Effective Health Care Program

Future Research Needs Paper
Number 27

Wireless Motility Capsule Versus Other Diagnostic Technologies for Evaluating Gastroparesis and Constipation: Future Research Needs

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

Future Research Needs Paper

Number 27

Wireless Motility Capsule Versus Other Diagnostic Technologies for Evaluating Gastroparesis and Constipation: Future Research Needs

Identification of Future Research Needs From Comparative Effectiveness Review No. 110

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-2007-10061-I

Prepared by:
Johns Hopkins University Evidence-based Practice Center
Baltimore, MD

Investigators:
Ellen Stein, M.D.
John O. Clarke, M.D.
Susan Hutfless, Ph.D.
Chad Boult, M.D., M.P.H., M.B.A.
David Doggett, Ph.D.
Oluwakemi A. Fawole, M.B., Ch.B., M.P.H.
Eric B. Bass, M.D., M.P.H.

AHRQ Publication No. 13-EHC018-EF
May 2013

This report is based on research conducted by the Johns Hopkins Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10061-I). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care researchers and funders of research make well-informed decisions in designing and funding research and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of scientific judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical research and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact EffectiveHealthCare@ahrq.hhs.gov.

None of the investigators have any affiliation or financial involvement that conflicts with the material presented in this report.

Suggested citation: Stein E, Clarke JO, Hutfless S, Boult C, Doggett D, Fawole OA, Bass EB. Wireless Motility Capsule Versus Other Diagnostic Technologies for Evaluating Gastroparesis and Constipation: Future Research Needs. Future Research Needs Paper No. 27. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-2007-10061-I.) AHRQ Publication No. 13-EHC018-EF. Rockville, MD: Agency for Healthcare Research and Quality. May 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

An important part of evidence reports is to not only synthesize the evidence, but also to identify the gaps in evidence that limited the ability to answer the systematic review questions. AHRQ supports EPCs to work with various stakeholders to identify and prioritize the future research that is needed by decisionmakers. This information is provided for researchers and funders of research in these Future Research Needs papers. These papers are made available for public comment and use and may be revised.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The evidence reports undergo public comment prior to their release as a final report.

We welcome comments on this Future Research Needs document. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director, EPC Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Christine Chang, M.D., M.P.H.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

Contributors

Thomas Abell, M.D.
University of Mississippi Medical Center
Jackson, MS

Mary-Angela DeGrazia-DiTucci
Patient and Advocate
Association of Gastrointestinal Motility Disorders, Inc.
Bedford, MA

Gianrico Farrugia, M.D.
Mayo Clinic
Rochester, MN

Lauren B. Gerson, M.D., M.Sc.
Stanford University School of Medicine
Palo Alto, CA

Richard McCallum, M.D.
Texas Tech University
El Paso, TX

Henry Parkman, M.D.
Temple University School of Medicine
Philadelphia, PA

Pankaj Jay Pasricha, M.D.
Stanford Medical Center
Palo Alto, CA

Satish Rao, M.D.
University of Iowa Health Care
Iowa City, IA

Holger Schunemann, M.D., Ph.D.
McMaster University
Hamilton, Ontario, Canada

Contents

Tables

Figures

Appendixes

Executive Summary

Background

Gastroparesis

Gastroparesis delays gastric emptying without a physical blockage.[1] Its symptoms are nausea, vomiting, early satiety, bloating, abdominal pain, and postprandial fullness.[2] Its prevalence is estimated to be 9.6 per 100,000 among men and 37.8 per 100,000 among women.[2] Between 1.5 and 3 million Americans are affected. Related hospitalizations increased by 158 percent between 1995 and 2004.[3] Evaluation may employ gastric emptying scintigraphy, antroduodenal manometry, and the wireless motility capsule (WMC) – and it guides nutritional, medical, and surgical therapies.

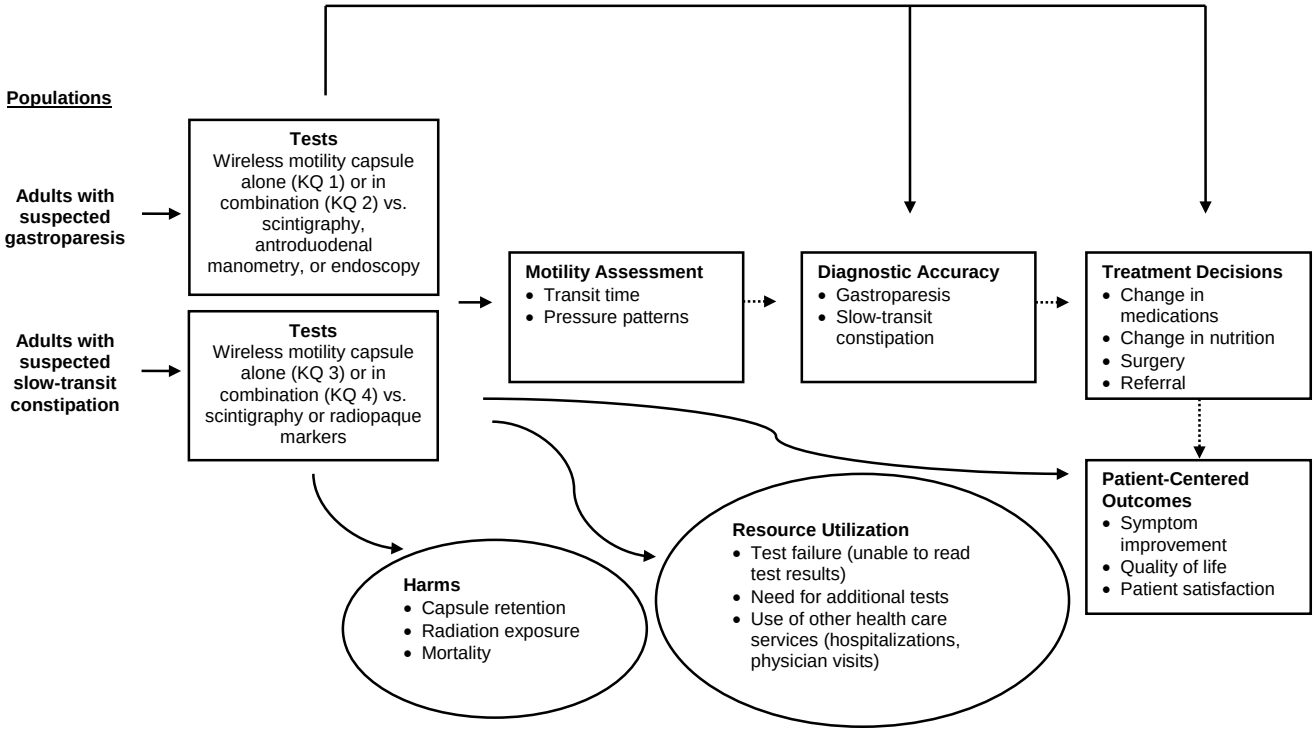
Constipation

Constipation is common, occurring in 15 to 20 percent of the population.[4-6] It is defined as fewer than two bowel movements per week, or a decrease in a person's normal frequency accompanied by straining, difficulty defecating, or passage of hard stools.[4] Patients with slow-transit constipation often have severe symptoms, with prolonged intervals between bowel movements, and may be refractory to standard therapies. Prevalence of slow-transit constipation is 0.03–0.17 percent.[7] For patients with refractory symptoms, colonic physiology testing may include radiopaque markers (ROM), colonic scintigraphy, manometry, and the WMC.[8,9]

Objectives of the Systematic Review

The WMC is a new modality for diagnosing gastric and colonic motility disorders. The Johns Hopkins Evidence-based Practice Center recently completed an Agency for Healthcare Research and Quality (AHRQ)-funded systematic review of the effectiveness of WMC compared with other tests of gastric and colonic motility (see analytic framework, Figure A). We also sought to define populations that would benefit most from motility testing.

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



KQ = Key Question

In the systematic review of this topic, we formulated Key Questions (KQ), reviewed the literature extensively, and obtained feedback from experts. The results are summarized in Table A.

Table A. Summary of the results from the systematic review on the wireless motility capsule

Topic and Key Questions	Key Findings	Strength of Evidence
Comparative diagnostic accuracy of WMC for gastroparesis. **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, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	Seven studies evaluated diagnosis of gastric emptying delay. We found low strength of evidence that WMC alone was comparable to scintigraphy for diagnostic accuracy, motility assessment, treatment decisions, and resource utilization.	**Low**
Incremental diagnostic accuracy for WMC in combination with other diagnostic methods for gastroparesis. **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 the WMC in terms of diagnostic accuracy of gastric emptying delay, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	We found two studies evaluating WMC as an add-on to other testing. The strength of evidence was low for diagnostic accuracy and motility assessment of WMC in combination with other modalities. The addition of WMC increased diagnostic yield.	**Low**
Comparative diagnostic accuracy of WMC for slow-transit constipation. **KQ 3:** In the evaluation of colonic dysmotility, how does the WMC alone compare with radiopaque markers and scintigraphy in terms of diagnostic accuracy of slow-transit constipation, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	Nine studies analyzed colon transit disorders and provided moderate strength of evidence for diagnostic accuracy, motility assessment, and harms. WMC was comparable to radiopaque markers. Few harms were reported. The evidence was insufficient to draw conclusions about effects of WMC on treatment decisions and resource utilization.	**Low**
Incremental diagnostic accuracy for WMC in combination with other diagnostic methods for slow-transit constipation. **KQ 4:** When a radiopaque marker or scintigraphy is used in the evaluation of colonic dysmotility, what is the incremental value of also using the WMC in terms of diagnostic accuracy of slow-transit constipation, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	No studies directly assessed use of WMC in combination with other tests to detect colon transit delay.	**Insufficient**

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

Overall, the strength of evidence regarding the ability of WMC to detect gastroparesis or slow-transit constipation was graded as low (see KQs listing in Table A). The main limitations were inconsistencies in reporting the performance of motility testing modalities. Great variability existed in administering diagnostic tests and in assessing those tests. No uniform standards define differences in diagnostic accuracy, so we arbitrarily chose a 10 percent difference in sensitivity or specificity for reference standards, such as gastric scintigraphy, and device concordance for non-reference standards, such as ROM.

Most of the "normal" subjects upon which the tests were validated were college-age men, while most of the patients with suspected gastroparesis or constipation were women over the age

of 50 years. Since the population of interest comprised motility patients, we excluded studies that included only nondiseased participants.

The major strength of the review was its comprehensiveness. We reviewed abstracts, queried industry sources for unpublished studies, and contacted study authors for missing data.

