Comparative Effectiveness Review
Number 110

Effective Health Care Program

Wireless Motililty Capsule Versus Other Diagnostic Technologies for Evaluating Gastroparesis and Constipation: A Comparative Effectiveness Review

Executive Summary

Gastroparesis

Definition and Prevalence

Gastroparesis is a condition in which patients experience symptoms of delayed gastric emptying in the absence of an actual physical blockage.[1] The most common symptoms are nausea, vomiting, early satiety, bloating, abdominal pain, and postprandial fullness.[2] Assessing gastric emptying delay is essential to diagnosing gastroparesis. In clinical research, the definition of gastroparesis is delayed gastric emptying as detected by clinical testing and the presence of symptoms of nausea and/or vomiting, postprandial fullness, early satiety, bloating, or epigastric pain for more than 3 months. Using this definition, the cumulative incidence of gastroparesis is 4.8 percent in people with type 1 diabetes, 1.0 percent in people with type 2 diabetes, and 0.1 percent in people without diabetes, who may have idiopathic gastroparesis or other etiologies.[2] A 2007 community-based study estimated the prevalence of gastroparesis to be 9.6 per 100,000 for men and 37.8 per 100,000 for women.[2] Newer estimates of prevalence report a higher rate

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at **www.effectivehealthcare.ahrq.gov/reports/final.cfm**.

AHRQ
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective Health Care

of 24.2 per 100,000 inhabitants. Some experts estimate that more than 1.5 to 3 million Americans may have gastroparesis.[3, 4]

Etiology and Clinical Course

The etiologies of gastroparesis are most often idiopathic, diabetic, or postsurgical, but can also be autoimmune, paraneoplastic, or neurologic. The condition is generally assessed in the outpatient setting, but some patients become severely ill with intractable vomiting and dehydration and are hospitalized. Hospitalizations for gastroparesis increased by 158 percent between 1995 and 2004.[5] In individuals with diabetes and gastroparesis, digestion of food is unpredictable, and wild swings in blood glucose can increase morbidity and necessitate medical care.

Evaluation of Possible Gastroparesis

A standard assessment for patients with typical symptoms (e.g., nausea, vomiting, bloating, abdominal pain, early satiety) of gastroparesis starts in the office of a physician, who takes a careful medical history and performs a physical examination.[6] First, the physician must rule out mechanical or medication-related dysfunction. Medications that commonly cause gastric emptying delay are opiates or glucagon-like peptide agonists. Second, the physician needs to test for gastric emptying. Methods of testing include gastric emptying scintigraphy, antroduodenal manometry, and now wireless motility capsule (WMC) technology. Motility disorders are difficult to diagnose. Multiple contributing factors make pathophysiology more complex, and physicians can have difficulty gathering a unifying diagnosis from a single test. In addition, most of the available tests have some inconsistency in performance, which can make their interpretation difficult.

Gastric Scintigraphy

Gastric scintigraphy is the ingestion of a meal commonly standardized to toast, jam, water, and radiolabeled egg whites. The egg whites are visible as they pass through the gastrointestinal tract during subsequent timed imaging, ideally 4 hours.[7, 8] Clinicians withhold interfering medications, such as opiates, motility agents, and glucagon-like peptide agonists, for 5 to 7 days before scintigraphic testing. Full 4-hour testing is more commonly available at regional referral centers or tertiary care centers with established practices of motility specialists.[7] Generally, physicians diagnose delayed gastric emptying if less than 90 percent of the gastric content has emptied at 4 hours, meaning that the patient has retained more than 10 percent of the content.

Antroduodenal Manometry

Antroduodenal manometry can provide information about gastric physiology. A manometry catheter, inserted through the pyloric channel with endoscopic guidance and patient sedation, measures pressure. Antroduodenal manometry may help differentiate myopathic and neuropathic etiologies of symptoms. Myopathy is present if amplitude muscle pressure falls below 30 mmHg, and neuropathy is present if uncoordinated bursts of muscle activity occur.

WMC

The United States Food and Drug Administration (FDA) approved WMC for identifying motility disorders. This device is a portable, one-time use, ingestible capsule that, when swallowed, records and transmits data to a receiver as it travels through the gut. A single device can detect specific transit times in the stomach, small bowel, and colon in a single test. The capsule can measure pH, pressure, and temperature to track location, gastric contents, and expulsion time from different regions of the bowel. The American Neurogastroenterology and Motility Society (ANMS) recommends its use and the American College of Gastroenterology considers it a technology that has great promise and should be watched.[9]

The patient takes the pill after eating a standardized meal and wears a small monitor that makes telemetry recordings. The established cutoff point for gastric emptying time is 300 minutes.[10] Disadvantages of the capsule include failure to capture data (requiring repeat testing) and delay or total failure to pass (requiring serial x rays to document passage or endoscopic or surgical removal, respectively). Another disadvantage is that it should not be used in patients with a possible stricture, altered anatomy, or severe pyloric stenosis.[11] Patients ideally should be able to tolerate not using proton pump inhibitors and histamine 2 blockers before testing.[11] Advantages include that it is wireless and painless and contains no radiation.[12, 13]

Use of Gastric Emptying Testing To Guide Treatment

Effective gastric-emptying-delay testing guides physicians in their recommendations for nutrition, medication, and surgical therapies. Testing informs physicians about the length and severity of delay, and this information can guide changes in diet to accommodate better gastric emptying. Recommended changes in diet may include a lowfat diet,

a low-residue diet (i.e., low fiber, easy to empty from the stomach), a liquid diet, or changing one's consumption pattern to multiple small meals per day. Testing can also inform physicians about the use of prokinetic medicines like metoclopramide or erythromycin, which are often used to treat gastroparesis. This is important because of the FDA black box warning about the side effects of using metoclopramide for more than 3 months. Both metoclopramide and erythromycin can cause profound tachyphylaxis, limiting any intended benefit. Similarly, domperidone (Motilium®) is not FDA-approved but is available in many countries outside the United States and is used in clinical care and research in the United States through an Investigational New Drug Application. Therefore, clear documentation of gastroparesis is important to physicians who are considering using a prokinetic. Patients with severe symptoms and severe emptying delay despite dietary changes may need feeding tubes, such as jejunostomy or gastrojejunostomy tubes, that bypass the stomach entirely. As patients undergo consideration for compassionate use of gastric stimulation therapy, one of the eligibility criteria is the presence of gastric emptying delay on testing. Thus, accurate diagnosis of gastroparesis is integral to decisions about management.

Outcomes

Major outcomes of interest are assessment of motility and diagnosis of gastric emptying delay. Other outcomes include the ability of testing to influence treatment decisions (e.g., changes in medications, nutrition), or to affect patient-centered outcomes (e.g., symptom improvement, need for surgery, quality of life, patient satisfaction). It is important to consider potential harms of testing such as capsule retention, radiation exposure, and mortality. Clinicians and policymakers may also be interested in the effects on resource utilization, such as the need for additional tests, physician services, or hospitalizations.

Constipation

Definition and Prevalence

Constipation is common, occurring in 15 to 20 percent of the U.S. population.[11, 14, 15] Multiple professional societies define constipation (with slight variation) as fewer than two bowel movements per week or a decrease in a person's normal frequency of stools accompanied by straining, difficulty passing stool, or passage of hard solid stools.[11] Physicians must assess patients with symptoms of constipation via their medical history and a physical examination to exclude malignant or organic causes of constipation. Clinicians should ask about warning signs such as new onset of symptoms, obstructive symptoms, rectal bleeding, unintentional weight loss, or family history of early colon cancer. A rectal examination can help to delineate rectal function and tone and exclude a low rectal cancer. Clinicians should perform a colonoscopy on all patients over 50 who have never received a screening colonoscopy, and those who have fecal occult blood, iron deficiency anemia, or any other warning signs.[16] However, the yield of colonoscopy in patients with constipation with warning signs is low. Once a physician has eliminated all organic causes for constipation, a diagnosis of functional constipation is appropriate. Physicians do not need to test an individual less than 50 years old and without "red flag" symptoms in order to diagnose constipation if the patient meets the Rome III criteria.

The Rome III criteria define functional constipation as follows:[17]

1. Two or more of the following:
    a. Straining during at least 25 percent of defecations
    b. Lumpy or hard stools in at least 25 percent of defecations
    c. Sensation of incomplete evacuation for at least 25 percent of defecations
    d. Sensation of anorectal obstruction/blockage for at least 25 percent of defecations
    e. Manual maneuvers to facilitate at least 25 percent of defecations (e.g., digital evacuation, support of the pelvic floor)
    f. Fewer than three defecations per week
2. Loose stools rarely present without the use of laxatives
3. Insufficient criteria for irritable bowel syndrome

A patient must have two or more of the above criteria for the last 3 months, with symptom onset being at least 6 months prior to diagnosis.

Clinically, patients with slow-transit constipation, also known as colonic inertia, often have the most severe symptoms of those patients with constipation, with prolonged periods of time between bowel movements. Often, standard medical therapies have failed these patients. The definition of slow-transit constipation is retention of greater than six radiopaque markers after 5 days from ingestion.[11, 18] The reported incidence of slow-transit constipation is 1 in 3,000 or 0.033 percent. Other studies list a prevalence of 0.17 percent.[19] The true incidence is likely unknown.

Etiology and Clinical Course

There are several types of chronic constipation including slow-transit, normal-transit, and dyssynergic defecation. There is also constipation-predominant irritable bowel syndrome.[11] Physicians should recommend lifestyle changes and medical management for all patients with symptoms of constipation. Lifestyle changes include drinking appropriate quantities of liquid, removing all possible offending medications, and eating the U.S. Department of Agriculture's recommended amount of vegetables, fruit, and fiber. Medical management includes avoiding constipating medications and initiating bulking agents (e.g., fiber supplements), stool softeners (docusate, mineral oil), osmotic and stimulant laxatives (e.g., lactulose, milk of magnesia, magnesium citrate, polyethylene glycol [Miralax®], PEG-3350, senna), or prokinetics (e.g., bisacodyl), and secretagogues/prokinetics (e.g., lubiprostone, linaclotide), or in other countries prucalopride (not yet FDA-approved), as indicated. Thus, the initial evaluation of constipation symptoms does not often involve colonic transit testing.

