using the Bravo pH-monitoring system', *Diseases of the Esophagus*, vol. 19, no. 3, pp. 177–182.

Pandolfino, JE, Richter, JE, Ours, T, Guardino, JM, Chapman, J & Kahrilas, PJ 2003, 'Ambulatory esophageal pH monitoring using a wireless system', *American Journal of Gastroenterology*, vol. 98, no. 4, pp. 740–749.

Pandolfino, JE, Schreiner, MA, Lee, TJ, Zhang, Q, Boniquit, C & Kahrilas, PJ 2005, 'Comparison of the Bravo (TM) wireless and Digitrapper (TM) catheter-based pH-monitoring systems for measuring esophageal acid exposure', *American Journal of Gastroenterology*, vol. 100, no. 7, pp. 1466–1476.

Park, EY, Choi, MG, Baeg, M, Lim, CH, Kim, J, Cho, Y, Park, J, Lee, I, Kim, S & Choi, K 2013, 'The value of early wireless esophageal pH monitoring in diagnosing functional heartburn in refractory gastroesophageal reflux disease', *Digestive Diseases & Sciences*, vol. 58, no. 10, pp. 2933–2939.

Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B & Dawes M 2001, Oxford Centre for Evidence-Based Medicine levels of evidence. Centre for Evidence-Based Medicine, Oxford, UK.

Prakash, C & Clouse, RE 2005, 'Value of extended recording time with wireless pH monitoring in evaluating gastroesophageal reflux disease', *Clinical Gastroenterology and Hepatology*, vol. 3, no. 4, pp. 329–334.

—— 2006, 'Wireless pH monitoring in patients with non-cardiac chest pain', *American Journal of Gastroenterology*, vol. 101, no. 3, pp. 446–452.

Prakash, C, Jonnalagadda, S, Azar, R & Clouse, RE 2006, 'Endoscopic removal of the wireless pH-monitoring capsule in patients with severe discomfort', *Gastrointestinal Endoscopy*, vol. 64, no. 5, pp. 828–832.

Remes-Troche, JM, Ibarra-Palomino, J, Carmona-Sanchez, RI & Valdovinos, MA 2005, 'Performance, tolerability, and symptoms related to prolonged pH monitoring using the Bravo system in Mexico', *American Journal of Gastroenterology*, vol. 100, no. 11, pp. 2382–2386.

Renteln, D, Kayser, T, Riecken, B & Caca, K 2008, 'An unusual case of Bravo capsule aspiration', *Endoscopy*, vol. 40 Suppl 2, Sep, p. E174.

Sackett, DL & Haynes, RB 2002, 'The architecture of diagnostic research', *BMJ,* vol. 324, no. 7336, pp. 539–541.

Scarpignato, C 2012, 'Poor effectiveness of proton pump inhibitors in non-erosive reflux disease: the truth in the end!', *Neurogastroenterology & Motility*, vol. 24, no. 8, pp. 697–704.

Scarpulla, G, Camilleri, S, Galante, P, Manganaro, M & Fox, M 2007, 'The impact of prolonged pH measurements on the diagnosis of gastroesophageal reflux disease: 4-day wireless pH studies', *American Journal of Gastroenterology*, vol. 102, no. 12, pp. 2642–2647.

Schneider, JH, Kramer, KM, Konigsrainer, A & Granderath, FA 2007, 'Ambulatory pH: monitoring with a wireless system', *Surgical Endoscopy*, vol. 21, no. 11, pp. 2076–2080.

Shaheen, NJ & Richter, JE 2009, 'Barrett's oesophagus', *Lancet*, vol. 373, no. 9666, pp. 850–861.

Shahid, S & Fisher, R 2011, 'Left, right and out: esophageal probe aspiration', *American Journal of Gastroenterology*, vol. 106, pp. S378–S379.

Sofi, A, Filipiak, C, Sodeman, T, Ahmad, U, Bawany, M & Daboul, I 2011, 'Comparison of esophageal placement of Bravo capsule system under direct endoscopic guidance with conventional placement method', *Turkish Journal of Gastroenterology*, vol. 22, no. 3, pp. 243–248.

Starr, DS 2006, 'Legal advisor. Case #1: misdiagnosis of smoker's chest pain', *Clinical Advisor for Nurse Practitioners*, vol. 9, no. 1, pp. 74–75.

—— 2007, 'Legal advisor. Case #1: past GERD hides critical new ailment', *Clinical Advisor for Nurse Practitioners*, vol. 10, no. 3, pp. 141–142.

Susman, JL 2009, ' "GERD" turns out to be heart disease', *The Journal of Family Practice*, vol. 58, no. 5, p. 288.

Sweetman, LL, Ng, YT & Kerrigan, JF 2007, 'Gelastic seizures misdiagnosed as gastroesophageal reflux disease', *Clinical Pediatrics*, vol. 46, no. 4, pp. 325–328.

Sweis, R, Fox, M, Anggiansah, A & Wong, T 2011, 'Prolonged, wireless pH-studies have a high diagnostic yield in patients with reflux symptoms and negative 24-h catheter-based pH-studies', *Neurogastroenterology & Motility*, vol. 23, no. 5, pp. 419–426.

Sweis, R, Fox, M, Anggiansah, R, Anggiansah, A, Basavaraju, K, Canavan, R & Wong, T 2009, 'Patient acceptance and clinical impact of Bravo monitoring in patients with previous failed catheter-based studies', *Alimentary Pharmacology & Therapeutics*, vol. 29, no. 6, pp. 669–676.

