I. Description

Esophageal pH monitoring, using wired or wireless devices, can record the pH of the lower esophagus for a period of several days. Impedance pH monitoring measures electrical impedance in the esophagus to evaluate reflux episodes concurrent with changes in pH. These tests are used for certain clinical indications in the evaluation of gastroesophageal reflux disease (GERD).

For individuals who have GERD who receive catheter-based pH monitoring, the evidence includes various cross-sectional studies evaluating test performance in different populations. Relevant outcomes are test accuracy and validity, symptoms, and functional outcomes. Positive pH monitoring tests correlate with endoscopically defined GERD and with GERD symptoms, but because there is no reference standard for clinical GERD, diagnostic characteristics cannot be determined. There are no studies of clinical utility showing improved outcomes, and the chain of evidence supporting the utility of the test is weak. The evidence is insufficient to determine that the technology improves health outcomes.

For individuals who have GERD who receive wireless pH monitoring, the evidence includes various cross-sectional studies evaluating test performance and diagnostic yield in different populations. Relevant outcomes are test accuracy and validity, symptoms, and functional outcomes. Positive wireless pH monitoring tests correlate with endoscopically defined GERD and GERD symptoms, but because there is no reference standard for clinical GERD, diagnostic characteristics cannot be determined. Some studies have shown higher positive test rates with prolonged wireless monitoring compared with catheter-based pH monitoring, but the effect of this finding on patient outcomes is uncertain. There are no studies of clinical utility showing improved outcomes, and the chain of evidence supporting the utility of the test is weak. The evidence is insufficient to determine that the technology improves health outcomes.

For individuals who have GERD who receive impedance pH testing, the evidence includes various cross-sectional studies evaluating test performance and diagnostic yield in different populations. Relevant outcomes are test accuracy and validity, symptoms, and functional outcomes. Positive
impedance pH tests correlate with endoscopically defined GERD and with GERD symptoms, but because there is no reference standard for clinical GERD, diagnostic characteristics cannot be determined. Some studies have shown higher positive test rates with impedance pH testing compared with pH testing alone, but the effect of this finding on patient outcomes is uncertain. There are no studies of clinical utility showing improved outcomes, and the chain of evidence supporting the utility of the test is weak. The evidence is insufficient to determine that the technology improves health outcomes.

Expert clinical opinion has suggested that catheter-based and wireless pH monitoring may aid in the diagnosis of GERD in patients who have an uncertain diagnosis after clinical evaluation and endoscopy. Esophageal pH monitoring is not considered a standard diagnostic test for most patients with GERD, but there is strong clinical support for its use in selected subpopulations for certain indications. Clinical guidelines support pH testing for patients with GERD being considered for surgical intervention. Wireless pH monitoring measurements appear to correlate closely to catheter-based monitoring and may be more comfortable for patients.

Background

Gastroesophageal Reflux Disease
Acid reflux is the cause of heartburn and acid regurgitation esophagitis, which can lead to esophageal stricture. Acid reflux can also cause or contribute to some cases of asthma, posterior laryngitis, chronic cough, dental erosions, chronic hoarseness, pharyngitis, subglottic stenosis or stricture, nocturnal choking, and recurrent pneumonia.

Diagnosis
Gastroesophageal reflux disease is most commonly diagnosed by clinical evaluation and treated empirically with a trial of medical management. For patients who do not respond appropriately to medications, or who have recurrent chronic symptoms, endoscopy is indicated to confirm the diagnosis and assess the severity of reflux esophagitis. In some patients, endoscopy is nondiagnostic, or results are discordant with the clinical evaluation (in these cases, further diagnostic testing may be of benefit).

Monitoring
Esophageal monitoring is done using a tube with a pH electrode attached to its tip, which is then passed into the esophagus to approximately 5 cm above the upper margin of the lower esophageal sphincter. The electrode is attached to a data recorder worn on a waist belt or shoulder strap. Every instance of acid reflux, as well as its duration and pH, is recorded over a 24-hour period. Wireless pH monitoring is achieved using endoscopic or manometric guidance to attach the pH measuring capsule to the esophageal mucosa using a clip. The capsule records pH levels for up to 96 hours and transmits them via radiofrequency telemetry to a receiver worn on the patient’s belt. Data from the recorder are uploaded to a computer for analysis by a nurse or doctor.