Evaluation of Possible Slow-Transit Constipation

For certain individuals with suspected slow-transit constipation, colon transit testing can provide valuable insight into the etiology of the constipation. Testing can explain why a patient fails basic therapy and can help identify or exclude patients as surgical candidates.[11] However, a single test may not reflect the full complexity of a patient's motility disturbances. For example, anorectal dysfunction can impact colonic transit, but must be assessed by anorectal manometry separate from other transit testing. Furthermore, most of the available tests have some inconsistency in performance, which makes their interpretation difficult in some cases. Transit disorders include slow colonic transit or colonic inertia, a hypomotile disorder of the colon where transit in the proximal colon is slow without evidence of retropulsion of the markers from the left colon and without evidence of anorectal dysfunction. Defecatory dysfunction (or functional outlet dysfunction) is the presence of uncoordinated motion of the anorectum muscles causing ineffective or weak expulsion of stool. Idiopathic megacolon (primary or secondary), a pathological enlargement of the colon, can also be present and may occur in conjunction with longstanding neurological diseases or Hirschsprung's disease, a failure of the development of the nerve cells within the colon wall.[20] The main diagnostic methods used to test for colonic motility are radiopaque marker (ROM) examination, colonic scintigraphy, colonic and anorectal manometry, and WMC testing.[21, 22] The nonreference standard is ROM.

ROM

The nonreference standard of ROM testing (commonly known as Sitz Markers) defines slow-transit constipation.[21, 22] In its simplest form, a patient ingests the ROMs on day zero and then receives an x ray at day 5, using overpenetrated films (110 kiloelectron volts) in order to reduce x-ray exposure. Gastroenterologists no longer focus on the areas of colon that have the greatest delays, since studies have shown that this does not predict pathophysiology or treatment. The only exception to this statement is the patient who accumulates markers in the rectum and does not pass them; this would strongly suggest a defecation disorder. Marker retention identifies patients with slow transit.[11, 18] One disadvantage to ROM testing is x-ray exposure. However, the test is valid and in practice since the late 1960s.[18]

Colonic Scintigraphy

Colon scintigraphy is rarely available outside of highly-specialized motility research centers. It follows an ingested radiolabeled meal or radiolabeled tracer from the upper to lower gastrointestinal tract. A disadvantage is that testing requires several days and entails radiation exposure. Studies have assessed the validity of colon scintigraphy relative to ROM.[23, 24] The ANMS guidelines endorse colon scintigraphy as a potential test for evaluating colon transit.

WMC

WMC testing assesses colonic transit time by measuring the time between cecal entry and rectal exit. Cecal entry produces a sustained drop in pH of greater than 1 unit that occurs more than 30 minutes after gastric emptying. Rectal exit produces a large temperature reduction.[11] One disadvantage is that 5 percent of tests do not record cecal entry time data, thus limiting the diagnostic potential of the study.[18] Camilleri has reported the use of the combined small bowel and colon transit time to allow for interpretation of the tests that do not report cecal entry.[25] Other disadvantages are that clinicians must use radiographic imaging to identify capsule retention when it fails to pass spontaneously, and that the device can fail at a rate up to 3 percent according to some studies. In addition, prolonged colon transit time with this technology does not necessarily distinguish slow transit from defecatory dysfunction.

Use of Colon Transit Testing To Guide Treatment

Most patients with chronic constipation see symptom improvement with medical therapy and/or lifestyle changes. For some patients, all measures fail and physicians must use colon transit testing to better understand the motility disorders. Physicians use anorectal

manometry to identify anorectal or outlet dysfunction, and treat with biofeedback therapy. Evidence of Hirschsprung's disease is an indication for surgical segmental resection. Megacolon requires medical therapy tailored to reducing gas formation, and reduction of fiber intake may paradoxically relieve symptoms. If these conservative measures fail, megacolon may require segmental or total colectomy. If testing confirms the presence of slow-transit constipation (colonic inertia) without the use of laxatives, then the next step in evaluation in some centers is transit testing with use of laxatives. Physicians should only consider surgery as a potential therapy after they have demonstrated colonic inertia.[26] Clear demonstration of severe total or segmental slow-transit constipation is an indication for colectomy; however, most clinicians reserve colectomy for patients with the most terminal or untreatable conditions.

Outcomes

A major outcome of interest to clinicians is the ability to characterize transit time and to diagnose slow-transit constipation. Other outcomes include the ability of testing to influence treatment decisions (e.g., change in medications, change in nutrition) or to affect patient-centered outcomes (e.g., symptom improvement, need for surgery, quality of life, patient satisfaction). It is important to consider potential harms such as capsule retention, radiation exposure, and mortality. Clinicians and policymakers may also be interested in the effects on resource utilization such as the need for additional tests, physician services, and hospitalizations.

Scope of Review and Key Questions

Our objective was to summarize the evidence on how useful current testing modalities for gastric and colonic motility are for diagnosing disease. We sought to determine whether WMC testing is useful in conjunction with or instead of other testing modalities for diagnosing and managing motility disorders. We also sought to define the populations that would benefit most from motility testing, including WMC testing. We listed our Key Questions (KQs) below and displayed them in Figure A.

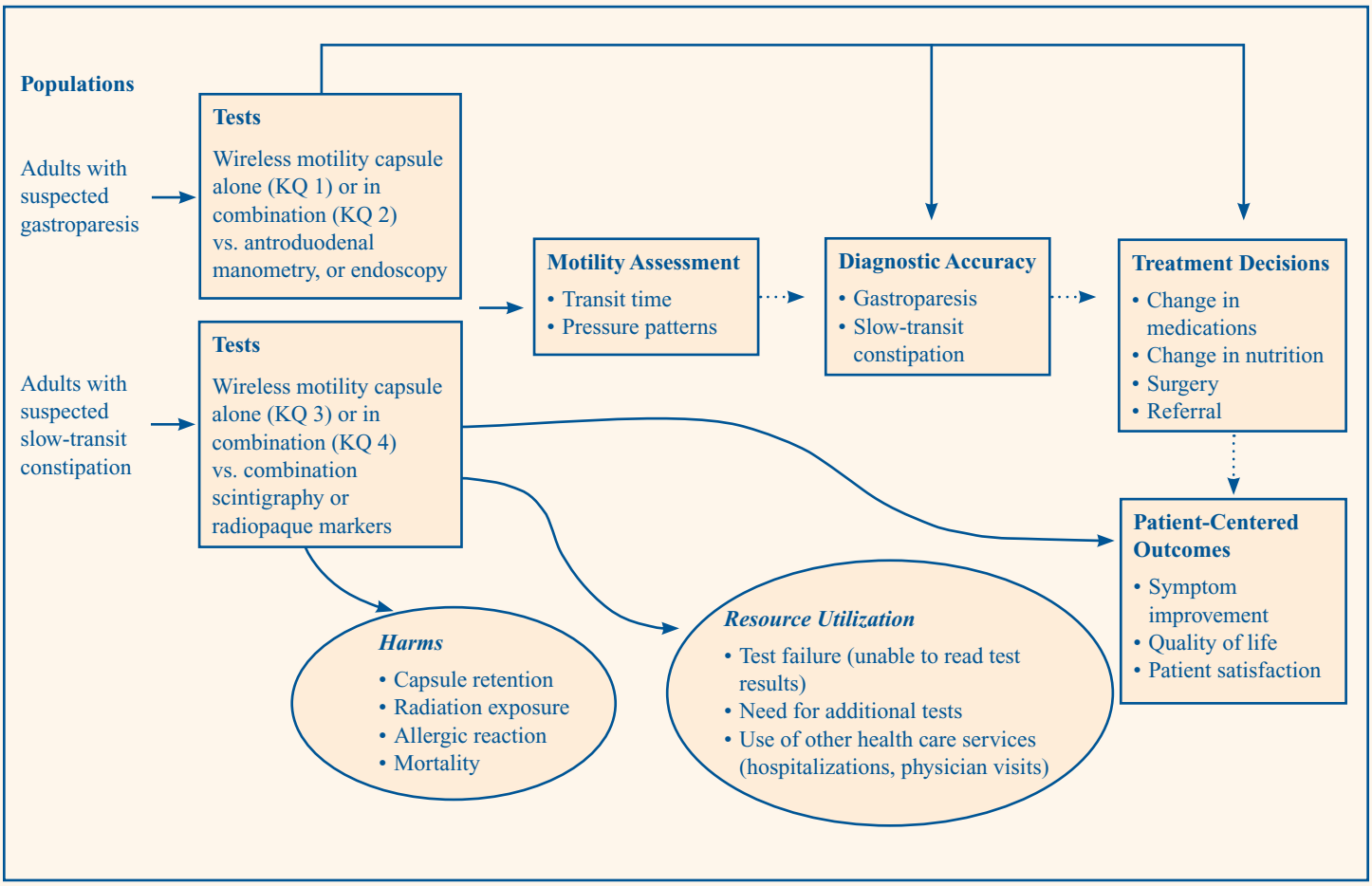
KQ 1. In the evaluation of gastric dysmotility, how does the WMC alone compare with gastric scintigraphy, antroduodenal manometry, and endoscopy, in terms of diagnostic accuracy of gastric emptying delay, accuracy of motility assessment, effect on treatment decisions,effect on patient-centered outcomes, harms, and effect on resource utilization?

KQ 2. When gastric scintigraphy, antroduodenal manometry, or endoscopy is used in the evaluation of gastric dysmotility, what is the incremental value of also using WMC, in terms of diagnostic accuracy of gastric emptying delay, accuracy of motility assessment, effect on treatment decisions, effect on patient-centered outcomes, harms, and effect on resource utilization?

KQ 3. In the evaluation of colonic dysmotility, how does WMC alone compare with ROM and scintigraphy in terms of diagnostic accuracy of slow-transit constipation, accuracy of motility assessment, effect on treatment decisions, effect on patient-centered outcomes, harms, and effect on resource utilization?