Conclusions of the Systematic Review

WMC is comparable to other modalities in use for detecting delayed gastric emptying and slow-transit constipation. Data are insufficient to determine the optimal timing of WMC in diagnostic algorithms.

Methods

The objectives of the Future Research Needs (FRN) project were to identify the evidence gaps highlighted by the results of the systematic review and to create a set of prioritized FRN to guide stakeholders in future decisions.

Evidence Gap Identification

Evidence gaps were identified in the review writing process based on the strength of evidence, applicability, and limitations of the review. Individuals who contributed to review writing met multiple times and circulated by email lists of potential questions to identify gaps. This process developed a list of research gaps to be presented to the stakeholders.

Stakeholder Engagement for Additional Gap Identification and Prioritization

Stakeholder Identification

Important stakeholder categories to include are patients/advocates, clinical experts, and payers. Stakeholders from these categories were identified from the Key Informants and Technical Expert Panel members who had been the most responsive for the systematic review, as well as new participants suggested by the review investigators.

Orienting Stakeholders

The stakeholders were provided by email a description of the project, the draft of the executive summary of the review, and a web link to the complete draft report.

Engagement Round 1, Gap List Review and Preliminary Prioritization

The Evidence-based Practice Center's (EPC) list of research gaps were presented to the stakeholders by email for review and for suggestions of additional gaps within the scope of the systematic review. They were instructed to carry out a preliminary prioritization of the gaps. To perform this preliminary prioritization, they were asked to choose their top 5 choices and rank them in priority from 1 (highest) to 5 based on the criteria of (1) importance (prevalence and severity of condition, lack of or inadequacy of alternatives, burden of condition to patients and healthcare system), and (2) impact (potential to change practice and/or to improve clinical and patient outcomes).

Engagement Round 2, Final Prioritization

The team incorporated the stakeholder comments and additional suggestions from Engagement 1 into a final list of gaps for final prioritization. This list included the preliminary ranking from the previous engagement. Each stakeholder was presented with this list and asked to choose their top 5 choices and prioritize them as described above. Stakeholders were also asked if they were aware of any ongoing studies addressing the gaps (duplication), and they were encouraged to comment on the feasibility of research addressing the gaps.

Top-Tier Future Research Needs

The individual priority ratings of the stakeholders were summed for each question to get global ratings that were used to sort the questions by priority. The priority ranking was inspected by the EPC team to determine if there was an obvious cutpoint between a top tier of gaps and the remainder. If the global ranking was a continuum with no apparent cutpoint, the top half of the gaps or the top 10, whichever was fewer, was chosen as the top tier and is considered the FRN.

External Literature Searches

To identify ongoing clinical trials that may have addressed our Key Questions, we searched the World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch/) and clinicaltrials.gov for trials registered since the search cutoff date of the review.

Results

Knowledge Gaps

Using the information from the systematic review of the comparative effectiveness of WMC testing, knowledge gaps were identified in numerous areas with low strength of evidence. The reason that most strength of evidence (SOE) was low was a lack of fundamental knowledge about WMC measurements—for both normal patients and those with suspected or documented motility disorders, as well as a lack of fundamental knowledge about scintigraphy, manometry, and ROM measurements. The range of WMC values in normal patients, those suspected of motility disorders, and those with documented motility disorders is undefined. These foundational gaps may need to be addressed before clinical gaps can be addressed productively. Therefore, the EPC team identified not only clinical knowledge gaps derived directly from the Key Questions of the review (Table B), but also methodological gaps in the foundational knowledge of test results from normal and diseased patients (Table C).

The gap topics and questions were reviewed by the stakeholder panel in two lists to allow them to comment, suggest and rank important issues for both the directly important clinical issues, and the foundational and methodological questions.

The clinical research gaps (Table B) were framed to reflect the clinical gaps in the main evidence report more specifically as relevant to clinical providers. The highest ranked question reflected the uncertainty in the role of the WMC in outcomes for patients with presumed gastroparesis and the role of the WMC as a replacement test versus adjunct test for diagnosis. Consensus guidelines suggest that WMC testing is a replacement for current testing methods, however additional research would lend more weight to that argument. It is currently very difficult in active practice to get access to the WMC for some patients due to lack of insurance

coverage, but further research might be more convincing to payers that there is a benefit in this new test and might enable greater access to the WMC. This is clearly a high priority.

The other top ranked question asked whether these same patients with suspected gastroparesis would have comparable results from scintigraphy, anteroduodenal manometry, plain x-ray after marker blind capsule ingestion, endoscopy, or a combination of tests. Similarly, the next ranked questions asked to identify which test should serve as the gold standard or reference standard for comparison, and whether a confirmatory test is required after use of the WMC. Reproducibility of the WMC was also thought to be important and ties into the previous questions which tried to establish the role of WMC testing. Beyond these questions on gastroparesis, the next set of questions focused on slow-transit constipation, including establishing the role of WMC testing in diagnosis compared with other standard tests for ability to diagnose, as well as accuracy and safety in diagnosis.

The next series of questions asked about the incremental value of WMC testing in addition to other testing methods for gastric emptying delay and slow-transit constipation. Interestingly, the top tier included questions about correlating pathology with clinical history and WMC findings. This echoes the questions from the foundational side asking about the same correlation. Also, a focus was the role of WMC testing in colonic dysmotility to predict outcomes, or to predict the effects of medical and/or surgical therapy on outcomes.

The most highly ranked foundational questions (Table C) made clear that basic data ranges for WMC testing have yet to be firmly established in non-diseased populations and those with suspected gastric or colonic dysmotility. With better established norms for diseased and non-diseased patients, there can be better framing of future questions and research endeavors. The next most highly ranked questions were very similar to the initial research questions addressed in our evidence report, which had only low strength of evidence. With additional high quality research, the panel thought that the strengths and weaknesses of WMC testing would be more apparent. Thus, they suggested a priority area for research would be future studies that focus on establishing the role of WMC testing comparatively with scintigraphy, manometry, and radiopaque markers or that focus on obtaining information about test failure and the need for additional tests. It was also considered important to assess diagnostic accuracy of the test when used by non-academic specialists or as a front-line test. Other priority items included establishing the thresholds of diagnostic accuracy and establishing the basic science connection between WMC results and histopathological findings from patients with known disease, if one exists. All of these basic foundational questions were ranked highly by participants in both rounds. These are by definition high-priority areas for future research.

Table B. Wireless motility capsule stakeholder prioritized clinical knowledge gaps

Clinical Knowledge Gaps	Priority Rank[a] (1 = Highest Priority)
Top Tier	
1. Comparative effectiveness of managing with WMC (gastric dysmotility/gastroparesis)	1
2. Comparative diagnostic accuracy and safety (gastric dysmotility/gastroparesis)	2
3. Gold standard (gastric dysmotility/gastroparesis)	3
4. Comparison of WMC and scintigraphy results (gastroparesis)	4
5. Gold standard (colonic dysmotility/slow-transit constipation)	5
6. Reproducibility of WMC results (gastroparesis)	5
7. Comparative diagnostic accuracy and safety (colonic dysmotility/slow-transit constipation)	6
8. Incremental diagnostic accuracy (gastric dysmotility/gastroparesis)	7
9. Incremental diagnostic accuracy (colonic dysmotility/slow-transit constipation)	8
Other Gaps	
10.Distinction of subtypes of gastroparesis (gastric dysmotility/gastroparesis)	9
11. Best test or combination for predicting outcomes after colectomy (colonic dysmotility/slow-transit constipation)	9
12. Comparative accuracy in predicting response to treatment (colonic dysmotility/slow-transit constipation)	9
13. Utility of colonic pressure (colonic dysmotility)	10
14. Differentiating IBS-C from idiopathic constipation (colonic dysmotility/slow-transit constipation)	10
15. Comparisons of WMC pressure profiles (gastroparesis)	10
16. Best test or combination for predicting outcomes without colectomy (colonic dysmotility/slow-transit constipation)	11
17. Comparative accuracy in predicting response to treatment (gastric dysmotility/gastroparesis)	12
18. Comparative value in monitoring response to treatment (gastric dysmotility/gastroparesis)	13
19. Comparative value in monitoring response to treatment (colonic dysmotility/slow-transit constipation)	13
20. Colonic pressure patterns after waking (colonic dysmotility)	13
21. Evaluation of high-amplitude contractions (colonic dysmotility)	13

C = clinical knowledge gap; WMC = wireless motility capsule

[a]It was possible for multiple gaps to have the same priority rank if the summation scores for computing the ranks were the same (Appendix C).

Table C. Wireless motility capsule stakeholder prioritized methodological knowledge gaps

Methodological Knowledge Gaps	Priority Rank (1 = Highest Priority)
Top Tier	
1. WMC data in patients with gastric and colonic dysmotility	1
2. WMC data in nondiseased populations	2
3. Rates of test failures and need for additional tests	3
4. WMC diagnostic accuracy and safety in gastroparesis subtypes	4
5. Thresholds	5
6. Correlation between WMC values and histopathology	6
Other Gaps	
7. Colonic pressure patterns following meal ingestion.	7
8. Relation of baseline WMC values with long-term followup values	8
9. WMC data compared with other tests in nondiseased populations	9
10. Concurrent tests	10
11. Masking interpreters of WMC data	11
12. Use of historical controls in WMC research	12

M = methodological knowledge gap; WMC = wireless motility capsule

Research Needs

The gap topics and questions were reviewed by the stakeholder panel in two lists to allow them to comment, suggest and rank important issues for both the directly important clinical issues, and the foundational and methodological questions. We received suggestions for additional research questions, which were solicited at the first round of survey, and then subject to general review at the second round of questions. Our expert panel readily identified these areas as necessary and important during both rounds of feedback and, specifically, they focused on certain questions of greater importance. The global rankings demonstrate a delineation of the most highly valued and frequently chosen topics from the least valued and least frequently selected items. The top half of the priority ranked questions were chosen as the top tier, and are considered the highest priority research needs. Next we developed a list of research questions based on the research needs (i.e., a top tier of prioritized evidence gaps) with sufficient detail for use by researchers and funders (in PICOTS format: population, intervention, comparisons, outcomes, timing, setting), including recommendations on research designs that would best suit each research question (Table D). The clinical questions focus on the basic clinical needs in research on WMC testing. The methodological questions address the relationships of WMC data to other diagnostic test result ranges and to the basic biology of the normal and diseased

digestive tract. Other questions address basic research principles needed to study the WMC properly. The answers to the latter questions will provide firm underpinnings for future research.