Another technology closely related to pH monitoring is impedance pH monitoring, which incorporates pH monitoring with measurements of impedance, a method of measuring reflux of liquid or gas of any pH. Multiple electrodes are placed along the length of the esophageal catheter.
The impedance pattern detected can determine the direction of flow and the substance (liquid or gas). Impedance monitoring can identify reflux events in which the liquid is only slightly acidic or nonacidic.

**Regulatory Status**

Esophageal pH electrodes are considered class I devices by the U.S. Food and Drug Administration (FDA) and are exempt from 510(k) requirements. A catheter-free, temporarily implanted device (Bravo™ pH Monitoring System; Medtronic [Minneapolis, MN], now Given Imaging) was cleared for marketing by FDA through the 510(k) process for the purpose of “gastroesophageal pH measurement and monitoring of gastric reflux in adults and children from 4 years of age.”

Several wireless and catheter-based (wired) esophageal pH monitoring devices have been cleared for marketing by FDA through the 510(k) process. Examples include the Bravo™ pH Monitoring System (Given Imaging), the Sandhill Scientific PediaTec™ pH Probe (Sandhill Scientific [Highlands Ranch, CO]), the ORION II Ambulatory pH Recorder (MMS, Medical Measurement Systems [the Netherlands]), and the TRIP CIC Catheter (Tonometrics). FDA product code: FFT.

**II. Criteria/Guidelines**

A. Esophageal pH monitoring using a wireless or catheter-based system is covered (subject to Administrative Guidelines) for one or more of the following indications for adults and adolescents or children able to report symptoms:
   1. Documentation of abnormal acid exposure in endoscopy-negative patients being considered for surgical antireflux repair;
   2. Evaluation of patients after antireflux surgery who are suspected of having ongoing abnormal reflux;
   3. Evaluation of patients with either normal or equivocal endoscopic findings and reflux symptoms that are refractory to proton pump inhibitor therapy;
   4. Evaluation of refractory reflux in patients with chest pain after cardiac evaluation and after a 1-month trial of proton pump inhibitor therapy;
   5. Evaluation of suspected otolaryngologic manifestations of gastroesophageal reflux disease (ie, laryngitis, pharyngitis, chronic cough) in patient who have failed to respond to at least 4 weeks of proton pump inhibitor therapy;

B. Twenty-four-hour catheter-based esophageal pH monitoring is covered (subject to Administrative Guidelines) for infants or children who are unable to report or describe symptoms of reflux with any of the following:
   1. Unexplained apnea
   2. Bradycardia
   3. Refractory coughing or wheezing, stridor, or recurrent choking (aspiration)
   4. Persistent or recurrent laryngitis; or
   5. Recurrent pneumonia
C. Catheter-based impedance pH monitoring is not covered for all indications, as it has not been shown to improve health outcomes.

D. Esophageal pH monitoring systems should be used in accordance with FDA-approved indications and age ranges.

III. Administrative Guidelines

A. Precertification is not required for this service. Documentation supporting the medical necessity should be legible, maintained in the patient's medical record and must be made available to HMSA upon request. HMSA reserves the right to perform retrospective review using the above criteria to validate if services rendered met payment determination criteria.

B. Manometry, when used for pH tip placement, is considered part of the pH recording and will not be paid separately.

C. The device may be placed with either endoscopic or manometry guidance. CPT codes 43235 (endoscopy) or 91010 (manometry) might be used, followed on a subsequent day with the code 91034 (nasal catheter) or 91035 (Bravo esophageal pH monitoring), which represents the interpretation of the recorded measurements.