KQ 4. When an ROM or scintigraphy is used in the evaluation of colonic dysmotility, what is the incremental value of also using WMC, in terms of diagnostic accuracy of slow-transit constipation, accuracy of motility assessment, effect pm treatment decisions, effect on patient-centered outcomes, harms, and effect on resource utilization?

Figure A. Analytic framework for research on the comparative effectiveness of diagnostic technologies for evaluating gastroparesis and constipation



KQ = Key Question

Methods

Literature Search Strategy

We searched the following databases for primary studies for the periods in parentheses: MEDLINE® (1966 to July 1, 2012) and Embase® (1974 to July 1, 2012). We developed a search strategy for MEDLINE, accessed via PubMed, based on an analysis of the medical subject headings (MeSH) terms and text words of key articles identified a priori. Additionally, we reviewed the reference lists of included articles and any relevant review articles. We asked the manufacturer of WMC about any published or unpublished randomized controlled trials or observational studies that evaluated WMC. The manufacturer submitted comments on the draft report but did not submit any new materials. We searched ClinicalTrials.gov to identify any relevant trials.

Study Selection

Two independent reviewers evaluated each title, abstract, and full article. We included studies that compared WMC with other diagnostic tests among patients with suspected gastroparesis or slow-transit constipation, in terms of diagnostic accuracy, accuracy of motility transit time assessment, effect on treatment decisions, effect on patient-centered outcomes, effect on resource utilization, or harms. Other diagnostic tests were gastric scintigraphy, antroduodenal manometry, and endoscopy for the evaluation of gastroparesis, and scintigraphy and ROM for slow-transit constipation. There were no language

restrictions. We resolved differences between investigators regarding eligibility through consensus adjudication.

Data Abstraction

We created and pilot tested standardized spreadsheets for data extraction. The study investigators performed double data abstraction on each article. The second reviewer confirmed the first reviewer's abstracted data for completeness and accuracy. We formed reviewer pairs that included personnel with both clinical and methodological expertise.

For all articles, the reviewers extracted information on study characteristics (e.g., study design, country, location of recruitment, start year of recruitment, multicenter vs. single center, length of followup, length of time in between diagnostic tests), characteristics of study participants (e.g., condition; age; gender; race; weight; prior diagnostic tests; blood sugar; smoking status; diabetes status; defecatory dysfunction status; and the use of prokinetics, opiates, antidepressants, proton pump inhibitors, and laxatives), eligibility criteria, characteristics of WMC testing (e.g., was the pill swallowed or placed; did the study provide a standardized meal; did the study provide Ensure® shakes, and if so, when?[a]), characteristics of the other diagnostic tests, outcome measures, definitions, and the results of each outcome, including measures of variability. For each of the diagnostic tests, we collected information on the criteria used to make a diagnosis of gastroparesis or slow-transit constipation, and on whether the study instructed patients to abstain from tobacco, prokinetics, opiates, antidepressants, proton pump inhibitors, or laxatives at the time of the test.

Quality Assessment

Two reviewers independently assessed article quality. We selected and modified the questions from the QUADAS-2 quality assessment tool.[27] We supplemented this tool with quality-assessment questions (i.e., to assess spectrum bias) based on recommendations in the Methods Guide for Medical Test Reviews.[28] Our quality assessment included items on: (1) whether the study excluded healthy subjects from the diagnostic accuracy comparison, (2) whether the study excluded severely affected patients, (3) whether the study enrolled a random sample of patients, (4) whether all patients received the same reference standard, (5) whether the study included all patients in the analysis, (6) whether the study interpreted results of the test independently, (7) whether the time period between tests was reasonably short (within 3 months) to ensure that the condition did not change, (8) whether the study established cut-off values for test positivity before the study started, (9) whether a stated aim of the study was to compare diagnostic accuracy between WMC testing and other diagnostic tests, (10) whether the study reported on conflicts of interest, (11) whether a commercial source related to motility testing funded the study, and (12) whether a commercial source related to motility testing employed or gave funding or fees to any of the authors. The two reviewers resolved differences in quality assessment.

Applicability

We assessed the applicability of studies in terms of the degree to which the characteristics of the study population (e.g., age, etiology, comorbidities, prior surgery or gastric pacer), diagnostic test procedures (e.g., use of opiates during testing, use of bowel motility-altering agents such as laxatives or prokinetic agents), outcomes, and settings (e.g., referral center) were typical for the treatment of individuals with suspected gastroparesis or slow-transit constipation.

Data Analysis and Synthesis

We had planned to conduct meta-analyses if sufficient data were available (at least five studies for hierarchical summary receiver operator characteristic curves for diagnostic accuracy and at least three studies for other outcomes) and if studies were sufficiently homogenous with respect to key variables (e.g., population characteristics, study duration, diagnostic test procedures). We qualitatively summarized studies not amenable to pooling.

We considered gastric scintigraphy and clinical symptoms to be reference standards and ROM to be a nonreference standard. For measures of diagnostic accuracy when there was a reference standard, we summarized the results in terms of sensitivity, specificity, and test concordance. For measures of diagnostic accuracy when there was a nonreference standard, we summarized the results in terms of positive percent agreement, negative test agreement, and test concordance.[29] When the reference standard was a clinical diagnosis, we chose a 10 percent difference between tests in sensitivity or specificity as a potentially important difference because key studies were powered to detect a 10 percent difference.[25] When the reference/

[a]Ensure® is a commercial nutritional drink that is given to subjects in some centers as part of the WMC protocol.

nonreference standard was another diagnostic test, we considered it similar if WMC had a test concordance of at least 80 percent.

We conducted a sensitivity analysis where we included data that was reported only in a conference abstract.

Rating the Body of Evidence

At the completion of our review, we graded the strength of the available evidence addressing the KQs by adapting an evidence grading scheme recommended in the "Methods Guide for Medical Test Reviews"[28] and in the "Methods Guide for Effectiveness and Comparative Effectiveness Reviews."[30, 31] Both of these evidence grading schemes are based on recommendations of the GRADE Working Group.[32] We applied evidence grades to the bodies of evidence about each diagnostic test comparison for each outcome. We assessed the strength of the available evidence by assessing the risk of bias, consistency, directness, and precision.

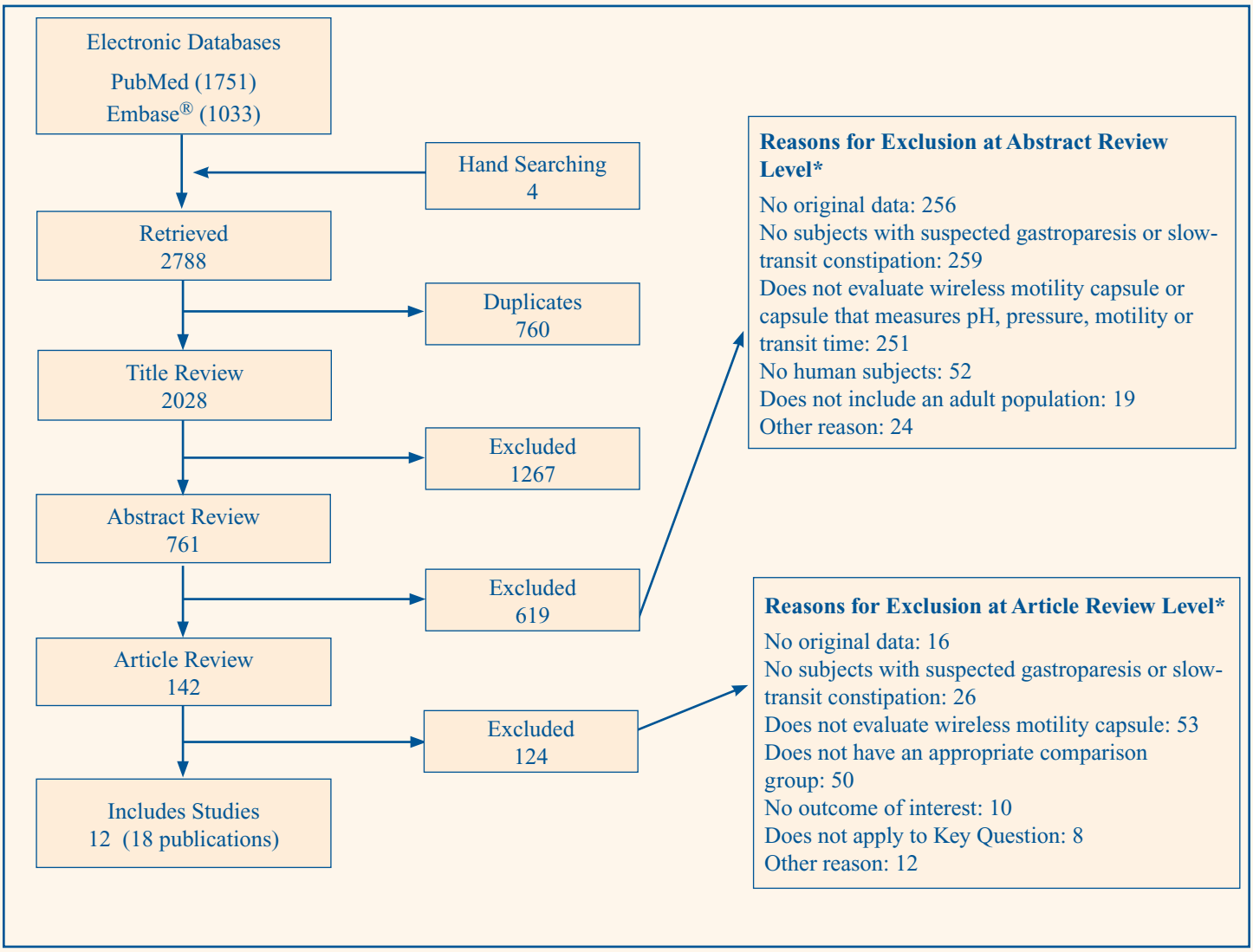
We classified evidence pertaining to the KQs into four basic categories: (1) "high" strength of evidence or SOE (indicating high confidence that the evidence reflects the true effect and that further research is very unlikely to change our confidence in the estimate of the effect); (2) "moderate" SOE (indicating moderate confidence that the evidence reflects the true effect and that further research may change our confidence in the estimate of the effect and may change the estimate); (3) "low" SOE (indicating low confidence that the evidence reflects the true effect and that further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate); and (4) "insufficient" SOE (indicating that evidence is unavailable or does not permit a conclusion).[32]

Results

Search Results

Figure B summarizes the results of our literature search. Our search retrieved 2,028 unique records. After reviewing the titles and abstracts, we considered 142 articles as potentially relevant and we reviewed the full text of the article for eligibility. We included a total of 12 studies (in 18 publications) in this review.[11, 25, 33-42] Seven studies (10 publications) evaluated WMC among patients with gastroparesis[33-39] and nine studies (14 publications) evaluated WMC among patients with slow-transit constipation.[11, 25, 33, 34, 36, 38, 40-42]

Figure B. Summary of literature search, with numbers of articles involved in each search step



* Total may exceed number in corresponding box, as articles could be excluded or more than one reason at this level.