Ongoing Research

We scanned trial registries (Appendix A) for any ongoing clinical trials which may have already addressed these high-priority areas (Appendix B). We found two clinicaltrials.gov references which had already published the results on populations outside of the scope of our review, spinal cord injury patients and critically ill intensive care patients. One clinicaltrials.gov reference was not assessing the role of wireless motility capsule in diagnosis of constipation or gastroparesis, but instead focused on acid measurement. We also identified a funded research protocol, which likely has yet to complete enrollment, regarding effect of medication for constipation on outcomes with wireless motility capsule. We await the outcome of this trial, but it is only representative of one of many treatment modalities available to these patients. No other ongoing research projects were identified which address the questions we designed.

Table D. Summary of research needs (top tier)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Clinical Questions								
1. Comparative effectiveness of managing with WMC (gastric dysmotility/ gastroparesis)	Among patients with presumed gastroparesis, how do clinical outcomes differ between patients managed with WMC data alone versus those managed with scintigraphy data alone?	Presumed gastro-paresis	WMC alone	4 hour gastric scintigraphy by standard protocol (ANMS consensus)	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization	Simul-taneous	Academic, multi-center practice	Randomized, controlled trial, prospective
2. Comparative diagnostic accuracy and safety (gastric dysmotility/ gastroparesis)	Among patients with suspected gastroparesis, how does the WMC compare with gastric scintigraphy, antroduodenal manometry, and endoscopy in its accuracy and safety in diagnosing gastroparesis?	Suspected gastro-paresis	WMC alone	4 hour gastric scintigraphy by standard protocol (ANMS consensus), or antroduodenal manometry and/or endoscopy	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic, multi-center practice	Randomized, controlled trial, prospective

Table D. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Clinical Questions (continued)								
3. Gold standard (gastric dysmotility/gastro-paresis)	What is an appropriate test (gold standard) to compare with WMC in the diagnosis and monitoring of patients with gastroparesis? (e.g., clinical diagnosis, 4-hour gastric scintigraphy, antroduodenal manometry, endoscopy, plain X-ray of the abdomen 5 hours after swallowing a "dumb" radiopaque pill, a combination of tests)	Suspected or known gastro-paresis subjects	WMC alone	Clinical history, standardized GI question-naire, 4 hour gastric emptying, antroduodenal manometry, endoscopy, plain x ray 5 hours after swallowing a patency capsule, or some combination above	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic, multicenter practice	Randomized controlled prospective trial, or prospectively enrolled but retrospectively reviewed cases, blinded interpretation of results
4. Comparison of WMC and scintigraphy results (gastroparesis)	In the clinical evaluation of patients with suspected gastroparesis, do abnormal WMC gastric emptying times correlate with 4-hour gastric scintigraphy results well enough to replace scintigraphy, or is a confirmatory scintigraphy required after an abnormal WMC test?	Suspected gastro-paresis	WMC alone	4 hour gastric scintigraphy or WMC plus gastric scinti-graphy	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic multicenter practice	Randomized controlled prospective trial, not device funded, blinded interpretation of results

Table D. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Clinical Questions (continued)								
5. Gold standard (colonic dysmotility/slow-transit constipation)	What is an appropriate test (gold standard) to compare with WMC in the diagnosis and monitoring of patients with slow-transit constipation? (e.g., clinical diagnosis, colonic scintigraphy, radiopaque markers, a combination of tests)	Suspected Slow transit constipation	WMC alone	Clinical history, GI question-naire, colonic scintigraphy, radiopaque markers, combinations of tests	Clinical outcomes – diagnostic accuracy, diagnostic gain with additional tests, safety, tolerability		Academic practice with scinti-graphy available – caveat that scinti-graphy is very expensive and this study may be prohibitively expensive	Randomized controlled prospective, blinded interpretation of results
6. Reproducibility of WMC results (gastroparesis)	In normal patients and those with gastroparesis, how reproducible are WMC studies performed 2 weeks apart?	Normal and suspected gastro-paresis	WMC	Repeat WMC at later time	Reproducibility of test results, change in specificity or sensitivity with two tests as compared with one		Any, preferably academic	Prospective, Blinded review, all patients receive same intervention.
7. Comparative diagnostic accuracy and safety (colonic dysmotility/slow-transit constipation)	Among patients with suspected slow-transit constipation how does the WMC compare with radiopaque markers and scintigraphy in accuracy in diagnosis and safety?	Suspected Slow transit constipation	WMC	Radiopaque markers, colonic scinti-graphy	Accuracy, Safety		Any	Prospective, blinded interpretation, followed at 1 and 30 days for harms
8. Incremental diagnostic accuracy (gastric dysmotility/gastroparesis)	What is the incremental value of WMC in addition to gastric scintigraphy, antroduodenal manometry or endoscopy to diagnose delayed gastric emptying or dysmotility?	Suspected gastric emptying delay	WMC	Gastric scinti-graphy, antro-duodenal manometry, endoscopy	Incremental value in diagnosis		Any	Blinded review of results, controlled trial, could examine cohorts

Table D. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
9. Best test or combination for predicting outcomes after colectomy (colonic dysmotility/slow-transit constipation)	Among patients with confirmed colonic dysmotility who undergo colectomy, is WMC data alone, or scintigraphy plus radiopaque markers, or the combination of all three the best predictor of whole gut dysmotility and clinical outcomes?	Confirmed colonic dysmotility patients who post test(s) have undergone or not undergone colectomy	WMC	Scintigraphy, radiopaque markers, combination of all three	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Any	Retrospective with prospectively collected data.
Methodological Questions								
1. WMC data in patients with gastric and colonic dysmotility	Among patients with gastric and colonic dysmotility, what are the ranges of WMC values for temperature, pH, pressure pattern and transit time in the stomach, intestines and colon?	Suspected gastric and suspected colonic dysmotility	WMC	Ranges of normal, no comparators needed	Temperature, pH, pressure patterns, transit time		Any	Prospective enrolled, blinded review, with confirmatory second reviewer to calculate kappa
2. WMC data in non-diseased populations	In non-diseased populations, what are the distributions of age-, race- and sex-specific values for pressure patterns/amplitude/frequency, temperature, pH, and transit time in the stomach, intestines and colon, as measured by the WMC?	Healthy patients, multiple ages, all sexes	WMC	Ranges of normal, comparators would be in patients of different ages and sexes	Temperature, ph , pressure patterns, transit time		Any	Prospective enrolled, blinded review

Table D. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Methodological Questions (continued)								
3. Rates of test failures and need for additional tests	How does the WMC compare with scintigraphy, manometry and radiopaque markers in the rates of test failures and need for additional tests?	Suspected gut dysmotility patients	WMC	Scintigraphy, manometry, radiopaque markers	Diagnostic accuracy, Test failure, need for additional tests		Any	Randomized, prospective, also could collect data prospectively and retrospectively analyze as ultimate outcome becomes available
4. WMC diagnostic accuracy and safety in gastroparesis subtypes	How do the diagnostic accuracy, safety and resource utilization of the WMC differ when used in ambulatory gastroenterology clinics and primary care offices - and when the WMC results are interpreted by primary gastroenterologists, primary care physicians or nurse practitioners?	Ambulatory general gastro-enterology patients and primary care physicians offices, suspected motility disorder with symptoms of gastric emptying delay or colonic transit abnormality	WMC	Clinical diagnosis	Diagnostic accuracy, safety, resource utilization		Ambulatory gastro-enterology practice and primary care offices	Prospective trial, observation in practice due to concerns over difficulties with randomizing at multiple local physicians offices.

Table D. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Methodological Questions (continued)								
5. Thresholds	What thresholds should be used to differentiate diagnostic accuracy between the WMC and another test? (e.g., sensitivity, specificity, % agreement, other)	Any popula-tions, healthy vs. diseased	WMC	Other testing modalities, what threshold differentiates effectively	Threshold optimal range for determining diagnostic accuracy		Any	Statistical analysis based on above data to determine optimal thresholds, would need data from other trials to calculate.
6. Correlation between WMC values and histopathology	Among patients with motility disorders, are WMC values correlated with the histopathological findings from full-thickness or standard biopsy?	Gastric or colonic dysmotility or general gut dysmotility patients	WMC	Histo-pathology samples or genetic samples from patients with known disease	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic center, with experience in genetics and tissue banking also neuro-gastro staining techniques	Prospectively collected registry information that can be correlated with ongoing genetic and advanced neurogastro-enterological pathology diagnostic techniques as they become available

ANMS = American Neurogastroenterology and Motility Society; GI = gastrointestinal; WMC = wireless motility capsule
[a]Timing field left blank if there were no salient timing issues.

Discussion

Our method of determining Future Research Needs has strengths and limitations. One issue was the narrow scope of our original evidence report, which focused on gastroparesis and constipation and the comparative role of WMC testing. Since our initial review did not try to analyze small bowel transit or whole gut transit abnormalities as part of the review process, some aspects of the potential benefit of WMC testing may not have been established by our work. Further, by focusing primarily on WMC testing, we may not have captured all of the needs for research on the evaluation of gastroparesis and constipation. However, after thorough analysis of the data, clear gaps were seen on the methodological side and on the clinical side of research on gastroparesis and slow-transit constipation with regard to WMC testing.

We used an abbreviated Delphi technique to determine the priorities of the stakeholders. In the first round, the stakeholders varied widely in the priorities they assigned to the various gaps. In the second round, the stakeholders still showed moderate variation in their priorities. This finding is not surprising, given that we intentionally recruited stakeholders having very different perspectives.

We included a limited number of stakeholders, and thus may not have a totally representative view of all relevant stakeholders. It would have taken much longer to collect information from more stakeholders, however, a larger group would have required review and approval by the Office of Management and Budget. We deliberately recruited individuals with different perspectives to gain insight from key stakeholder groups. The participating stakeholders included clinical experts, methodologic experts, and a patient/consumer advocate. Collectively, their thoughtful comments were indispensable in creating this FRN report.

The highest priority clinical and methodological gaps and questions are shown in Table E. To develop clinically accurate recommendations for providers and payers, we urgently need additional studies of these questions to help guide the development of formal evidence-based guidelines to replace the current consensus guidelines, which are based on limited evidence.