<table>
<thead>
<tr>
<th>CPT Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>91034</td>
<td>Esophagus, gastroesophageal reflux test; with nasal catheter pH electrode(s) placement, recording, analysis and interpretation</td>
</tr>
<tr>
<td>91035</td>
<td>Esophagus, gastroesophageal reflux test; with mucosal attached telemetry pH electrode placement, recording, analysis and interpretation</td>
</tr>
<tr>
<td>91038</td>
<td>Esophageal function test, gastroesophageal reflux test with nasal catheter intraluminal impedance electrode(s) placement, recording, analysis and interpretation; prolonged (greater than 1 hour, up to 24 hours)</td>
</tr>
</tbody>
</table>

IV. Scientific Background

This policy was created in July 1996 and has been updated regularly with searches of the MEDLINE database. The most recent literature update was performed through September 11, 2017. No new studies were identified in the latest literature review that would change our conclusions.

Assessment of diagnostic technology typically focuses on 3 categories of evidence: (1) technical reliability (test-retest reliability or intrarater reliability); (2) clinical validity (sensitivity, specificity, positive and negative predictive values) in relevant populations of patients; and (3) clinical utility (ie, demonstration that the diagnostic information can be used to improve patient outcomes).

Gastroesophageal reflux disease

Technical Reliability
We did not identify any literature assessing the technical performance of esophageal pH monitoring.

Clinical Validity

Catheter-Based pH Monitoring Systems
Esophageal pH monitoring for 24 hours with catheter-based systems is primarily used in patients who have gastroesophageal reflux disease (GERD) that have not responded symptomatically to a program of medical therapy (including proton pump inhibitors); monitoring is also conducted in patients with refractory extra-esophageal symptoms. Although established technology, aspects of these catheter-based systems’ use as a diagnostic test for GERD are problematic, and thus make it difficult to determine its utility or the utility of potential alternative tests.

There is no independent reference standard for GERD for specific populations. Traditional pH monitoring has been evaluated in patients with endoscopically diagnosed GERD, where it has been shown to be positive 77% to 100% of the time. However, in clinically defined but endoscopically negative patients, the test is positive from 0% to 71% of the time. In normal control populations, traditional pH monitoring is positive in 0% to 15% of subjects. Thus, the test is imperfectly sensitive and specific in patients with known presence or absence of disease. The current evidence for the diagnostic capability of catheter-based pH monitoring led Kahrilas and Quigley, authors of a technical review, “…to conclude that ambulatory pH studies quantify esophageal acid exposure but that this has an imperfect correlation with reflux-related symptoms, esophageal sensitivity, or response to acid suppressive therapy.”

Without a reference standard for GERD, it is difficult to compare the diagnostic test performance of different types of tests. While it is possible to determine the degree to which the 2 tests correlate, it is difficult to determine if one is better than the other.

Wireless pH Monitoring
Several observations of relevance to this review are based on a 2006 TEC Special Report on wireless esophageal pH monitoring. Six case series reviewed in the report demonstrated success rates of over 90% in completing a 48-hour pH study. Two studies that surveyed patients who received wireless pH monitoring and patients who received traditional catheter monitoring showed less discomfort, less disruption of daily activities, and higher overall satisfaction with the wireless test. Studies that evaluated test positivity in clinically diagnosed GERD cases and normal controls showed similar results (results were also similar in patients using traditional pH monitoring). Studies that directly compared the performance of traditional catheter and wireless pH monitoring in the same patients revealed a fairly close correlation between the 2 types of studies after correcting for calibration differences; however, the ideal cut-point for test positivity differed for the tests.