Study Design Characteristics

Seven of the 12 studies were prospective,[10, 11, 25, 35, 37, 41, 42] 4 studies were retrospective,[33, 34, 36, 38] and 1 did not specify a study design.[40] All prospective studies applied the tests concurrently. Six studies appeared in meeting abstracts,[35-38, 40, 41] the remainder were in peer-reviewed publications.

All studies that reported the study location occurred in the United States.[10, 11, 25, 33-35, 37, 38] One study took place in multiple countries including the United States.[25] All studies that reported the location of recruitment occurred in tertiary centers.[11, 33-38]

Length of followup for the prospective studies and those with unspecified designs included the day of the testing only,[35, 37, 38, 40, 43] 3 days,[10] 5 days,[41, 42] 14 days,[25] and 21 days.[11]

Prospective studies included patients with known gastroparesis[10, 35, 37] or constipation.[11, 25, 42] Four retrospective studies included patients with suspected gastroparesis or constipation[33, 34, 36, 38] and one included patients with known constipation exclusively.[40] Six of the prospective studies also included patients without gastroparesis or constipation,[10, 11, 35, 37, 41, 42] whereas one study included only patients with known constipation.[25] Three studies that included patients with constipation used

the Rome III criteria as inclusion criteria.[11, 25, 42] Three studies reported age restrictions. One allowed patients 18 to 80 years of age[25] and two others included patients older than 65 years of age.[41, 42]

Study Population Characteristics

No gender restrictions were made in the inclusion criteria, although most of participants with gastroparesis or constipation were female. The mean age was 40 or greater in all studies that reported an average.[11, 25, 33, 34, 40, 41] Three studies reported on race or ethnicity.[10, 25, 34] More than 80 percent of the participants were white in these studies. No study reported a measure of weight, blood sugar, or smoking status at baseline. Two studies reported on the percent of patients with diabetes,[33, 39] reporting 15 and 37 percent with the disease, respectively. Two studies reported on defecatory dysfunction.[33, 40] In one study, 20 of 32 subjects had defecatory dysfunction,[40] and in another study 64 percent of patients had this dysfunction.33 Studies rarely reported on prior or concurrent use of medications, including prokinetics, opiates, antidepressants, proton pump inhibitors, and laxatives. Diagnostic testing prior to the study included scintigraphy[10, 33, 34, 37] and ROM.[33, 34]

Characteristics of Diagnostic Tests

We summarized the characteristics of the tests used in the studies, taking into consideration how the evaluation of gastrointestinal motility is dependent on multiple factors, including not only the types of test but also the specific protocols the studies employed, which were often not standardized. Our criteria for study assessment suggested that "best practice" studies would report on smoking, use of prokinetics, use of selective serotonin reuptake inhibitors, use of antacids, and the specific timing of ingestion of test meals. However, only a few of the studies with larger populations specified a predetermined meal and meal schedule for patients undergoing WMC testing. Several of the studies also specified that participants did not use prokinetics within the immediate timeframe of WMC testing. Clinicians most frequently performed gastric scintigraphy using the consensus protocol.[8] The community referral practice coordinated the ROM studies as per their local standards or the study made reference to a variation of the Metcalf protocol, wherein patients ingest ROMs and then receive an interval x ray and assessment of the marker location and number.[11, 44-46] Few articles gave more specific test characteristics for ROM testing. Most abstracts did not report on any of these characteristics.

Study Quality

We reported study quality separately for the full-length publications and the abstracts, because the abstracts had limited information about study methods. Overall, study quality was fair among the 11 full-length publications we assessed.[10, 11, 25, 33, 34, 39, 42, 47-50] Half of them used a uniform reference standard.[10, 11, 25, 47, 48] Only three studies interpreted the WMC results independently from the reference standard.[11, 25, 34] In another three studies that did not report blinding, we were able to confirm, after contacting the authors, that the studies interpreted results independently.[10, 39, 47]

KQ 1. Evaluation of Gastric Dysmotility: WMC Alone Versus Other Diagnostic Tests; and KQ 2. Evaluation of Gastric Dysmotility: WMC in Combination With Other Diagnostic Tests Versus Other Diagnostic Tests Alone

We summarized the results for KQ 1 and KQ 2 in Table A.

Table A. Summary of the strength of evidence (SOE) and main findings of studies comparing WMC alone (KQ 1) or in combination (KQ 2) with other diagnostic tests for the evaluation of gastroparesis

KQ	Comparison	Outcome(s)	SOE*	# of Studies	Main Findings
KQ 1	WMC vs. scintigraphy	Diagnostic accuracy	Low	7	Diagnostic accuracy of WMC is similar to scintigraphy. The sensitivity of WMC compared with clinical gastroparesis ranged from 65 to 68% and the specificity ranged from 82 to 87%. Sensitivity of WMC compared with gastric scintigraphy ranged from 59 to 86 percent and specificity ranged from 64 to 81 percent.
KQ 1	WMC vs. other modalities (antroduodenal manometry, endoscopy)	All outcomes	Insufficient	0	No studies addressed these comparisons.
KQ 1	WMC vs. scintigraphy	Motility assessment: Transit	Low	2	Transit data obtained via WMC are similar to scintigraphy.
KQ 1	WMC vs. scintigraphy	Motility assessment: pressure patterns	Low	3	WMC can measure pressure patterns and measurement of pressure patterns adds to diagnostic accuracy.
KQ 1	WMC vs. scintigraphy	Treatment decisions	Low	3	WMC testing alters management in patients with suspected gastroparesis (50-69% change in management for medicine, diet, or surgery).
KQ 1	WMC vs. scintigraphy	Resource utilization	Low	1	WMC testing may reduce the need for other studies, but this conclusion is based on one study with a high risk of bias. Need for anorectal manometry may not be reduced by WMC.
KQ 1	WMC vs. scintigraphy†	Harms	Low	2	Harms associated with WMC are minimal and no major safety issues were reported.
KQ 1	WMC vs. scintigraphy	Patient-centered outcomes	Insufficient	0	No studies reported on patient-centered outcomes for this comparison.
KQ 2	WMC in combination with other tests vs. scintigraphy	Diagnostic accuracy	Low	2	Adding WMC to conventional motility testing improves diagnostic accuracy in patients with suspected gastroparesis (sensitivity scintigraphy 42-51%; WMC 60-66%).

Table A. Summary of the strength of evidence (SOE) and main findings of studies comparing WMC alone (KQ 1) or in combination (KQ 2) with other diagnostic tests for the evaluation of gastroparesis (continued)

KQ	Comparison	Outcome(s)	SOE*	# of Studies	Main Findings
KQ 2	WMC in combination with other tests vs. scintigraphy	Motility assessment	Low	5	Adding WMC to conventional motility testing improves assessment of motility parameters in patient with suspected gastroparesis. (Scintigraphy does not measure pressure patterns.)
KQ 2	WMC in combination with other tests vs. scintigraphy	Treatment decisions, utilization, patient-centered outcomes, harms	Insufficient	0	No studies addressed these outcomes for these comparisons.

*The SOE was defined as follows: High = High confidence that the evidence reflects the true effect. Further research is unlikely to change our confidence in the estimate of the effect. Moderate = Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of the effect and may change the estimate. Low = Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate. Insufficient = Evidence is unavailable or does not permit a conclusion.

†Findings were based on observational studies that did not include a direct comparison of WMC with gastric scintigraphy.

KQ 3. Evaluation of Colonic Dysmotility: WMC Alone Versus Other Diagnostic Tests; and KQ 4. Evaluation of Colonic Dysmotility: WMC in Combination With Other Diagnostic Tests Versus Other Diagnostic Tests Alone

We summarized the results from KQ 3 and KQ 4 in Table B.

Table B. Summary of the SOE and main findings of studies comparing WMC alone (KQ 3) or in combination (KQ 4) with other diagnostic tests for the evaluation of slow-transit constipation

KQ	Comparison	Outcome	SOE*	# of Studies	Main Findings
KQ 3	WMC vs. ROM	Diagnostic accuracy	Low	5	Diagnostic accuracy of WMC is similar to ROM. Concordance between ROM and WMC was approximately 80% in 3 larger studies. The sensitivity for WMC compared with clinical suspicion ranged from 32 to 46% and specificity ranged from 95 to 100%. The sensitivity of day-5 ROM ranged from 28 to 37% and specificity ranged from 95 to 100%.
KQ 3	WMC vs. ROM	Motility assessment: Transit	Low	3	WMC was comparable with ROM in judgment of colonic transit time and identification of slow-transit constipation.
KQ 3	WMC vs. ROM†	Treatment decisions	Low	2	Very small numbers made comparison difficult for treatment decisions. Studies reported 7.1% change in nutrition, 21% referral to surgery, and 4% change in nutritional and behavioral therapies with WMC.
KQ 3	WMC vs. ROM	Resource utilization	Low	4	WMC testing may reduce the need for other tests, but this conclusion is based on one study with a high risk of bias. WMC does not replace anorectal manometry.
KQ 3	WMC vs. ROM†	Harms	Low	5	Harms and adverse events were infrequently reported for WMC or ROM. WMC is comparable to ROM with regard to harms. ROM involves exposure to at least one x ray. Day 21 x ray was required in a small proportion of patients who received WMC by protocol if the capsule had not spontaneously passed. Technical failures were reported in prototype devices the range of 3 to 10% in some series.[11]
KQ 3	WMC vs. ROM	Patient-centered outcomes	Insufficient	0	No studies addressed this outcome.
KQ 3	WMC vs. colonic scintigraphy	Diagnostic accuracy	Insufficient	0	No studies assessed the role of WMC versus these other modalities in the population of interest for this outcome.
KQ 4	WMC in combination with other diagnostic tests vs. other tests alone	Diagnostic accuracy	Insufficient	0	No studies addressed this question.