Conclusions

Evidence exists to support use of WMC testing for the detection, diagnosis, treatment and management of gastric and colonic dysmotility, but much of the evidence is either low strength or insufficient for making evidence-based recommendations. We focused on creating a compelling set of research questions, which was reviewed by a group of stakeholders and revised into a tiered list of priorities for future research. Future research to answer these questions will help to improve care by establishing a stronger evidence base for the use of WMC testing, potentially leading to improved diagnosis, detection, treatment and management of motility disorders, which have had few tools for accurate, reliable, portable, non-radiating, standardized diagnosis in the past. Although colonic and gastric dysmotility are not as common as high blood pressure or diabetes, the burden of disease for gastroparesis and severe colonic dysmotility is great. Accurate and rapid diagnosis on a consistent basis would be a cornerstone for conducting future research on the clinical outcomes of homogeneous groups of dysmotility patients.

Table E. Top-tier research priorities listed in order

Clinical Questions
Among patients with presumed gastroparesis, how would clinical outcomes differ between patients managed with WMC data alone versus those managed with scintigraphy data alone?
Among patients with suspected gastroparesis, how does the WMC compare with gastric scintigraphy, antroduodenal manometry, and endoscopy in its accuracy and safety in diagnosing gastroparesis?
What is an appropriate comparison test (gold standard) to which the WMC should be compared in the diagnosis and monitoring of patients with gastroparesis (e.g., clinical diagnosis, 4-hour gastric scintigraphy, antroduodenal manometry, endoscopy, plain X-ray of the abdomen 5 hours after swallowing a "dumb" radiopaque pill, a combination of tests)?
In the clinical evaluation of patients with suspected gastroparesis, do abnormal WMC gastric emptying times correlate with 4-hour gastric scintigraphy results well enough to replace scintigraphy, or is a confirmatory scintigraphy required after an abnormal WMC test?
What is an appropriate comparison test (gold standard) to which the WMC should be compared in the diagnosis and monitoring of patients with slow-transit constipation (e.g., clinical diagnosis, colonic scintigraphy, radiopaque markers, a combination of tests)?
In normal patients and those with gastroparesis, how reproducible are WMC studies performed 2 weeks apart?
Among patients with suspected slow-transit constipation how does the WMC compare with radiopaque markers and scintigraphy in its accuracy and safety of diagnosis?
What is the incremental value of using the WMC in addition to gastric scintigraphy, antroduodenal manometry or endoscopy to diagnose delayed gastric emptying or dysmotility?
Among patients with confirmed colonic dysmotility who undergo colectomy, is WMC data alone, or scintigraphy plus radiopaque markers, or the combination of all three the best predictor of whole gut dysmotility and clinical outcomes?
Methodological Questions
Among patients with gastric and colonic dysmotility, what are the ranges of WMC values for temperature, pH, pressure pattern and transit time in the stomach, intestines and colon?
In non-diseased populations, what are the distributions of age-, race- and sex-specific values for pressure patterns/amplitude/frequency, temperature, pH, and transit time in the stomach, intestines and colon, as measured by the WMC?
How does the WMC compare with scintigraphy, manometry, and radiopaque markers in the rates of test failures and need for additional tests?
How do the diagnostic accuracy, safety, and resource utilization of the WMC differ when used in ambulatory gastroenterology clinics and primary care offices - and when the WMC results are interpreted by primary gastroenterologists, primary care physicians, or nurse practitioners?
What thresholds should be used to differentiate diagnostic accuracy between the WMC and another test (e.g., sensitivity, specificity, % agreement, other)?
Among patients with motility disorders, are WMC values correlated with the histopathological findings from full-thickness or standard biopsy?

WMC = wireless motility capsule

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. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.

Background

Context

Gastroparesis

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] Detection of gastric emptying delay is the essence of diagnosing gastroparesis. The prevalence of gastroparesis was estimated by a community-based study in 2007 to be 9.6 per 100,000 for men and 37.8 per 100,000 for women.[2] Hospitalizations for gastroparesis increased by 158 percent between 1995 and 2004.[3] Standard assessment for patients with typical symptoms begins with exclusion of mechanical causes of disease. Methods of testing include gastric emptying scintigraphy, antroduodenal manometry, and now wireless motility capsule (WMC) technology. Documentation of gastric emptying delay guides physicians in their recommendations for nutrition, medication, and surgical therapies.

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 such as changes in medications or nutrition, or to affect patient-centered outcomes such as symptom improvement, need for surgery, quality of life, and 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

Constipation is common, occurring in 15 to 20 percent of the U.S. population.[4-6] It is defined as fewer than two bowel movements per week or a decrease in a person's normal frequency of stools that is accompanied by straining, difficulty passing stool, or passage of hard solid stools.[4] Patients with symptoms of constipation must be assessed by their medical history and a physical examination to exclude malignant or organic causes of constipation. For individuals who are less than 50 years of age without "red flag" symptoms, no testing is required to make a diagnosis of constipation if they meet the Rome III criteria. 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. Reported incidence of slow-transit constipation is 1 in 3000. Other studies list an incidence of 0.17 percent.[7] The true incidence is likely unknown.[4] Lifestyle changes and medical management should be used for all patients with symptoms of constipation. Thus, the initial evaluation of constipation symptoms does not often involve colonic transit testing. For certain individuals with suspected slow-transit constipation, colon transit testing can provide insight into the etiology of the constipation. The main diagnostic methods used to test for colonic motility are radiopaque marker examination, colonic scintigraphy, colonic and anorectal manometry, and WMC testing.[8,9] The reference standard has been radiopaque markers.

Most patients with chronic constipation have improvement of symptoms with medical therapy and/or lifestyle changes. If testing confirms the presence of slow-transit constipation (colonic inertia) without use of laxatives, then surgery could be considered as a potential

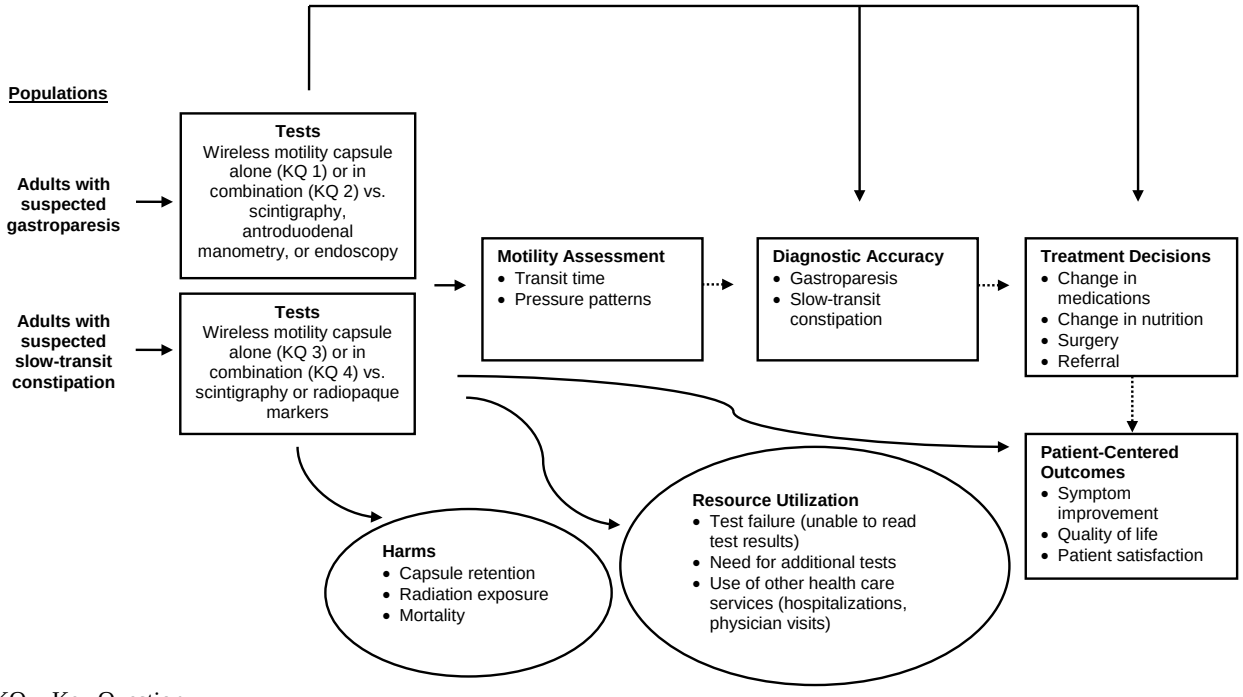
therapy.[10] Most clinicians reserve colectomy for patients with the most terminal or untreatable conditions.

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 such as change in medications or change in nutrition or to affect patient-centered outcomes such as symptom improvement, need for surgery, quality of life, and 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.

Objectives of the Original Systematic Review

The WMC is a new modality for diagnosing gastric and colonic motility disorders. The Johns Hopkins Evidence-based Practice Center (EPC) recently completed an Agency for Healthcare Research and Quality (AHRQ) funded systematic review of comparative effectiveness of WMC as compared with other tests for diagnosing and managing gastric and colonic motility disorders, instead of or in conjunction with other testing modalities (see analytic framework Figure 1). We also sought to define the populations that would benefit most from motility testing, including WMC testing. Overall, we showed that there was low or insufficient strength of evidence to answer most of the parts of the clinical questions we posed. A summary of the Key Questions from the review and the evidence is listed in Table 1.

Figure 1. Analytic framework of the comparative effectiveness of diagnostic technologies for evaluating gastroparesis and constipation



KQ = Key Question

Table 1. Summary of the results from the systematic review on the wireless motility capsule

Topic and Key Questions	Key Findings	Strength of Evidence
Comparative diagnostic accuracy of WMC for gastroparesis. **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, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	Seven studies evaluated diagnosis of gastric emptying delay. We found low strength of evidence that WMC alone was comparable to scintigraphy for diagnostic accuracy, motility assessment, treatment decisions, and resource utilization.	**Low**
Incremental diagnostic accuracy for WMC in combination with other diagnostic methods for gastroparesis. **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 the WMC in terms of diagnostic accuracy of gastric emptying delay, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	We found two studies evaluating WMC as an add-on to other testing. The strength of evidence was low for diagnostic accuracy and motility assessment of WMC in combination with other modalities. The addition of WMC increased diagnostic yield.	**Low**
Comparative diagnostic accuracy of WMC for slow-transit constipation. **KQ 3:** In the evaluation of colonic dysmotility, how does the WMC alone compare with radiopaque markers and scintigraphy in terms of diagnostic accuracy of slow-transit constipation, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	Nine studies analyzed colon transit disorders and provided moderate strength of evidence for diagnostic accuracy, motility assessment, and harms. WMC was comparable to radiopaque markers. Few harms were reported. The evidence was insufficient to draw conclusions about effects of WMC on treatment decisions and resource utilization.	**Low**
Incremental diagnostic accuracy for WMC in combination with other diagnostic methods for slow-transit constipation. **KQ 4:** When a radiopaque marker or scintigraphy is used in the evaluation of colonic dysmotility, what is the incremental value of also using the WMC in terms of diagnostic accuracy of slow-transit constipation, motility assessment, treatment decisions, patient-centered outcomes, harms, and resource utilization?	No studies directly assessed use of WMC in combination with other tests to detect colon transit delay.	**Insufficient**

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

Conclusions of the Systematic Review

WMC is similar to current modalities in use for detection of slow-transit constipation and gastric emptying delay, and therefore is another viable diagnostic modality. Little data is available to determine the optimal timing of WMC in diagnostic algorithms.