Some studies have attempted to support an argument that a longer monitoring time with a wireless monitor would result in a superior test performance; however, without a reference standard, or showing superior patient outcomes based on the longer test, such an argument cannot be made. The longer monitoring period usually results in a larger proportion of tests that
are classified as positive, depending on the method of determining a positive test. Prakash and Clouse (2005) compared the diagnostic yield for a single day of monitoring with the complete 2 days of monitoring. They reported that the second day of recording time increased the proportion of subjects with symptoms by 6.8%. However, this study had several methodologic flaws. Ideally, a study that compares the diagnostic performance of an additional day of monitoring would require an independent reference standard or demonstration of improved patient outcomes when managing patients with a 1-day vs a 2-day study. In this study, the 2-day study was essentially considered the “reference test,” and there was no discussion of how the second day of monitoring was used to improve patient management in this heterogeneous group of patients. In addition, in their statistical analysis, the authors eliminated patients who did not report any symptoms during the testing period, thus deflating the denominator and inflating the yield of the additional day of testing. Finally, the 1-day test was essentially a component of the 2-day test, and thus the 2 monitoring periods were not independent, further limiting any comparison between them. A greater number of positive tests produced by a longer duration of test is not evidence of a superior test.

Studies published since the 2006 TEC Special Report have shown similar findings on the correlation between wireless pH monitoring and standard catheter monitoring. Wenner et al (2007), in a study of 64 patients with GERD and 50 asymptomatic controls, showed a sensitivity of 59% to 65% when setting the specificity to 90% to 95%. The sensitivity of wireless monitoring was noted to be worse than other studies of traditional pH monitoring, but the patient population may have had less severe disease. A study by Schneider et al (2007) revealed a similar diagnostic performance of wireless and traditional pH monitoring. Hakanson et al (2009) evaluated simultaneous wireless and traditional pH testing in 92 patients. Wireless pH testing showed consistently lower estimates of acid exposure than traditional pH testing. The 2 techniques correlated \( r^2 = 0.66 \); however, the range between limits of agreement was wide. The techniques were concordant on the final diagnosis 82.1% of the time.

Additional studies have replicated findings that a longer period of monitoring increases the proportion of positive tests. Scarpulla et al (2007) attempted 96-hour monitoring in 83 patients. Monitoring for the full 96 hours was successful in 41% of patients. In them, the proportion showing some degree of pathologic acid exposure increased as monitoring time increased. Garrean et al (2008) studied the use of 96-hour pH testing where during the first 2 days of monitoring, patients were off therapy, and during the second 2 days, they were prescribed proton pump inhibitors. As expected, during the second and third days, fewer patients showed reflux symptoms. It is difficult to determine from data analysis how such a testing protocol improves the diagnosis of GERD. Grigolon et al (2011) showed that, in 51 patients receiving prolonged monitoring, the 96-hour test reduced the number of indeterminate tests from 11 to 5. In this particular study, comparison of outcomes for patients who received wireless monitoring, and a matched control group who received traditional catheter monitoring, showed similar outcomes and satisfaction.

**Impedance pH Testing**
Evidence on the use of impedance pH testing suffers from issues similar to the evaluation of wireless pH testing: lack of a reference standard and lack of evidence that shows improved patient outcomes. Many studies have argued that an increase in positive tests, or diagnostic yield, is by itself evidence that supports the validity of the test. However, the increase in positive tests, if it indicates increased sensitivity, may decrease specificity. The net effect on patient management and patient outcomes is uncertain.

Several studies have demonstrated a higher yield for positive tests when using impedance pH testing and identifying reflux events that are nonacidic or only weakly acidic (and thus would not be detected using pH testing alone). Bajbouj et al (2007) studied 41 patients with atypical GERD symptoms with numerous tests. The test that produced the highest number of positive findings was impedance pH testing. Bredenoord et al (2006) did a similar study in 48 patients. A higher proportion of subjects had positive tests when using impedance pH data (77%) than when using pH data alone (67%). A study by Mainie et al (2006) showed similar findings.

Studies have examined performing impedance pH testing while patients are on acid-suppression therapy. Vela et al (2001) demonstrated that, during acid-suppressive therapy, the total number of reflux episodes is similar, but fewer episodes of acidic reflux occur.

Although impedance-pH testing produces a higher number of positive tests, particularly compared with traditional or wired pH testing in the setting of concurrent acid-suppressive therapy, there is insufficient evidence that these test results are more accurate.