KQ = Key Question; ROM = radiopaque markers; WMC = wireless motility capsule.

*The SOE was defined as follows: High = High confidence that the evidence reflects the true effect. Further research is unlikely to change our confidence in the estimate of the effect. Moderate = Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of the effect and may change the estimate. Low = Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate. Insufficient = Evidence is unavailable or does not permit a conclusion.

†Findings were based on observational studies that did not include a direct comparison of WMC with ROM.

Discussion

Potential Niche for WMC

WMC is a potential improvement over previous testing modalities for patients with possible gastroparesis or slow-transit constipation because it is small and can be transported to patients wherever they live. Also, the capsule does not contain any radioactive material or entail x-ray exposure, and can record information about pressure, transit, and location simultaneously. Other testing modalities for gastric emptying and colonic motility assessment do not share these characteristics. Certain academic centers use scintigraphy to assess gastric transit abnormalities and evaluate whole gut motility; however, this procedure involves radiation exposure, significant patient time, and significant cost. Antroduodenal manometry assesses gastric pressure parameters but has limited availability and is more invasive than other testing modalities; thus, physicians commonly use it as an investigative tool rather than as a clinical test. ROMs are portable and small, but require radiation exposure, access to fluoroscopy, and radiology interpretation. In addition, all other methods for evaluating either gastric or colonic motility evaluate either transit or pressure, but not both; yet both are involved in disease pathogenesis. Since WMC can evaluate both transit and pressure simultaneously, it could allow more optimal assessment of motility than evaluation of either parameter independently. Likewise, by recording both parameters, WMC has the potential to replace a combination of modalities and provide more accurate diagnosis with less resource utilization and enhanced patient convenience.

In light of this potential niche, WMC is becoming much more readily available in both academic and community centers. However, questions remain about the position of WMC in the diagnostic algorithm for suspected motility disorders such as gastroparesis and slow-transit constipation. Is WMC equivalent to conventional testing? Is it superior? Is it more likely to establish a concrete diagnosis or guide medical therapy than conventional motility testing? Should it be used as a stand-alone test? What should be done when WMC results are normal but clinical suspicion remains? Recommendations from the ANMS practice guidelines suggest that WMC can be useful in the diagnostic work up of patients with suspected gastroparesis and slow-transit constipation as well as those with more generalized motility disorders, but these are consensus guidelines. There is no specific or clear information about when or how physicians should utilize a WMC.

We must also consider potential limitations of WMC. The manufacturer lists severe gastroparesis as a contraindication to capsule placement due to fear of capsule retention. In addition, by definition, WMC evaluates motility at only a single point, as opposed to antroduodenal manometry, which has multiple recording points, or scintigraphy, which looks at transit of an entire meal. One assumes that the single point of measurement is representative of motility parameters as a whole; however, this is an assumption only and is not clearly established in the literature. When assessing constipation, one cannot distinguish patients with slow-transit constipation from those with defecatory dysfunction based on only colonic transit time, so we need further motility testing with anorectal manometry and clinical judgment to evaluate defecation. Finally, parameters of motility for a nondigestible solid are different from those for either liquids or a meal—so that patients can have abnormalities that would be detected with one modality but that would not be seen with another. In short, while the potential of WMC testing is exciting, many questions remain as to its appropriate place in the diagnostic algorithm.

Key Findings and Implications

Few studies met our criteria for evaluation. The paucity of full-length articles with independent data limited our ability to answer the KQs definitively.

Key Question 1. Evaluation of Gastric Dysmotility: WMC Alone Versus Other Diagnostic Tests

WMC Versus Scintigraphy

We found low SOE from seven studies[10, 33-35, 37-39] that WMC has comparable diagnostic accuracy with gastric scintigraphy. The sensitivity was moderately greater in some studies, but some studies reported slightly lower specificity. The test agreement and diagnostic gain were moderate. Diagnostic agreement between WMC and gastric scintigraphy ranged from 58 to 86 percent for positive test agreement and from 64 to 81 percent for negative test agreement.

We found low SOE from five studies[10, 34, 35, 37, 39] that transit data obtained via WMC testing correlates well with scintigraphic gastric emptying. The reporting of the results in these studies was heterogeneous. One study reported a correlation coefficient of 0.73 between gastric emptying time measured by the WMC and 4-hour gastric emptying measured by gastric scintigraphy.10 When comparing WMC with gastric scintigraphy, one should keep in mind that WMC measures emptying of an indigestible object after the emptying of a meal, while gastric scintigraphy

measures emptying of a meal. In a sense, then, WMC indirectly measures what gastric scintigraphy measures. Good correlation between the two tests indicates that delayed meal emptying generally translates into delayed indigestible object emptying. Other studies reported sensitivity, specificity, and device agreement between WMC transit data and gastric scintigraphy.[34, 37, 39] All three studies examining transit time showed similar sensitivity and specificity for WMC and scintigraphy, and some studies reported increased diagnostic gain of sensitivity with WMC.

Low SOE from two studies supports the utility of WMC versus scintigraphy in measuring pressure profiles.[37, 39] A WMC detects pressure patterns, whereas scintigraphy cannot. It does appear, however, that abnormalities are more likely with WMC than scintigraphy--especially if one adds assessment of pressure patterns to the equation. However, based on the literature there remain questions as to whether increased diagnostic detection has clinical implications.

Overall, we had graded the SOE for many outcomes addressing KQ 1 to be low because we considered the evidence to have medium risk of bias, consistent reporting, direct nature of the data, and imprecise findings. The main limitation weighting the risk of bias was that studies did not prespecify patient enrollment or perform it in a random fashion; in fact many studies did not report how they selected patients for testing and study. Another limitation was the lack of advance prespecification of criteria and values of positivity of the tests the studies used. The final major limitation was that few studies mentioned whether they had selected a person without conflict of interest to manage data collection. Most studies had limited followup duration, which hampers our ability to draw conclusions about some of the outcomes that are really important to patients. A major strength of the full-length articles was that analysis involved an independent review of the results.

We could not conduct a meta-analysis because of the heterogeneity of the data and patient populations in the studies. Our ability to compare studies was limited by lack of consistency in the definition of reference standards. Studies often reported the reference standard as community-based gastric scintigraphy testing performed within 2 years of enrollment into a study. Local standards for scintigraphy vary greatly, and this introduced heterogeneity into the patient populations under investigation. Many studies had different definitions for key outcomes such as diagnostic agreement, sensitivity, and specificity, as well as different diagnoses based on similar test results. This latter discrepancy is likely due to changes over time in cut-off values for detecting gastroparesis using a WMC. It is uncertain if the available examinations of motility testing captured the full spectrum of patients, as academic referral centers were the primary recruitment site for studies. Overall, seven studies with 560 patients addressed the question of diagnostic accuracy.[33-39] For a rare illness, the large number of patients that researchers have included for evaluation reflects the great lengths that they have gone to in order to assess the quality of this modality.

Several studies suggested that there was some diagnostic gain with WMC as compared with scintigraphy, assuming that all the additional cases they identified were correct and not false positives.[10, 33, 34, 37, 39] The investigators attempted to minimize the impact of having a heterogeneous population by employing simultaneous scintigraphy and WMC at the time of assessment; sensitivity and specificity for both scintigraphy and WMC compared with symptoms in these studies is expectedly low given the issues above and the fact that the denominator may not have truly represented only gastroparetic patients. Device agreement is a more useful parameter to measure in these papers than sensitivity and specificity.[28] However, agreement is likely to be imperfect because these two modalities look at different mechanisms of transit.

Regarding treatment decisions, we did find that, in three studies, WMC testing altered management in patients with suspected gastroparesis (50 to 69 percent change in management for medicine, diet, or surgery). However, the SOE was low (i.e., likely to be changed by future evidence).

The evidence was insufficient to permit conclusions regarding the differences or similarities between gastric scintigraphy and WMC with regard to patient-centered outcomes or resource utilization. Very little research examined resource utilization, and no studies specifically examined this outcome with any rigor.

The findings contained in the literature are consistent with what would be expected based on the pathophysiology of gastroparesis and the comparative methods of WMC and gastric scintigraphy. Comparing scintigraphy with WMC is fundamentally a challenging endeavor. Both modalities evaluate different parameters. Scintigraphy looks at transit of a test meal and does not assess pressure. When the stomach processes a meal, fundic accommodation is followed by antral contractions that break up the food into small particles that are then propelled from the antrum to

the duodenum. In comparison, the WMC is not digested and is believed to exit the stomach when the gastric motility patterns change from a fed to fasting state and migratory motility complexes resume. As such, these two technologies are evaluating different parameters and a direct comparison may be challenging if one looks at transit alone.

WMC Anteroduodenal Manometry or Endoscopy

We did not find any head-to-head comparisons of antroduodenal manometry (which can record pressure patterns) and WMC in patients with suspected gastroparesis in our review. This makes it difficult to make a more definitive assessment of the ability of WMC to detect abnormalities in pressure patterns in our defined populations. Similarly, we did not find any studies that compared WMC with endoscopy among patients with suspect gastroparesis.