Evidence Gaps in the Systematic Review

Overall, the evidence was graded as low to address the ability of WMC to detect gastroparesis or slow-transit constipation. The low strength of evidence was due to the limitations in the literature identified in the systematic review. The main limitations resulted from inconsistency in reporting on motility testing modalities. Great variation existed between methods of administration of diagnostic tests, and also in assessment of those tests. No unified

standards existed to determine improvement in diagnostic accuracy. In fact, radiopaque marker testing is considered a non-reference standard, not consistent enough to be considered a gold or reference standard for assessment of slow-transit constipation. We arbitrarily chose a 10 percent difference in sensitivity or specificity for reference standards or device agreement/concordance for non-reference standards, such as radiopaque markers (ROM). Most of the "normal" subjects upon which the studies were validated were young college aged men, and most of the affected population under investigation were 50-something women.

We had excluded studies that included non-diseased participants exclusively, since our population of focus was subjects with suspected gastroparesis or constipation. Many of the studies in the literature did report on non-diseased participants, and were thus excluded. The major strength of our review was its comprehensiveness.

Little data is available to support the timing of WMC in the diagnostic and therapeutic approach to patients with symptoms of possible gastroparesis or slow transit constipation. Further work needs to be done to classify the types of patients within subgroups of gastroparesis or slow-transit constipation to identify severe cases that may need more urgent evaluation. Finally, little is known about whether testing should be used 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.

Our aim was to compare the diagnostic accuracy of the WMC to other testing modalities to diagnose and manage gastroparesis and slow transit constipation. The identified literature limited our ability to answer our Key Questions for several reasons. We comprehensively reviewed the literature in a systematic fashion to accomplish this goal. However, we excluded studies that included non-diseased participants exclusively as 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. Thus, we excluded a number of studies that evaluated characteristics of the WMC. Other limitations included the fact that few studies prospectively addressed the goal of tabulating the incremental value of WMC in addition to other modalities for diagnosis of gastroparesis or constipation. No studies appeared to definitively identify non-inferiority or superiority of a diagnostic evaluation with WMC instead of other modalities. Most of the identified studies were from major academic referral centers, which may have led to spectrum bias. The sensitivity and specificity of the 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. We were unable to compare the results of studies with and without industry or investigator conflicts of interest because most studies were sponsored by the company that manufactures the WMC. 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.

Methods

The aim of the Future Research Needs (FRN) project was to develop a prioritized list (or multiple lists) of research needs within the scope of the original systematic review, with considerations for potential research designs with sufficient detail for researchers and funders to use for developing research proposals or solicitations. As the resulting research is meant to improve healthcare decisions, stakeholders included patients/advocates, clinicians, and third-party payers.

The research needs were based on the research gaps identified in the systematic review writing process and contributed by the stakeholders. The methods for identifying evidence gaps and developing them into a prioritized list of research needs and feasible researchable questions involved the steps described in the immediately following subsections.

Identification of Evidence Gaps

A subset of seven of the original systematic review authors constituted the EPC FRN team. Evidence gaps were identified in the systematic review writing process based on the strength of evidence, applicability, and limitations of the systematic review. Evidence gaps were defined as parts of the systematic review Key Questions that had low strength of evidence or insufficient evidence. The FRN team met multiple times and circulated by email lists of potential questions to identify gaps with specific reference to study design and the PICOTS (lack of information or insufficient evidence for: sub-populations/whole populations; interventions, comparisons of interventions to each other; outcomes; timing of interventions or comparisons of interventions; and settings). The FRN team used this process to develop a list of research gaps to be presented to a stakeholder panel for review, as described in the following subsections.

Engagement of Stakeholders, Researchers, and Funders

Stakeholder Identification

We recruited a group of nine stakeholders to participate in the identification and prioritization of research gaps. We sought input from patients/advocates, clinical experts, and payers. Five stakeholders were chosen from the systematic review Key Informants and Technical Expert Panel members. These were chosen because of their expertise, familiarity with the systematic review, and because they had been particularly responsive and helpful with the systematic review process. In addition, four new participants were chosen to fill out the nine member stakeholder panel. Some of these were suggested by the systematic review investigators. We also searched websites of advocacy organizations to identify patient advocates who appear to be independent of payers and manufacturers according to the voting membership requirements and funding mechanisms of their organizations. The list was summarized in a table of their individual strengths and the list was presented to stakeholders who accepted an email invitation explaining the project and inviting them to participate signed a Conflict of Interest (COI) form declaring professional activities and financial ties relevant to the clinical area. It was made clear to them that accepting the invitation and returning the COI form constituted agreement to be identified as a stakeholder contributor to the final document. Manufacturers were not solicited to be part of the stakeholder panel, but they were informed of their ability to comment during the 4 week public posting period for this report.

Orienting Stakeholders

By emailed letters we provided the stakeholders with a description of the FRN Project, and how it is derived from and relates to the systematic review. We also sent them the draft of the executive summary of the systematic review. A Web link to the complete draft report was provided noting that the draft is only temporarily available, and that reading the executive summary should be sufficient to meaningfully contribute to the process to identify evidence gaps and FRN.

Stakeholder Engagement for Additional Gap Identification and Prioritization

We used an approach performed in two rounds of engagement with the stakeholders by means of emailed questionnaires.

Engagement Round 1: Gap List From Systematic Review and Preliminary Prioritization

The FRN team's list of research gaps, derived from the systematic review as described above, was presented to the stakeholders by email for review and for suggestions of additional gaps within the scope of the systematic review. They were instructed to carry out a preliminary prioritization of the gaps, including any additional gaps they added to the list. To perform this preliminary prioritization they were asked to use the criteria and ranking method described in the next subsection.

Criteria for Prioritization

Prior to engaging our stakeholders, we developed a draft framework consisting of criteria we considered important for prioritizing topics for future research. These draft criteria were adapted from AHRQ's Effective Health Care Program Topic Selection Criteria and included:

1. **Importance.** The importance of the condition to patients (including consideration of whether that gap is of particular relevance to priority subpopulations such as pediatric patients, elderly patients, vulnerable and disparity populations)
2. **Impact.** The extent to which new research with definitive findings could potentially impact decision-making by patients, providers, or policy makers

Other prioritization criteria of AHRQ/EHC were determined to be less useful or relevant for future research need prioritization.

Uncertainty is not a useful criterion, because all identified evidence gaps are uncertain by definition.

Feasibility of research on an evidence gap is a secondary concern that is independent of the need for evidence. Evidence gaps and the need for research to close such gaps are innate to the area of interest. They are gaps and needs regardless of whether research is possible or feasible. There is value in determining the *absolute* importance and potential impact of closing each of the gaps. The feasibility of carrying out the required research to close a gap is *relative*, depending on the difficulty of the research, funding sources, availability of adequate numbers of patients, incentives for researchers, the convenience of the needed length of followup, attractiveness of the research to researchers, etc. A question of the highest priority may secondarily be deemed worth pursuing in spite of the difficulty and cost of the research; whereas, similar difficulty and

cost may render research on a lesser priority gap "unfeasible." A funding source or research group with substantial resources may consider a research question feasible, while those with limited resources may not – circumstances beyond our knowledge or control. Therefore, we did not want relative feasibility to enter into the absolute priority decisions of the stakeholders. Nevertheless, we asked them to comment on feasibility, and we discuss it as a secondary aspect of the FRN questions.

We did not attempt to give our two criteria quantitative values or to break these major criteria into their multiple factors for individual weighting or priority ranking and combination by a mathematical formula. That would require validation of the weights. We did not consider that practical within the scope of this project. Summing multiple factors by an arbitrary mathematical formula would give an undue appearance of objectivity, accuracy and precision. Instead we instructed the stakeholders to consider these two major criteria, *importance* and *impact*, in their priority decisions.

In the round 1 questionnaire each stakeholder was presented with two lists of gaps, one for clinical questions from the systematic review, and another for methodological questions. The stakeholders were asked to choose their top five in each list and rank them by priority from 1 (highest) to 5, based on the criteria described above. Stakeholders were also asked to suggest additional gaps within the scope of the systematic review, if they were aware of any ongoing studies addressing a gap, and to comment on feasibility of research to address the gap.

In order to sort the gaps by priority, the ratings of 1 (highest) to 5 were inverted, so a priority rating of 1 corresponded to a point value of 5 for highest priority. Then these inverted individual stakeholder scores for each gap were summed and sorted from highest sum (highest priority) to lowest. If multiple gaps achieved the same sum, they were given the same priority rank. This priority sorted list was considered the preliminary priority ranked list of gaps.

Engagement Round 2: Final Prioritization

The FRN team incorporated the stakeholder comments and additional suggestions from engagement round 1 into a list of gaps for final prioritization. This list included the preliminary ranking from the previous round. Again this was two lists, one for clinical questions, and another for methodological questions. Each stakeholder was presented with these lists by email and asked to choose their top 6 choices in the clinical list and prioritize them from 1 (highest) to 6. And they were asked to choose their top seven from the methodology list and rate them from 1 (highest) to 7. We again asked them to base their ratings on the same criteria as in round 1, *importance*, and *potential impact*. Stakeholders were again asked if they are aware of any ongoing studies addressing the gaps (duplication), and they were able to comment on the feasibility of research addressing the gaps.

Top-Tier Future Research Needs

A global priority ranking of the gaps was calculated from the stakeholder individual ratings as described above. Appendix C shows the inverted scores, sums, and priority ranks. The global ranking was inspected by the EPC team to determine if there was an obvious cutpoint between a top tier of questions and the remainder. If the global ranking was a continuum with no apparent cutpoint, the top half of the gaps or the top 10, whichever is fewer, were to be chosen as the top tier and considered the high-priority FRN.