Section Summary: Clinical Validity
The tests under consideration all correlate with symptoms of GERD or endoscopically defined GERD; however, without a true reference standard for clinical GERD, the diagnostic characteristics of catheter-based pH monitoring, wireless pH monitoring, and impedance-pH testing are uncertain.

Clinical Utility
Clinical utility of pH testing can be determined by studies that directly compare strategies of diagnosis and treatment using pH testing with strategies of diagnosis and treatment not using pH testing, or by a compelling indirect chain of evidence that supports improved outcomes with the use of pH testing. We did not identify studies evaluating the clinical utility of any of the tests. A chain of evidence supporting pH testing has not been clearly formulated and argued.

Summary of evidence
For individuals who have GERD who receive catheter-based pH monitoring, the evidence includes various cross-sectional studies evaluating test performance in different populations. Relevant outcomes are test accuracy and validity, symptoms, and functional outcomes. Positive pH monitoring tests correlate with endoscopically defined GERD and with GERD symptoms, but because there is no reference standard for clinical GERD, diagnostic characteristics cannot be determined. There are no studies of clinical utility showing improved outcomes, and the chain of evidence supporting the utility of the test is weak. The evidence is insufficient to determine that the technology improves health outcomes.
For individuals who have GERD who receive wireless pH monitoring, the evidence includes various cross-sectional studies evaluating test performance and diagnostic yield in different populations. Relevant outcomes are test accuracy and validity, symptoms, and functional outcomes. Positive wireless pH monitoring tests correlate with endoscopically defined GERD and GERD symptoms, but because there is no reference standard for clinical GERD, diagnostic characteristics cannot be determined. Some studies have shown higher positive test rates with prolonged wireless monitoring compared with catheter-based pH monitoring, but the effect of this finding on patient outcomes is uncertain. There are no studies of clinical utility showing improved outcomes, and the chain of evidence supporting the utility of the test is weak. The evidence is insufficient to determine that the technology improves health outcomes.

For individuals who have GERD who receive impedance pH testing, the evidence includes various cross-sectional studies evaluating test performance and diagnostic yield in different populations. Relevant outcomes are test accuracy and validity, symptoms, and functional outcomes. Positive impedance pH tests correlate with endoscopically defined GERD and with GERD symptoms, but because there is no reference standard for clinical GERD, diagnostic characteristics cannot be determined. Some studies have shown higher positive test rates with impedance pH testing compared with pH testing alone, but the effect of this finding on patient outcomes is uncertain. There are no studies of clinical utility showing improved outcomes, and the chain of evidence supporting the utility of the test is weak. The evidence is insufficient to determine that the technology improves health outcomes.

**Supplemental information**

**Clinical input from physician specialty societies and academic medical centers**

While the various physician specialty societies and academic medical centers may collaborate with and make recommendations during this process, through the provision of appropriate reviewers, input received does not represent an endorsement or position statement by the physician specialty societies or academic medical centers, unless otherwise noted.

In response to requests, input was received from 1 physician specialty society (2 reviewers) and 3 academic medical centers while this policy was under review for 2010. Input was mixed. Most reviewers indicated that the wireless device was more comfortable and allowed patients to have more varied activities during the recording. One reviewer cited problems with availability of the catheter-based systems. Moreover, most reviewers agreed that a link between wireless monitoring and improved health outcome had not been demonstrated.

**Practice guidelines and position statements**

**American College of Gastroenterology**

The American College of Gastroenterology (ACG) released practice guidelines on esophageal reflux testing in 2007. The literature up to 2006 was reviewed. Although the literature on wireless pH testing was extensively reviewed, the recommendations for testing made no distinction between wireless and traditional pH monitoring. An indirect endorsement of wireless monitoring
might be inferred from a statement that a 48-hour study would produce a greater diagnostic yield from a symptom-association test. Symptom-association tests require statistical testing of the data, and a 48-hour test produces more data points. However, these statistical correlation tests are not perfect, because the guidelines state that such measures “do not ensure a response to either medical or surgical antireflux therapies.” No studies were cited that indicated superior outcomes for patients for treatment guided by wireless pH testing vs traditional pH testing. The major advantage for the wireless system cited was patient tolerability.