Key Question 2. Evaluation of Gastric Dysmotility: WMC in Combination With Other Diagnostic Tests Versus Other Diagnostic Tests Alone

WMC Plus Gastric Scintigraphy Versus Gastric Scintigraphy Alone

Two studies[34, 39] assessed the incremental value of using WMC with gastric scintigraphy. We found low SOE to suggest that WMC is associated with modest improvement in diagnostic accuracy over use of scintigraphy alone for patients with suspected gastroparesis. We also found low SOE to support the incremental benefit of WMC in evaluation of transit times and pressure patterns. The two studies that did attempt to address this question had a method of data collection that may not have allowed for full understanding of diagnostic discrepancy. Discrepancy exists when one test shows disease and the other test does not show disease. The authors assumed that in a population of patients with gastroparesis, diagnostic gain (when WMC was positive but scintigraphy was not) was always present when there was discrepancy with results.[34] This assumption is difficult to confirm without an independent gold standard for establishing the diagnosis.

While few studies addressed this question specifically, the ones that did were among the better-quality studies, and demonstrated independent review of WMC and scintigraphy. We assessed risk of bias as medium and felt these studies were consistent and direct. We felt that precision was low but this is difficult to gauge for this question. The overall SOE was low for this KQ.

It is very hard to prove an incremental benefit of the test when studies use it in addition to other testing modalities because it is hard to determine how the study performed clinical decisionmaking. It may be unclear which test the clinician used to form an opinion of the case, and it may be unclear how much the incremental information contributed to the decisionmaking process. The retrospective nature of studies also limited the strength of evidence (SOE).

In addition, understanding the incremental benefit of WMC when added to gastric scintigraphy should take into account the fact that eligibility criteria for these studies required a previous positive test for gastric emptying scintigraphy and documented gastroparetic symptoms. Therefore, added WMC testing showed incremental sensitivity over scintigraphy alone in such a population, which one should take into account when judging these results' clinical applicability.

The incremental benefit for WMC in diagnostic evaluation of suspected gastroparesis is consistent with the nature of the disorder and the tests, since WMC offers pressure data and motility data that scintigraphy alone cannot detect, as well as lower gastrointestinal motility data, which can be implicated as a cause of symptoms in patients with combinations of motility disorders. One may obtain measurable benefit from the additional reported information in combination with scintigraphy, especially with regard to identification of a more diffuse motility disorder. The evidence was limited and there was no information to guide any conclusions regarding treatment decisions, utilization, patient-centered outcomes, or harms when evaluating the incremental value of also using WMC.

Incremental Value of WMC Compared with Antroduodenal Manometry Alone or Endoscopy Alone

We did not find any studies that evaluated the incremental value of adding the WMC test to testing with either antroduodenal manometry or endoscopy in patients with suspected gastroparesis.

Key Question 3. Evaluation of Colonic Dysmotility: WMC Alone Versus Other Diagnostic Tests

WMC Versus ROM

The SOE was low from five studies (306 total patients) comparing WMC with ROM in terms of their ability to accurately diagnose slow-transit constipation,[11, 25, 33, 34, 42] The diagnostic accuracy of WMC was similar to scintigraphy. (Concordance was about 80 percent in two of the larger studies.) Sensitivity and specificity were estimated to be 46 and 95 percent for WMC compared

with a symptom-based diagnosis of clinical constipation, and 37 and 95 percent for ROM.[11] WMC was comparable to ROM in assessing diagnostic accuracy, and matched the sensitivity in different target populations in a reliable way.

The SOE was low to suggest that the colonic transit time estimated by WMC correlates well with the colonic transit times recorded by ROM. The correlation coefficients between these two measures ranged from 0.69 to 0.71.

The SOE was low regarding the effect of WMC testing on treatment decisions based on ROM testing. We graded the SOE as low because only two retrospective chart reviews offered information about change in management for WMC compared with ROM.[33, 34] These two studies differed in the patient populations and the reporting of the outcomes. One of the studies reported few events, providing imprecise results. The data was further limited because not all patients underwent both diagnostic tests of interest. We found low SOE that WMC can affect resource utilization.

The SOE was low in the five studies reporting on any harms relevant to WMC or ROM.[11, 25, 34, 40, 42] Studies infrequently reported harms and adverse events for WMC or ROM. WMC is comparable to ROM with regard to low frequency of harms, as no studies reported serious adverse events or mortality. ROM testing involves exposure to at least one x ray by definition. A small proportion of patients who received WMC needed x rays on day 21 by protocol when the capsule had not spontaneously passed, but this may not be necessary in practice if someone witnesses capsule passage. Prototype devices suffered technical failure rates of 3 and 10 percent, depending on the study.[11] Studies also reported harms or adverse events, such as dysphagia, abdominal discomfort, bloating, or nausea, which happened infrequently. These all resolved spontaneously when reported.[25]

The SOE was insufficient to permit any conclusions about patient-centered outcomes like symptom improvement, quality of life, or patient satisfaction. No included studies addressed these outcomes of interest. These are difficult outcomes to assess without using dedicated symptom scores or mining large sources of data on hospital and physician visits. We will need longer-duration studies to address questions about change in quality of life or symptoms, which requires assessment along multiple time points.

Many factors contributed to the overall grading of evidence for outcomes we assessed as having low SOE in reference to KQ 3. We considered the evidence to have moderate risk of bias because many of the studies were retrospective, lacked random patient selection, did not report if there was blinding of assessment, and did not apply the same reference standard to all the patients. Furthermore, many studies recruited patients from academic referral centers; it is uncertain if the available examinations of motility testing captured the full spectrum of patients. Most studies had limited followup duration, which hampered our ability to draw conclusions about some of the outcomes that are important to patients such as patient satisfaction or change in symptom scores. We had only imprecise estimates of the effects on treatment decisions and harms. Our conclusions were limited by how studies defined the nonreference standards. The non-reference standard test was often a community-based ROM study of varying protocol. The multiple protocols had different assessment methods, which could have influenced the results. We could not conduct a meta-analysis because of the heterogeneity of reported data and patient populations in the studies. Although the SOE was low, it is impressive how well these devices correlated given limitations of the studies.

Much like scintigraphy as compared to WMC, ROM and WMC assess different components of transit. Some of the points of assessment coincide and provide comparable data, but the additional pressure and transit data offered by WMC make it a different and possibly complementary modality. Overall, the studies showed diagnostic agreement between WMC and ROM for assessment and diagnosis of slow-transit constipation.

WMC Versus Colonic Scintigraphy

We found no evidence to evaluate the WMC in comparison with colonic scintigraphy in patients with suspected slow-transit constipation. We excluded existing studies on scintigraphy from our analysis because they compared testing in healthy subjects separately from those with constipation or slow-transit constipation and thus were not eligible for inclusion.

KQ 4. Evaluation of Colonic Dysmotility: WMC in Combination With Other Diagnostic Tests Versus Other Diagnostic Tests Alone

No studies directly addressed any outcomes of interest related to KQ 4. The small amounts of data that were available from small trials about these outcomes were heterogeneous and did not specify the specific patient populations of interest; thus, it was impossible to generalize based on these data. One could use diagnostic gain to assess the incremental value of a new technology. However, when trying to judge whether a new test can be a

replacement or an adjunct to an old test, it is difficult to get a clear picture of which test was most helpful in making a diagnosis without a blinded comparison or without a followup study capable of assessing the validity of the diagnosis and or treatment effects over time.

Applicability

Limiting the application of the literature is the fact that all studies occurred at referral centers and that all prospective studies involved patients with known disease (thereby providing no prospective testing of WMC as a diagnostic tool). When a study used a comparison group without constipation or gastroparesis, it included "healthy" controls instead of patients who may have similar presenting symptoms but who do not have constipation or gastroparesis. These controls tended to be college-age men compared with middle-age females with suspected disease. Additionally, it is unclear how previous treatments or comorbidity, including diabetes, affect test performance or how the test results ultimately affect management.

Limitations and Strengths of Our Review Process

Our review had three major limitations:

1. No standards exist in the field of motility assessment for determining the minimum improvement of diagnostic accuracy that will identify one test as superior to another test. There are also no standards to establish the equivalence of motility tests. We arbitrarily chose a 10 percent difference in sensitivity or specificity as a potentially important difference between tests.[25] We felt that this threshold was a conservative minimum improvement over a reference standard with moderate diagnostic accuracy (between 50 and 80 percent). If the reference standard had a larger diagnostic accuracy (90 percent or greater), a 10 percent absolute difference is too large to expect.
2. We excluded studies that included non-diseased participants exclusively, because our review focused on studies that compared the diagnostic accuracy of the tests for patients with gastroparesis or slow-transit constipation. We recognize that many of the most commonly cited studies in the field included non-diseased participants exclusively.[12, 13, 51-64] Thus, we excluded a number of studies that evaluated characteristics of WMC.
3. Experts in the field acknowledge that scintigraphy and ROM have imperfect diagnostic accuracy. There are several options to account for the imperfection of the reference standard.65 We chose to incorporate two of these in our review: (1) We presented the results as if the reference standard had no measurement error and acknowledged this imperfection. (2) We presented concordance of the test results when available. We did not attempt to adjust the results to correct for the measurement error. This adjustment would have required assumptions that we did not have sufficient data to justify. Another option is to examine patient outcomes according to WMC. We had included patient outcomes (need for medications, additional tests) as outcomes in our review. Unfortunately, we found few studies evaluating these outcomes.

The major strength of our review process was its comprehensiveness. We included abstracts, contacted industry for unpublished studies, and contacted study authors for missing data.

Limitations of the Identified Literature

Our aim was to compare the diagnostic accuracy of WMC with other testing modalities to diagnose and manage gastroparesis and slow transit constipation. The identified literature limited our ability to answer our KQs for several reasons:

1. No study directly addressed the incremental value of using WMC in addition to ROM or scintigraphy in the evaluation of colonic dysmotility (KQ 4). Only limited data addressed the incremental value of using WMC in addition to gastric scintigraphy, antroduodenal manometry, or endoscopy in the evaluation of gastric dysmotility (KQ 2).
2. All study sites were referral centers that tend to have patients with more severe disease. The study results have limited generalizability to general gastroenterology or primary care clinics where there is a greater spectrum of disease severity. The sensitivity and specificity of WMC may be different in referral center settings than in other settings, and the positive and negative predictive values will be different when the prevalence of disease is different.
3. Many studies included nondiseased patients in the comparison of the diagnostic accuracy of WMC with other tests, using a clinical diagnosis of disease as the reference standard rather than the results of the other diagnostic tests.
4. The non-diseased participants had demographic characteristics very different from the gastroparesis and slow-transit-constipation patients. For example, the majority of the non-diseased participants were

college-age males, whereas the gastroparesis and slow-transit-constipation patients were middle-age women. Using clinical diagnosis as the reference standard, it is difficult to determine if WMC and other tests are distinguishing disease from non-disease or measuring differences in motility by demographic differences such as age and sex.