We also took into account the research needs that were prioritized highly by each stakeholder. After determining a preliminary top-tier cutpoint (Appendix C) according to the

score sums as described, we reviewed and analyzed the individual stakeholder responses for the questions on either side of the cutpoint. We then assessed which questions had received top votes and next-to-top votes from each stakeholder, and we tracked how many of these number 1 and number 2 priorities each question near the cutpoint received. Using this method we verified that the sum score cutpoint we chose reflected the high-priority items that the stakeholders sought to identify (Appendix C). If a stakeholder scored an item highest priority it was placed on the high-priority list automatically. There were 9 designated top priority items in the Clinical Questions list, and 6 in the Methodology Questions list, for a total of 15.

Research Needs Development and Research Design Considerations

We developed questions addressing the FRN (top tier of knowledge gaps), including PICOTS information. These were circulated to the entire group by email, and discussed and further developed in multiple meetings of the FRN team. These FRN were presented to the stakeholders along with the corresponding evidence gaps in the second round of prioritization.

When we solicited information from stakeholders, we asked them to assess factors such as resource use, ethical considerations, data availability, recruitment or feasibility issues, and validity. We also offered them space to comment on any specific study designs relevant to a particular question. To complete the FRN report, we listed and assessed one to three study design suggestions for each of the final FRN, with limited textual description as suggested by the AHRQ guidance.[11] We considered information from comments on each question that were provided to us by stakeholders via the initial questionnaire, and many of these were incorporated in the study design section.

Approach to External Literature Searches

To identify ongoing clinical trials that may address our FRN questions, we searched ClinicalTrials.gov, NIH Reporter, the Canadian Institutes of Health Research, the World Health Organization Clinical Trials Registry, and the European Union Clinical Trials Register since the search cutoff date of the systematic review. Our search terms are presented in Appendix A. Each article was reviewed by two people for inclusion, applying the same inclusion/exclusion criteria used in the comparative effectiveness systematic review. For each included trial, we abstracted the trial identification number, date of registry, the expected date of completion, the study name, status, medications compared, any published results, and determined the Key Question the study is likely to address (Appendix B).

Analytic Framework

We used an analytic framework to describe research gaps using the same format as for the systematic review (Figure 1).

Results

Stakeholders

The stakeholders broadly represented clinical, research and advocacy perspectives on gastroparesis and constipation. The panel comprised nine stakeholders, including a patient advocate and non-motility specialist statistician. We generated lists of potential stakeholders via a review of Technical Expert Panel members from the systematic review, and added several new members to our team of stakeholders who did not have any corporate interest in the development of WMC. Six of nine stakeholders are considered clinical experts in gastrointestinal motility and neurogastroenterological disorders, one is a general expert in gastroenterology who designs clinical trials, and the remaining two members offered a public perspective and research methodology background which were very useful.

Knowledge Gaps

Using the information from the Johns Hopkins University EPC draft systematic review, Wireless Motility Capsule versus Other Diagnostic Technologies for Evaluating Gastroparesis and Constipation, a series of key areas of unanswered questions or gaps in knowledge was identified by the FRN team from the areas with low or insufficient strength of evidence noted in the systematic review. We sought to represent all these evidence gaps with the design of a series of questions for prioritization.

A lack of basic and fundamental knowledge about WMC measurements for both normal patients and those with suspected or real motility disorders, as well as a lack of basic and fundamental knowledge about the comparison tests of scintigraphy, manometry, and ROMs affected the strength of evidence. For example, what is the range of values for measurements with the WMC in both normal patients, those suspected of motility disorders, and those with actual motility disorders? Many of these values have yet to be firmly established, and thus the FRN team felt that some of these foundational natural history, epidemiology and methodological gaps would help lay the framework needed for more systematic research and discovery into the clinical, topical questions. Therefore, in analyzing the evidence gaps, the FRN team identified two types of gaps: (1) clinical topical gaps and (2) foundational epidemiology gaps, which we refer to as methodological gaps. Because of the distinct nature of these two types of gaps, and consistent with AHRQ guidance, (www.effectivehealthcare.ahrq.gov/ehc/products/378/1039/MFRN9_PresentationofFutureResearchNeeds_FinalReport_20120418.pdf) these two types of gaps were prioritized separately by the stakeholder panel. See Tables 2 and 3 for complete lists of clinical and methodological gaps.

The gap topics were reviewed by the stakeholder panel in two lists to allow them to comment, suggest and rank important issues for both the directly important clinical issues for WMC and the methodological questions. We received suggestions for additional research questions, which were solicited at the first round of survey, and then subject to general review at the second round of questions. Our stakeholder panel readily identified these areas as necessary and important during both rounds of feedback and specifically they focused on certain questions of greater importance and priority.

The highest ranked clinical evidence gap reflected the uncertainty in the role of WMC in outcomes for patients with presumed gastroparesis and the role of WMC as a replacement test versus adjunct test for diagnosis. Consensus guidelines suggest that WMC is a replacement test

for current testing methods, however additional research would lend more weight to that argument. It is currently very difficult in active practice to get access to the WMC for some patients due to lack of insurance coverage. But further evidence of benefit may facilitate wider access to the WMC. This is clearly a high priority. The other top-ranked question asked whether these same patients with suspected gastroparesis would have comparable results from scintigraphy, antroduodenal manometry, plain x ray after marker blind capsule ingestion, endoscopy, or combination of tests in diagnosis of gastroparesis. Similarly the next ranked gaps addressed the question of which test should serve as the gold standard or reference standard for comparison, and whether a confirmatory test is required after use of WMC. Reproducibility of the WMC was also thought to be important and ties into the previous questions which tried to establish the role of WMC in testing. Beyond these knowledge gaps on gastroparesis the next set of questions addressed slow-transit constipation, including establishing the role of WMC in diagnosis compared with other standard tests for ability to diagnose, as well as accuracy and safety in diagnosis.

The next series of gaps focused on the incremental value of WMC in addition to other testing methods for gastric emptying delay and slow-transit constipation. Interestingly, there were also ranked questions in the top tier asking about correlating pathology with clinical history and WMC findings. This echoes the questions from the methodological side asking about the same correlation. Also a focus was the role of WMC in colonic dysmotility to predict outcomes, or to predict the effects medical and/or surgical therapy on outcomes.

The most highly ranked methodological gaps (Table 3) made clear the basic data ranges for WMC in patients have yet to be firmly established in non-diseased populations and those with suspected gastric or colonic dysmotility. With better established norms for diseased and non-diseased patients there can be better framing of future questions and research endeavors. The next most highly ranked gaps were very similar to the initial research questions asked by the original systematic review, which had only low strength of evidence. With additional quality research, the panel thought that the strengths and weaknesses of WMC would be made more apparent. Thus, they suggested a priority area for research would be future studies that focus on establishing the role of WMC comparatively with scintigraphy, manometry, and ROMs as well as to find out information about test failure and need for additional tests. It was also considered important to assess diagnostic accuracy of the test when used by nonacademic specialists or as a front-line test. Other priority items included establishing the thresholds of diagnostic accuracy and establishing the basic science connection between WMC results and histopathological findings from patients with known disease, if one exists. All of these basic foundational or methodological questions were ranked highly by participants in both rounds. These are by definition high-priority areas for future research.

Table 2. Wireless motility capsule stakeholder prioritized clinical knowledge gaps

Clinical Knowledge Gaps	Priority Rank[1] (1 = Highest Priority)
Top Tier	
1. Comparative effectiveness of managing with WMC (gastric dysmotility/gastroparesis)	1
2. Comparative diagnostic accuracy and safety (gastric dysmotility/gastroparesis)	2
3. Gold standard (gastric dysmotility/gastroparesis)	3
4. Comparison of WMC and scintigraphy results (gastroparesis)	4
5. Gold standard (colonic dysmotility/slow-transit constipation)	5
6. Reproducibility of WMC results (gastroparesis)	5
7. Comparative diagnostic accuracy and safety (colonic dysmotility/slow-transit constipation)	6
8. Incremental diagnostic accuracy (gastric dysmotility/gastroparesis)	7
9. Incremental diagnostic accuracy (colonic dysmotility/slow-transit constipation)	8
Other Gaps	
10.Distinction of subtypes of gastroparesis (gastric dysmotility/gastroparesis)	9
11. Best test or combination for predicting outcomes after colectomy (colonic dysmotility/slow-transit constipation)	9
12. Comparative accuracy in predicting response to treatment (colonic dysmotility/slow-transit constipation)	9
13. Utility of colonic pressure (colonic dysmotility)	10
14. Differentiating IBS-C from idiopathic constipation (colonic dysmotility/slow-transit constipation)	10
15. Comparisons of WMC pressure profiles (gastroparesis)	10
16. Best test or combination for predicting outcomes without colectomy (colonic dysmotility/slow-transit constipation)	11
17. Comparative accuracy in predicting response to treatment (gastric dysmotility/gastroparesis)	12
18. Comparative value in monitoring response to treatment (gastric dysmotility/gastroparesis)	13
19. Comparative value in monitoring response to treatment (colonic dysmotility/slow-transit constipation)	13
20. Colonic pressure patterns after waking (colonic dysmotility)	13
21. Evaluation of high-amplitude contractions (colonic dysmotility)	13

C = clinical knowledge gap; WMC = wireless motility capsule

It was possible for multiple gaps to have the same priority rank if the summation scores for computing the ranks were the same (Appendix C).

Table 3. Wireless motility capsule stakeholder prioritized methodological knowledge gaps

Methodological Knowledge Gaps	Priority Rank (1 = Highest Priority)
Top Tier	
1. WMC data in patients with gastric and colonic dysmotility	1
2. WMC data in nondiseased populations	2
3. Rates of test failures and need for additional tests	3
4. WMC diagnostic accuracy and safety in gastroparesis subtypes	4
5. Thresholds	5
6. Correlation between WMC values and histopathology	6
Other Gaps	
7. Colonic pressure patterns following meal ingestion.	7
8. Relation of baseline WMC values with long-term followup values	8
9. WMC data compared with other tests in non-diseased populations	9
10. Concurrent tests	10
11. Masking interpreters of WMC data	11
12. Use of historical controls in WMC research	12

M = methodological knowledge gap; WMC = wireless motility capsule

Future Research Needs and Study Design Considerations

We next developed a list of research questions based on the FRN, (i.e., top tier of prioritized evidence gaps) with sufficient detail for use by researchers and funders (i.e., research questions, PICOTS information and considerations of pros and cons of various research designs). Our FRN focused on broader topics for review by stakeholders, which are summarized in Table 4. Our FRN team has recommended study design characteristics based on the previous systematic review and feedback from the stakeholders.