Impedance pH monitoring was cited as “may be useful” (a lower category of recommendation than for pH monitoring) for evaluation of patients with insufficient response to medical therapy in whom documentation of nonacid reflux would alter clinical management. Moreover, ACG suggested that impedance monitoring has a greater yield for findings than pH monitoring when performed on proton pump inhibitor (PPI) therapy. The last statement of the guidelines specified that the implications of an abnormal impedance test are unproven at this time.

In 2013, ACG published guidelines on the diagnosis and management of gastroesophageal reflux disease (GERD). The guidelines stated, “ambulatory esophageal reflux monitoring is indicated before consideration of endoscopic or surgical therapy in patients with nonerosive disease, as part of the evaluation of patients refractory to PPI therapy, and in situations when the diagnosis of GERD is in question.” This was a strong recommendation based on a low level of evidence. The ACG guidelines noted there is limited evidence and lack of clear consensus on how testing should be performed (eg, catheter-based pH, wireless pH, or impedance pH) for refractory GERD.

**American Gastroenterological Association**

The American Gastroenterological Association released a medical position statement and accompanying technical review on the management of GERD in 2008. Ambulatory impedance pH, catheter pH, and wireless pH monitoring were all supported as methods to evaluate patients with suspected GERD with otherwise normal endoscopy and no response to PPI therapy. The guidelines had a grade B recommendation, denoting fair evidence that the practice improves health outcomes. The guidelines additionally stated that the wireless pH monitor has superior sensitivity to catheter pH monitoring because of the extended period of recording.

However, as noted previously, an increase in positive tests has been documented in other reports as producing both increased sensitivity and decreased specificity relative to the reference standard used in the particular study. Thus, taking into account both characteristics of diagnostic performance, it is unclear whether patient outcomes are improved.

**National Institute for Health and Care Excellence**

The National Institute for Health and Care Excellence released a technology appraisal guidance on catheterless esophageal pH monitoring in 2006. This guidance indicated catheterless esophageal pH monitoring appears to be safe and effective and is commonly indicated for GERD symptoms refractory to PPIs and for GERD symptom recurrence after antireflux surgery.
In 2015, the Institute published recommendations on the diagnosis and management of GERD in children and young people. The recommendations specific to esophageal pH monitoring included: “Consider performing an esophageal pH study (or combined esophageal pH and impedance monitoring if available) in infants, children and young people with:

- Suspected recurrent aspiration pneumonia
- Unexplained apneas
- Unexplained non-epileptic seizure-like events
- Unexplained upper airway inflammation
- Dental erosion associated with a neurodisability
- Frequent otitis media
- A possible need for fundoplication
- A suspected diagnosis of Sandifer’s syndrome.

Consider performing an esophageal pH study without impedance monitoring in infants, children and young people if, using clinical judgement, it is thought necessary to ensure effective acid suppression.”

**Medicare national coverage**

There is no national coverage determination. In the absence of a national coverage determination, coverage decisions are left to the discretion of local Medicare carriers.

## V. Important Reminder

The purpose of this Medical Policy is to provide a guide to coverage. This Medical Policy is not intended to dictate to providers how to practice medicine. Nothing in this Medical Policy is intended to discourage or prohibit providing other medical advice or treatment deemed appropriate by the treating physician.

Benefit determinations are subject to applicable member contract language. To the extent there are any conflicts between these guidelines and the contract language, the contract language will control.

This Medical Policy has been developed through consideration of the medical necessity criteria under Hawaii’s Patients’ Bill of Rights and Responsibilities Act (Hawaii Revised Statutes §432E-1.4), generally accepted standards of medical practice and review of medical literature and government approval status. HMSA has determined that services not covered under this Medical Policy will not be medically necessary under Hawaii law in most cases. If a treating physician disagrees with HMSA’s determination as to medical necessity in a given case, the physician may request that
HMSA reconsider the application of the medical necessity criteria to the case at issue in light of any supporting documentation.
VI. References