5. Variability in the administration of the motility tests and outcome assessments may explain some of the heterogeneity in the study results. Many studies used similar protocols to perform WMC testing and other tests, but with slight modifications such as the contents of the meal. Frequently, the timing of the motility assessment differed for WMC and the alternative test within and between studies, which may explain differences in the test results and the diagnostic accuracy differences between studies.
6. The abstracts we included did not report enough data to allow us to fully understand the study population, answer our KQs, and assess the quality of the studies.
7. We were unable to compare the results of studies with and without industry or investigator conflicts of interest because the company that manufactures the WMC sponsored most of the studies. The other studies did not report on conflicts of interest. No study stated that it was performed independent of industry sponsorship with authors who had no previous or current financial relationships with the manufacturer of the WMC.
8. Many studies included patients with gastroparesis defined by clinical symptoms and a prior abnormal gastric scintigraphy via local standards; however, symptoms of gastroparesis can be non-specific and many local facilities do not follow a standardized gastric scintigraphy protocol. As such, it is difficult, based on the data, to separate patients with gastroparesis from those with functional dyspepsia or other functional gastrointestinal disorders. This may have, to some degree, affected data with regards to sensitivity, specificity, and device correlation.
9. We attempted to assess publication bias by contacting the manufacturer of the WMC and requesting any unpublished data, but received no response.
10. Not all studies reported sufficient numbers to describe all the combinations of test results; some only provided means or medians. This hampered our ability to perform analyses, especially when analyzing combinations of tests.
11. Very few studies reported on patient-centered outcomes, limiting our abilities to draw conclusions on these outcomes.

Future Research Needs

Future research should ideally concentrate on finding a cure to these diseases that is nontoxic, cheap, easily available, and safe without major surgery or implanted devices. As far as diagnostic testing, the goal is always to find accurate, effective, and inexpensive tools to diagnose or exclude cases and qualify their severity in a reproducible way, especially when treatment is expensive, unavailable, or accompanied by great risks. Studies that compare the diagnostic modalities should have blinded interpretation of the results and make every attempt to classify patients by identical criteria and standardized protocols that other centers can repeat and verify. We recommend that research focus more on prospectively studied patients in larger numbers with an appropriate spectrum of symptoms and adequate followup to determine whether the diagnosis was accurate over time. Due to the difficulty enrolling patients, studies should carefully craft retrospective analyses.

We need research studies that evaluate how clinicians should use the WMC in combination with or instead of other testing modalities for evaluating slow-transit constipation. The studies we reviewed used alternative measures to assess anorectal function, such as anorectal manometry, as WMC does not capture data about this region reliably. Thus, clinicians will likely use WMC in combination with this test.

Eventually, we need outcomes studies to see if testing helps to improve quality of life or symptom control. It is unclear at present whether a more sensitive diagnostic test might just provide lead-time bias—or apparent superiority for an earlier diagnosis—but not actually change the outcomes or management steps overall for the patient. As we identify other targeted therapies, we will need to reassess the value of testing. We are aware that a new therapy is in Stage II trials for patients with diabetes and gastric emptying delay, which may increase the need for research into this area if it becomes available for use.[66] Currently, most patients with nausea- and vomiting-predominant symptoms of gastroparesis receive similar first-line treatment with antiemetics or prokinetics. As treatment options for gastroparesis expand (some at great expense), then more accurate detection of disease prior to initiation of therapy may play a more prominent role in disease management. The literature does not currently report resource utilization with and without WMC—we will need more studies evaluating these measures.

Little data is available to determine the optimal timing of WMC testing in the diagnostic and therapeutic approach to patients with symptoms of possible gastroparesis or slow-transit constipation. We need to do further work to classify the types of patients within subgroups of gastroparesis or slow-transit constipation in order to identify severe cases that may need more urgent evaluation. Finally, little is known about whether physicians should use testing to assess the effectiveness of treatment or if subsequent testing would offer any benefit in long-term management of patients. Currently, symptoms and symptom resolution guide therapeutic decisions, but these require careful interpretation.

Conclusions

Based on the current literature, WMC appears to be accurate in detection of gastroparesis and slow-transit constipation and may provide increased diagnostic gain as compared with standard motility testing. While the SOE is low, the data were relatively consistent and suggested that this modality is no less sensitive than conventional testing. The evidence is insufficient to determine whether use of WMC will improve outcomes of care.

References

1. Hasler WL. Gastroparesis: pathogenesis, diagnosis and management. Nat Rev Gastroenterol Hepatol. 2011;8(8):438-53. PMID: 21769117.
2. Camilleri M, Bharucha AE, Farrugia G. Epidemiology, mechanisms, and management of diabetic gastroparesis. Clin Gastroenterol Hepatol. 2011;9(1):5-12; quiz e7. PMID: 20951838.
3. Jung HK. The incidence, prevalence, and survival of gastroparesis in Olmsted County, Minnesota, 1996-2006 (gastroenterology 2009;136:1225-1233). J Neurogastroenterol Motil. 2010 Jan;16(1):99-100. PMID: 20535336.
4. Rey E, Choung RS, Schleck CD, et al. Prevalence of hidden gastroparesis in the community: the gastroparesis "iceberg". J Neurogastroenterol Motil. 2012 Jan;18(1):34-42. PMID: 22323986.
5. Wang YR, Fisher RS, Parkman HP. Gastroparesis-related hospitalizations in the United States: trends, characteristics, and outcomes, 1995-2004. Am J Gastroenterol. 2008;103(2):313-22. PMID: 18047541.
6. Quigley EM, Hasler WL, Parkman HP. AGA technical review on nausea and vomiting. Gastroenterology. 2001;120(1):263-86. PMID: 11208736.
7. Guo JP, Maurer AH, Fisher RS, et al. Extending gastric emptying scintigraphy from two to four hours detects more patients with gastroparesis. Dig Dis Sci. 2001;46(1):24-9. PMID: 11270790.
8. Abell TL, Camilleri M, Donohoe K, et al. Consensus recommendations for gastric emptying scintigraphy: a joint report of the American Neurogastroenterology and Motility Society and the Society of Nuclear Medicine. Am J Gastroenterol. 2008;103(3): 753-63. PMID: 18028513.
9. Wang TC, Fleischer DE, Kaufman PN, et al. The best of times and the worst of times: sustaining the future of academic gastroenterology in the United States--Report of a Consensus Conference Conducted by the AGA Institute Future Trends Committee. Gastroenterology. 2008 Feb;134(2):597-616. PMID: 18242223.
10. Kuo B, McCallum RW, Koch KL, et al. Comparison of gastric emptying of a nondigestible capsule to a radio-labelled meal in healthy and gastroparetic subjects. Aliment Pharmacol Ther. 2008;27(2):186-96. PMID: 17973643.
11. Rao SS, Kuo B, McCallum RW, et al. Investigation of colonic and whole-gut transit with wireless motility capsule and radiopaque markers in constipation. Clin Gastroenterol Hepatol. 2009;7(5): 537-44. PMID: 19418602.
12. Williams RE, 3rd, Bauman WA, Spungen AM, et al. SmartPill technology provides safe and effective assessment of gastrointestinal function in persons with spinal cord injury. Spinal Cord. 2012;50(1):81-4. PMID: 21876549.
13. Cassilly D, Kantor S, Knight LC, et al. Gastric emptying of a non-digestible solid: assessment with simultaneous SmartPill pH and pressure capsule, antroduodenal manometry, gastric emptying scintigraphy. Neurogastroenterol Motil. 2008;20(4):311-9. PMID: 18194154.
14. Stewart WF, Liberman JN, Sandler RS, et al. Epidemiology of constipation (EPOC) study in the United States: relation of clinical subtypes to sociodemographic features. Am J Gastroenterol. 1999 Dec;94(12):3530-40. PMID: 10606315.
15. Suares NC, Ford AC. Prevalence of, and risk factors for, chronic idiopathic constipation in the community: systematic review and meta-analysis. Am J Gastroenterol. 2011 Sep;106(9):1582-91; quiz 1, 92. PMID: 21606976.
16. Qureshi W, Adler DG, Davila RE, et al. ASGE guideline: guideline on the use of endoscopy in the management of constipation. Gastrointest Endosc. 2005;62(2):199-201. PMID: 16046978.
17. Longstreth GF, Thompson WG, Chey WD, et al. Functional bowel disorders. Gastroenterology. 2006 Apr;130(5):1480-91. PMID: 16678561.
18. Rao SS, Camilleri M, Hasler WL, et al. Evaluation of gastrointestinal transit in clinical practice: position paper of the American and European Neurogastroenterology and Motility Societies. Neurogastroenterol Motil. 2011;23(1):8-23. PMID: 21138500.
19. Ribas Y, Saldana E, Marti-Rague J, et al. Prevalence and pathophysiology of functional constipation among women in Catalonia, Spain. Dis Colon Rectum. 2011 Dec;54(12):1560-9. PMID: 22067186.
20. Wald A. Pathophysiology, diagnosis and current management of chronic constipation. Nat Clin Pract Gastroenterol Hepatol. 2006;3(2):90-100. PMID: 16456575.

21. Rao SS, Ozturk R, Laine L. Clinical utility of diagnostic tests for constipation in adults: a systematic review. Am J Gastroenterol. 2005;100(7):1605-15. PMID: 15984989.

22. Brandt LJ, Prather CM, Quigley EM, et al. Systematic review on the management of chronic constipation in North America. Am J Gastroenterol. 2005;100 Suppl 1:S5-S21. PMID: 16008641.