Table 4. Summary of research needs (top tier)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Clinical Questions								
1. Comparative effectiveness of managing with WMC (gastric dysmotility/ gastroparesis)	Among patients with presumed gastroparesis, how do clinical outcomes differ between patients managed with WMC data alone versus those managed with scintigraphy data alone?	Presumed gastro-paresis	WMC alone	4 hour gastric scintigraphy by standard protocol (ANMS consensus)	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization	Simul-taneous	Academic, multi-center practice	Randomized, controlled trial, prospective
2. Comparative diagnostic accuracy and safety (gastric dysmotility/ gastroparesis)	Among patients with suspected gastroparesis, how does the WMC compare with gastric scintigraphy, antroduodenal manometry, and endoscopy in its accuracy and safety in diagnosing gastroparesis?	Suspected gastro-paresis	WMC alone	4 hour gastric scintigraphy by standard protocol (ANMS consensus), or antro-duodenal manometry and/or endoscopy	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic, multi-center practice	Randomized, controlled trial, prospective

Table 4. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Clinical Questions (continued)								
3. Gold standard (gastric dysmotility/gastro-paresis)	What is an appropriate test (gold standard) to compare with WMC in the diagnosis and monitoring of patients with gastroparesis? (e.g., clinical diagnosis, 4-hour gastric scintigraphy, antroduodenal manometry, endoscopy, plain X-ray of the abdomen 5 hours after swallowing a "dumb" radiopaque pill, a combination of tests)	Suspected or known gastro-paresis subjects	WMC alone	Clinical history, standardized GI question-naire, 4 hour gastric emptying, antro-duodenal manometry, endoscopy, plain x ray 5 hours after swallowing a patency capsule, or some combination above	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic, multicenter practice	Randomized controlled prospective trial, or prospectively enrolled but retrospectively reviewed cases, blinded interpretation of results
4. Comparison of WMC and scintigraphy results (gastroparesis)	In the clinical evaluation of patients with suspected gastroparesis, do abnormal WMC gastric emptying times correlate with 4-hour gastric scintigraphy results well enough to replace scintigraphy, or is a confirmatory scintigraphy required after an abnormal WMC test?	Suspected gastro-paresis	WMC alone	4 hour gastric scintigraphy or WMC plus gastric scinti-graphy	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic multicenter practice	Randomized controlled prospective trial, not device funded, blinded interpretation of results

Table 4. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Clinical Questions (continued)								
5. Gold standard (colonic dysmotility/slow-transit constipation)	What is an appropriate test (gold standard) to compare with WMC in the diagnosis and monitoring of patients with slow-transit constipation? (e.g., clinical diagnosis, colonic scintigraphy, radiopaque markers, a combination of tests)	Suspected Slow transit constipation	WMC alone	Clinical history, GI question-naire, colonic scintigraphy, radiopaque markers, combinations of tests	Clinical outcomes – diagnostic accuracy, diagnostic gain with additional tests, safety, tolerability		Academic practice with scinti-graphy available – caveat that scinti-graphy is very expensive and this study may be prohibitively expensive	Randomized controlled prospective, blinded interpretation of results
6. Reproducibility of WMC results (gastroparesis)	In normal patients and those with gastroparesis, how reproducible are WMC studies performed 2 weeks apart?	Normal and suspected gastro-paresis	WMC	Repeat WMC at later time	Reproducibility of test results, change in specificity or sensitivity with two tests as compared with one		Any, preferably academic	Prospective, Blinded review, all patients receive same intervention.
7. Comparative diagnostic accuracy and safety (colonic dysmotility/slow-transit constipation)	Among patients with suspected slow-transit constipation how does the WMC compare with radiopaque markers and scintigraphy in accuracy in diagnosis and safety?	Suspected Slow transit constipation	WMC	Radiopaque markers, colonic scinti-graphy	Accuracy, Safety		Any	Prospective, blinded interpretation, followed at 1 and 30 days for harms

Table 4. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Clinical Questions (continued)								
8. Incremental diagnostic accuracy (gastric dysmotility/ gastroparesis)	What is the incremental value of WMC in addition to gastric scintigraphy, antroduodenal manometry or endoscopy to diagnose delayed gastric emptying or dysmotility?	Suspected gastric emptying delay	WMC	Gastric scinti-graphy, antro-duodenal manometry, endoscopy	Incremental value in diagnosis		Any	Blinded review of results, controlled trial, could examine cohorts
9. Best test or combination for predicting outcomes after colectomy (colonic dysmotility/ slow-transit constipation)	Among patients with confirmed colonic dysmotility who undergo colectomy, is WMC data alone, or scintigraphy plus radiopaque markers, or the combination of all three the best predictor of whole gut dysmotility and clinical outcomes?	Confirmed colonic dysmotility patients who post test(s) have undergone or not undergone colectomy	WMC	Scinti-graphy, radiopaque markers, combination of all three	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Any	Retrospective with prospectively collected data.
Methodological Questions								
1. WMC data in patients with gastric and colonic dysmotility	Among patients with gastric and colonic dysmotility, what are the ranges of WMC values for temperature, pH, pressure pattern and transit time in the stomach, intestines and colon?	Suspected gastric and suspected colonic dysmotility	WMC	Ranges of normal, no comparators needed	Temperature, pH, pressure patterns, transit time		Any	Prospective enrolled, blinded review, with confirmatory second reviewer to calculate kappa

Table 4. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Methodological Questions (continued)								
2. WMC data in non-diseased populations	In non-diseased populations, what are the distributions of age-, race- and sex-specific values for pressure patterns/amplitude/frequency, temperature, pH, and transit time in the stomach, intestines and colon, as measured by the WMC?	Healthy patients, multiple ages, all sexes	WMC	Ranges of normal, comparators would be in patients of different ages and sexes	Temperature, ph , pressure patterns, transit time		Any	Prospective enrolled, blinded review
3. Rates of test failures and need for additional tests	How does the WMC compare with scintigraphy, manometry and radiopaque markers in the rates of test failures and need for additional tests?	Suspected gut dysmotility patients	WMC	Scintigraphy, manometry, radiopaque markers	Diagnostic accuracy, Test failure, need for additional tests		Any	Randomized, prospective, also could collect data prospectively and retrospectively analyze as ultimate outcome becomes available

Table 4. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Methodological Questions (continued)								
4. WMC diagnostic accuracy and safety in gastroparesis subtypes	How do the diagnostic accuracy, safety and resource utilization of the WMC differ when used in ambulatory gastroenterology clinics and primary care offices - and when the WMC results are interpreted by primary gastroenterologists, primary care physicians or nurse practitioners?	Ambulatory general gastro-enterology patients and primary care physicians offices, suspected motility disorder with symptoms of gastric emptying delay or colonic transit abnormality	WMC	Clinical diagnosis	Diagnostic accuracy, safety, resource utilization		Ambulatory gastro-enterology practice and primary care offices	Prospective trial, observation in practice due to concerns over difficulties with randomizing at multiple local physicians offices.
5. Thresholds	What thresholds should be used to differentiate diagnostic accuracy between the WMC and another test? (e.g., sensitivity, specificity, % agreement, other)	Any popula-tions, healthy vs. diseased	WMC	Other testing modalities, what threshold differentiates effectively	Threshold optimal range for determining diagnostic accuracy		Any	Statistical analysis based on above data to determine optimal thresholds, would need data from other trials to calculate.

Table 4. Summary of research needs (top tier) (continued)

Gap	Question	Population	Intervention	Comparator	Outcomes	Timing[a]	Setting	Potential Study Designs
Methodological Questions (continued)								
6. Correlation between WMC values and histopathology	Among patients with motility disorders, are WMC values correlated with the histopathological findings from full-thickness or standard biopsy?	Gastric or colonic dysmotility or general gut dysmotility patients	WMC	Histo-pathology samples or genetic samples from patients with known disease	Clinical outcomes – diagnostic accuracy, harms, motility assessment, tolerability, treatment decisions, patient centered outcomes – symptom improvement, quality of life resource utilization		Academic center, with experience in genetics and tissue banking also neuro-gastro staining techniques	Prospectively collected registry information that can be correlated with ongoing genetic and advanced neurogastro-enterological pathology diagnostic techniques as they become available

ANMS = American Neurogastroenterology and Motility Society; GI = gastrointestinal; WMC = wireless motility capsule
[a]Timing field left blank if there were no salient timing issues.

Study Design Considerations

Clinical research questions (Table 4) were framed to be relevant to clinical providers. The stakeholders offered general guidance towards better study design. Many stakeholders remarked that prospective trials would be ideal to help resolve some of the methodological issues. Other stakeholders wanted simultaneous assessment of controls and suspected cases within the same study undergoing analysis, which is standard diagnostic test methodology. There were some concerns about how expensive or non-feasible a particular project may be. This was true for all the longitudinal projects suggested within the questions. There were significant concerns expressed about using colonic scintigraphy, because as a modality it is only available in two centers and it is very difficult to complete as well as expensive and taxing on the patients. There were also concerns about the expense of tandem test study designs, where multiple testing modalities were to be performed on the same patient. Few stakeholders suggested that further comparative effectiveness trials were the only area of research need, since there were so many methodological issues which also need attention.

From the original systematic review, several areas of focus on methodological quality were apparent, and the comments of the stakeholders for the FRN project echoed these same concerns. Testing modalities should have standardized protocols so that they can be compared between studies (Table 5). For instance, prior to WMC, proton pump inhibitors should ideally be held for 7 days. The test should be performed using the WMC ingested at appropriate times as per manufacturer's instructions. For gastric emptying study performance to be optimal and reproducible, the patients should have controlled blood sugar at the time of testing, should not have smoked cigarettes within 24 hours, and should be off all prokinetics for 7 days prior to testing. Studies that are performing these tests as part of clinical research should collect data about whether or not testing was carried out to the conventional standards recommended by the American Neurogastroenterology and Motility Society. In an ideal world, all tested motility patients should have been weaned off all narcotics prior to any motility testing, as narcotics directly influence bowel transit. Some of these requirements are difficult to achieve in real life patients, many of whom have longstanding poor diabetes control or narcotic dependence. But in the ideal research setting close attention should be paid to these details.