Taddio, A, Bersanini, C, Basile, L, Fontana, M & Ventura, A 2011, 'Gastroesophageal reflux disease at any cost: a dangerous paediatric attitude', *Acta Paediatrica, International Journal of Paediatrics*, vol. 100, no. 10, pp. e178–e180.

Tankurt, E, Tuzomay, N, Sengul, B, Kinik, E & Acil, B 2011, 'Safety and tolerability of BRAVO capsule pH monitoring', *Turkish Journal of Gastroenterology*, vol. 22, no. 3, pp. 352–353.

Tharavej, C, Hagen, JA, Portale, G, Hsieh, CC, Gandamihardja, TA, Lipham, JC, Peters, JH, DeMeester, SR, Crookes, PF, Bremner, CG & DeMeester, TR 2006, 'Bravo capsule induction of esophageal hypercontractility and chest pain', *Surgical Endoscopy*, vol. 20, no. 5, pp. 783–786.

Triester, SL, Leighton, JA, Budavari, AI, Crowell, MD & Fleischer, DE 2005, 'Severe chest pain from an indwelling Bravo pH probe', *Gastrointestinal Endoscopy*, vol. 61, no. 2, pp. 317–319.

Tseng, D, Rizvi, AZ, Fennerty, MB, Jobe, BA, Diggs, BS, Sheppard, BC, Gross, SC, Swanstrom, LL, White, NB, Aye, RW & Hunter, JG 2005, 'Forty-eight-hour pH monitoring increases sensitivity in detecting abnormal esophageal acid exposure', *Journal of Gastrointestinal Surgery*, vol. 9, no. 8, pp. 1043–1051; discussion 1051–1042.

Tu, CH, Lee, YC, Wang, HP, Wu, MS, Chiu, HM & Lin, JT 2004, 'Ambulatory esophageal pH monitoring by using a wireless system: a pilot study in Taiwan', *Hepatogastroenterology*, vol. 51, no. 60, pp. 1586–1589.

Turner, BG, Saltzman, JR, Hua, L, Maurer, R, Feldman, N, Carr-Locke, DL, Burakoff, R & Liu, JJ 2007, 'Endoscopic pH monitoring for patients with suspected or refractory gastroesophageal reflux disease', *Canadian Journal of Gastroenterology*, vol. 21, no. 11, pp. 737–741.

Tutuian, R 2008, 'Reflux monitoring: current status', *Current Gastroenterology Reports*, vol. 10, no. 3, pp. 263–270.

Ward, EM, Devault, KR, Bouras, EP, Stark, ME, Wolfsen, HC, Davis, DM, Nedrow, SI & Achem, SR 2004, 'Successful oesophageal pH monitoring with a catheter-free system', *Alimentary Pharmacology & Therapeutics*, vol. 19, no. 4, pp. 449–454.

Weber, C, Davis, CS & Fisichella, PM 2011, 'Current applications of evolving methodologies in gastroesophageal reflux disease testing', *Digestive & Liver Disease*, vol. 43, no. 5, pp. 353–357.

Wenner, J, Hall, M, Hoglund, P, Johansson, J, Johnsson, F & Oberg, S 2008, 'Wireless pH recording immediately above the squamocolumnar junction improves the diagnostic performance of esophageal pH studies', *American Journal of Gastroenterology*, vol. 103, no. 12, pp. 2977–2985.

Wenner, J, Johansson, J, Johnsson, F & Oberg, S 2007b, 'Optimal thresholds and discriminatory power of 48-h wireless esophageal pH monitoring in the diagnosis of GERD', *American Journal of Gastroenterology*, vol. 102, no. 9, pp. 1862–1869.

Wenner, J, Johnsson, F, Johansson, J & Oberg, S 2007a, 'Wireless esophageal pH monitoring is better tolerated than the catheter-based technique: results from a randomized cross-over trial', *American Journal of Gastroenterology*, vol. 102, no. 2, pp. 239–245.

Whiting, PF, Rutjes, AW, Reitsma JB, Bossuyt PM & Kleijnen J 2003, 'The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews', *BMC Medical Research Methodology,* vol. 3, p. 25.

Whiting, PF, Rutjes, AW, Westwood, ME, Mallett, S, Deeks, JJ, Reitsma, JB, Leeflang, MM, Sterne, JA & Bossuyt, PM 2011, 'QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies', *Annals of Internal Medicine*, vol. 155, no. 8, pp. 529–536.

Wiener, GJ, Morgan, TM, Copper, JB, Wu, WC, Castell, DO, Sinclair, JW & Richter, JE 1988, 'Ambulatory 24-hour esophageal pH monitoring: reproducibility and variability of pH parameters', *Digestive Diseases & Sciences*, vol. 33, no. 9, pp. 1127–1133.

Wileman SM, McCann, S, Grant Adrian, M, Krukowski Zygmunt, H & Bruce, J 2010, 'Medical versus surgical management for gastro-oesophageal reflux disease (GORD) in adults', *Cochrane Database of Systematic Reviews*, no. 3, DOI 10.1002/14651858.CD003243.pub2.

Wilson, JA & Vela, MF 2008, 'New esophageal function testing (impedance, Bravo pH monitoring, and high-resolution manometry): clinical relevance', *Current Gastroenterology Reports*, vol. 10, no. 3, pp. 222–230.

Wong, WM, Bautista, J, Dekel, R, Malagon, IB, Tuchinsky, I, Green, C, Dickman, R, Esquivel, R & Fass, R 2005, 'Feasibility and tolerability of transnasal/per-oral placement of the wireless pH capsule vs traditional 24-h oesophageal pH monitoring: a randomized trial', *Alimentary Pharmacology Therapeutics*, vol. 21, no. 2, pp. 155–163.