23. van der Sijp JR, Kamm MA, Nightingale JM, et al. Radioisotope determination of regional colonic transit in severe constipation: comparison with radio opaque markers. Gut. 1993;34(3):402-8. PMID: 8472991.

24. Pomerri F, Frigo AC, Grigoletto F, et al. Error count of radiopaque markers in colonic segmental transit time study. AJR Am J Roentgenol. 2007;189(2):W56-9. PMID: 17646438.

25. Camilleri M, Thorne NK, Ringel Y, et al. Wireless pH-motility capsule for colonic transit: prospective comparison with radiopaque markers in chronic constipation. Neurogastroenterol Motil. 2010;22(8):874-82, e233. PMID: 20465593.

26. Tack J, Muller-Lissner S, Stanghellini V, et al. Diagnosis and treatment of chronic constipation--a European perspective. Neurogastroenterol Motil. 2011;23(8):697-710. PMID: 21605282.

27. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36. PMID: 22007046.

28. Agency for Healthcare Research and Quality. Methods Guide for Medical Test Reviews. Rockville, MD. (published draft). Final: AHRQ Publication No 12-EHC017. Rockville, MD; June 2012 Chapters available at: www.effectivehealthcare.ahrq.gov

29. Food and Drug Administration. Guidance for Industry and FDA Staff: Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests. 2007. Available at: www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071148.htm.

30. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. J Clin Epidemiol. 2010;63(5):513-23. PMID: 19595577.

31. Singh S, Chang SM, Matchar DB, et al. Chapter 7: Grading a Body of Evidence on Diagnostic Tests. J Gen Intern Med. 2012;Jun(27):47-55.

32. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6. PMID: 18436948.

33. Kuo B, Maneerattanaporn M, Lee AA, et al. Generalized transit delay on wireless motility capsule testing in patients with clinical suspicion of gastroparesis, small intestinal dysmotility, or slow transit constipation. Dig Dis Sci. 2011;56(10):2928-38. PMID: 21625964.

34. Rao SS, Mysore K, Attaluri A, et al. Diagnostic utility of wireless motility capsule in gastrointestinal dysmotility. J Clin Gastroenterol. 2011;45(8):684-90. PMID: 21135705.

35. Brun M, Wilding GE, Surjanhata B, et al. Performance characteristics of gastric emptying test: Variability of gastric emptying results by two different technologies-gastric emptying scintigraphy (GES) and Wireless Motility Capsule (WMC) in healthy and gastroparetics. Gastroenterology. 2011;140(5):S804-S5.

36. Lee A, Michalek W, Wiener SM, et al. Clinical impact of a wireless motility capsule - A retrospective review. Gastroenterology. 2010;138(5):S481.

37. Reddymasu S, Semler JR, McCallum R. Postprandial gastric motility parameters assessed by the wireless motility capsule method are complimentary to gastric transit time measurement of a standardized meal for the diagnosis of gastroparesis. Gastroenterology. 2010;138(5):S714.

38. Lee A, Michalek W, Wong C, et al. Clinical impact of an ambulatory motility capsule-retrospective review. Neurogastroenterol Motil. 2009;21:73.

39. Lee A, Wilding G, Kuo B. Variable abnormal physiological motility in the proximal upper gastrointestinal tract in gastroparesis. Neurogastroenterol Motil. 2012 Mar 14. PMID: 22417117.

40. Rao SS, Paulson JA, Donahoe R, et al. Can assessment of colonic motility with wireless ph/pressure capsule (SmartPill®) distinguish subtypes of chronic constipation? Gastroenterology. 2009;136(5):A223.

41. Rao SS, Paulson JA, Saad RJ, et al. Assessment of colonic, whole gut and regional transit in elderly constipated and healthy subjects with a novel wireless pH/pressure capsule (SmartPill®). Gastroenterology. 2009;136(5):A144.

42. Rao SS, Coss-Adame E, Valestin J, et al. Evaluation of constipation in older adults: Radioopaque markers (ROMs) versus wireless motility capsule (WMC). Arch Gerontol Geriatr. 2012 May 7. PMID: 22572600.

43. Paulson J, Rao S, Donahoe R, et al. Can wireless pH/pressure capsule (SmartPill® (SP)) distinguish subtypes of chronic constipation? Neurogastroenterol. Motil. 2009;21:39.

44. Hinton JM, Lennard-Jones JE, Young AC. A new method for studying gut transit times using radioopaque markers. Gut. 1969 Oct;10(10):842-7. PMID: 5350110.

45. Metcalf AM, Phillips SF, Zinsmeister AR, et al. Simplified assessment of segmental colonic transit. Gastroenterology. 1987 Jan;92(1):40-7. PMID: 3023168.

46. Evans RC, Kamm MA, Hinton JM, et al. The normal range and a simple diagram for recording whole gut transit time. Int J Colorectal Dis. 1992 Feb;7(1):15-7. PMID: 1588218.

47. Saad RJ, Rao SS, Koch KL, et al. Do stool form and frequency correlate with whole-gut and colonic transit? Results from a multicenter study in constipated individuals and healthy controls. Am J Gastroenterol. 2010;105(2):403-11. PMID: 19888202.

48. Hasler WL, Saad RJ, Rao SS, et al. Heightened colon motor activity measured by a wireless capsule in patients with constipation: relation to colon transit and IBS. Am J Physiol Gastrointest Liver Physiol. 2009;297(6):G1107-14. PMID: 19808653.

49. Mysore KR, Attaluri A, Valestin J, et al. How useful is wireless motility capsule in diagnosis of gastrointestinal dysmotility? Neurogastroenterol. Motil. 2010;22:37-8.

50. Mysore KR, Attaluri A, Valestin J, et al. Evaluation of diagnostic utility of a wireless motility capsule in gastrointestinal dysmotility. Gastroenterology. 2010;138(5):S233.

51. Brun M, Michalek W, Surjanhata B, et al. Small bowel transit time (Sbtt) by Wireless Motility Capsule (WMC): Normal values and analysis of pressure profiles in different subgroups of patients with slow sbtt. Gastroenterology. 2011;140(5):S865.

52. Brun R, Michalek W, Surjanhata BC, et al. Comparative analysis of phase III migrating motor complexes in stomach and small bowel using wireless motility capsule and antroduodenal manometry. Neurogastroenterol Motil. 2012 Apr;24(4):332-e165. PMID: 22292793.

53. Maqbool S, Parkman HP, Friedenberg FK. Wireless capsule motility: comparison of the SmartPill GI monitoring system with scintigraphy for measuring whole gut transit. Dig Dis Sci. 2009;54(10):2167-74. PMID: 19655250.

54. Michalek W, Kuo B. Analysis of upper GI migrating motor complexes using invasive and non-invasive techniques. Neurogastroenterol. Motil. 2010;22:64-5.

55. Michalek W, Neuman S, Kloetzer L, et al. Impact of acid suppression on upper gastrointestinal function as measured by a non-invasive wireless pH and motility capsule. Gastroenterology. 2009;136(5):A186-A7.

56. Michalek W, Semler JR, Kuo B. Impact of acid suppression on upper gastrointestinal pH and motility. Dig Dis Sci. 2011;56(6):1735-42. PMID: 21086166.

57. Mikolajczyk A, Surma B, Rubin D. Assessment of tandem measurements of PH and total gut transit time in healthy volunteers. Am J Gastroenterol. 2011;106:S502-S3.

58. Mreyoud A, Rozov I, Moore J, et al. Assessment of drug effects on gastric emptying and contractility using wireless capsule manometry. Gastroenterology. 2009;136(5):A536.

59. Saad RJ, Semler JR, Wilding GE, et al. The effect of age on regional and whole gut transit times in healthy adults. Gastroenterology. 2010;138(5):S127.

60. Sarosiek I, Alvarez A, Romero R, et al. Prolonged cecal residence time identified by wireless technology: A new symptoms explantation for some patients with chronic constipation. Neurogastroenterol Motil. 2011;23:22-3.

61. Timm DA, Willis H, Thomas W, et al. The use of a new wireless motility device (SmartPill(registered trademark)) for measurement of gastrointestinal transit time after dietary fiber intervention. Gastroenterology. 2010;138(5):S462.

62. Timm D, Willis H, Thomas W, et al. The use of a wireless motility device (SmartPill(R)) for the measurement of gastrointestinal transit time after a dietary fibre intervention. Br J Nutr. 2011;105(9): 1337-42. PMID: 21138605.

63. Willis HJ, Thomas W, Willis DJ, et al. Feasibility of measuring gastric emptying time, with a wireless motility device, after subjects consume fiber-matched liquid and solid breakfasts. Appetite. 2011;57(1):38-44. PMID: 21435365.

64. Zarate N, Mohammed SD, O'Shaughnessy E, et al. Accurate localization of a fall in pH within the ileocecal region: validation using a dual-scintigraphic technique. Am J Physiol Gastrointest Liver Physiol. 2010;299(6):G1276-86. PMID: 20847301.

65. Trikalinos TA, Balion CM. Chapter 9: options for summarizing medical test performance in the absence of a "gold standard". J Gen Intern Med. 2012 Jun;27 Suppl 1:S67-75. PMID: 22648677.

66. ClinicalTrials.gov. Phase 2 Study to Evaluate Safety & Efficacy of RM-131 Administered to Patients With Diabetic Gastroparesis, Rhythm Pharmaceuticals, Inc. Bethesda (MD): National Library of Medicine (U.S.). 2012;2012 May 4.

Full Report

This executive summary is part of the following document: Stein E, Berger Z, Hutfless S, Shah L, Wilson LM, Haberl E, Bass EB, Clarke JO. Wireless Motility Capsule Versus Other Diagnostic Technologies for Evaluating Gastroparesis and Constipation: A Comparative Effectiveness Review. Comparative Effectiveness Review No. 110. (Prepared by Johns Hopkins Evidence-based Practice Center under Contract No. 290 2007 10061-I.) AHRQ Publication No. 13-EHC060-EF. Rockville, MD: Agency for Healthcare Research and Quality; May 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

AHRQ Pub. No. 13-EHC060-1
May 2013