Table 5. Test Protocols for the diagnosis of gastroparesis and slow-transit constipation

Test	Duration of Test	Timing of Meal or X Ray	Prespecified Criteria for Inclusion
Gastric scintigraphy	4 hour testing	Prespecified meal with radiolabeled egg	Off tobacco, off narcotics, off prokinetics, glucose controlled below 250, off antidepressants or stable for 6 months
WMC	As per protocol	Prespecified timing of meal Smartbar vs. Eggbeaters	Off tobacco, Off prokinetics, Off proton pump inhibitor,
Colon transit testing	As per protocol, prespecify which protocol and timing of ingestion	Prespecify x-ray timing to document stool passage, or lack thereof	Off prokinetics, Off laxatives, Stable antidepressants

WMC = wireless motility capsule

There was overlap among the questions, and some of the same research questions could be answered within the same study design if performed on appropriate populations, and with appropriate intervention comparisons (Table 6). We elected to keep a greater number of questions and focus on better study design for these types of studies.

Table 6. Major clinical and methodological questions were within same few categories

Categories of Clinical Questions	Categories of Methodological Questions
• Comparative diagnostic accuracy and safety of gastric and colonic dysmotility testing • Establishing a gold standard for gastric dysmotility • Comparison of WMC testing and scintigraphy • Establishing a gold standard for colonic dysmotility • Reproducibility of WMC testing for gastroparesis • Incremental diagnostic accuracy for both gastric and colonic dysmotility • Distinction of subtypes of gastroparesis and determining the best test or combination for predicting outcomes after colectomy for colonic dysmotility • Comparative accuracy in predicting response to treatment for colonic dysmotility	• WMC test data in patients with gastric and colonic dysmotility and non-diseased populations • Rates of test failures and need for additional tests • WMC diagnostic accuracy and safety in gastroparesis subtypes • Thresholds of analysis and correlation between WMC test values and histopathology

WMC = wireless motility capsule

Ongoing Research

We scanned the literature for any ongoing clinical trials which may have already addressed these high-priority areas. We found two clinical trials.gov references which had already published the results on populations outside of the scope of our review, spinal cord injury patients and critically ill intensive care patients. One clinicaltrials.gov reference was not assessing the role of WMC in diagnosis of constipation or gastroparesis, but instead focused on acid measurement. We also identified a funded research protocol which likely has yet to complete enrollment regarding effect of medication for constipation on outcomes with WMC. We await the outcome of this trial, but it is only representative of one of many treatment modalities available to these patients. No other identified research projects were identified which had already answered the questions we designed.

Discussion

The diagnosis and management of motility disorders is complex. In part it is difficult because our current testing modalities are suboptimal and hard to standardize. In part it is complicated by patient factors, like medications, which interfere with testing, or medical conditions like diabetes which tend to make patients worse over time. Also, patients often have more than one disorder, and a substantial percentage will have both gastric and colonic dysmotility at the same time. In addition, although patients with gastroparesis and constipation are often lumped in one category clinically, these patients likely represent different clinical phenotypes with different underlying pathophysiology and presumably different responses to treatment. It is unclear how objective measures of dysmotility (such as impaired gastric emptying) are related to symptoms as many studies have not shown a clear correlation between treatment response and objective motility improvement.

There is a burgeoning population of patients with gastroparesis, estimated at 1.5 to 3 million Americans, by some models. Based on current trends in obesity and diabetes, this problem has the potential to compound over the next 10 to 20 years. Measures taken now to invest in establishing good diagnostic tests and deciding about whether or not WMC is a useful modality will have a large impact on costs and may have potential impact on management of these diseases.

Our method of determining FRN has strengths and limitations. One problem was the narrow scope of our original evidence report focusing on gastroparesis and constipation and the comparative role of WMC testing. Since our initial review did not include small bowel transit or whole gut transit abnormalities in the review, some aspects of the potential benefit of WMC testing may not have been established by our work. By focusing primarily on WMC testing, we may not have captured all of the needs for research on the evaluation of gastroparesis and constipation. However, after thorough analysis of the data, clear gaps were seen in the methodological side and clinical side of research on gastroparesis and slow transit constipation with relation to WMC testing. We did not include gaps in basic science of neurogastroenterology, but we know there are many areas of ongoing research into this area which may also need to be assessed. This was beyond the scope of the original review.

Due to time constraints related to the desire to prepare a report on FRN as soon as possible after completion of the evidence report, we used an abbreviated Delphi technique to determine the priorities of the stakeholders. In the first round of the group judgment process, the stakeholders varied widely in the priorities they assigned to the gaps. In the second round of the process, the stakeholders still showed moderate variation in their priorities, indicating that there was not total consensus on the priorities. This finding is not surprising, however, given that we intentionally included stakeholders having very different perspectives. Nonetheless, the process revealed a number of gaps that different stakeholders considered very important to address.

We had a limited number of stakeholders, and thus may not have a totally representative view of all relevant stakeholders. It would have taken longer to collect information from more stakeholders because a larger group would require review and approval by the Office of Management and Budget. Our stakeholders represented the important perspectives, by including clinical experts, methodologic experts, a patient/consumer advocate. Together, they identified high priorities for future research. Their thoughtful comments were very useful in creating the FRN report.

The summative results of the high-priority research items showed a consistent pattern of identifying many of the same issues and questions that were identified in the original systematic review. The original questions posed were often selected again here for further research inquiry.

Conclusion

There is evidence for WMC in many areas of clinical use for the detection, diagnosis, treatment and management of gastric and colonic dysmotility, however much of the evidence was either low strength or insufficient for making evidence-based recommendations. We focused on creating a viable set of research questions, which was reviewed by an expert panel and revised into a tiered list of priorities for future research. These FRN will help to immediately impact care by establishing firmly the evidence-base for use of WMC and potentially lead to improved diagnosis, detection, treatment and management of motility disorders which have had few tools for accurate, reliable, portable, non-radiating standardized diagnosis in the past. Although colonic and gastric dysmotility are not as common as high blood pressure or diabetes, the burden of disease for gastroparesis and severe colonic dysmotility is great, and accurate and rapid diagnosis on a consistent basis would be the cornerstone of developing research projects to reliably compare patients with similar disease characteristics to each other. We urgently need these additional studies to help guide the development of formal evidence-based guidelines as compared with the current consensus guidelines based on limited evidence to develop clinically meaningful recommendations for providers and payers. The top tier of research priorities are listed in Table 7.

Table 7. Top-tier research priorities listed in order

Clinical Questions
Among patients with presumed gastroparesis, how would clinical outcomes differ between patients managed with WMC data alone versus those managed with scintigraphy data alone?
Among patients with suspected gastroparesis, how does the WMC compare with gastric scintigraphy, antroduodenal manometry, and endoscopy in its accuracy and safety in diagnosing gastroparesis?
What is an appropriate comparison test (gold standard) to which the WMC should be compared in the diagnosis and monitoring of patients with gastroparesis (e.g., clinical diagnosis, 4-hour gastric scintigraphy, antroduodenal manometry, endoscopy, plain X-ray of the abdomen 5 hours after swallowing a “dumb” radiopaque pill, a combination of tests)?
In the clinical evaluation of patients with suspected gastroparesis, do abnormal WMC gastric emptying times correlate with 4-hour gastric scintigraphy results well enough to replace scintigraphy, or is a confirmatory scintigraphy required after an abnormal WMC test?
What is an appropriate comparison test (gold standard) to which the WMC should be compared in the diagnosis and monitoring of patients with slow-transit constipation (e.g., clinical diagnosis, colonic scintigraphy, radiopaque markers, a combination of tests)?
In normal patients and those with gastroparesis, how reproducible are WMC studies performed 2 weeks apart?
Among patients with suspected slow-transit constipation how does the WMC compare with radiopaque markers and scintigraphy in its accuracy and safety of diagnosis?
What is the incremental value of using the WMC in addition to gastric scintigraphy, antroduodenal manometry or endoscopy to diagnose delayed gastric emptying or dysmotility?
Among patients with confirmed colonic dysmotility who undergo colectomy, is WMC data alone, or scintigraphy plus radiopaque markers, or the combination of all three the best predictor of whole gut dysmotility and clinical outcomes?

Table 7. Top-tier research priorities listed in order (continued)

Methodological Questions
Among patients with gastric and colonic dysmotility, what are the ranges of WMC values for temperature, pH, pressure pattern and transit time in the stomach, intestines and colon?
In non-diseased populations, what are the distributions of age-, race- and sex-specific values for pressure patterns/amplitude/frequency, temperature, pH, and transit time in the stomach, intestines and colon, as measured by the WMC?
How does the WMC compare with scintigraphy, manometry and radiopaque markers in the rates of test failures and need for additional tests?
How do the diagnostic accuracy, safety and resource utilization of the WMC differ when used in ambulatory gastroenterology clinics and primary care offices - and when the WMC results are interpreted by primary gastroenterologists, primary care physicians or nurse practitioners?
What thresholds should be used to differentiate diagnostic accuracy between the WMC and another test (e.g., sensitivity, specificity, % agreement, other)?
Among patients with motility disorders, are WMC values correlated with the histopathological findings from full-thickness or standard biopsy?

WMC = wireless motility capsule

References

10. Hasler WL. Gastroparesis: pathogenesis, diagnosis and management. Nat Rev Gastroenterol Hepatol. 2011;8(8):438-53. PMID: 21769117.

11. 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.

12. 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.

13. 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.

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. 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.

17. 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.

18. 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.

19. 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.

20. Carey T, Sanders GD, Viswanathan M, Trikalinos TA, Kato E, Chang S. Framework for Considering Study Designs for Future Research Needs. Methods Future Research Needs Paper No. 8. (Prepared by the RTI–UNC Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 12-EHC048-EF. Rockville, MD: Agency for Healthcare Research and Quality. March 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Acronyms and Abbreviations

AHRQ	Agency for Healthcare Research and Quality
ANMS	American Neurogastroenterology and Motility Society
COI	Conflict of Interest
EHC	Effective Health Care
EPC	Evidence-based Practice Center
FDA	Food and Drug Administration
FRN	Future Research Needs
KQ	Key Question
PICOTS	Population, Intervention, Comparison, Outcome, Timing, Setting
ROM	Radiopaque Markers
SOE	Strength of evidence
WMC	Wireless motility capsule

Appendix A. Search Strategy for Ongoing Studies

Resource **URL**	**Search Parameters**	**Search Terms/Strategy**
ClinicalTrials.gov clinicaltrials.gov/	Advanced search, Conditions field used	Wireless motility capsule OR Smartpill
EU Clinical Trials Register www.clinicaltrialsregister.eu/	Not applicable	Wireless motility capsule OR Smartpill
NIH Reporter projectreporter.nih.gov/reporter.cfm	Projects field searched	Wireless motility capsule OR Smartpill
Canadian Institutes of Health Research www.cihr-irsc.gc.ca/	Funding Decisions Data field searched	Wireless motility capsule OR Smartpill
World Health Organization International Clinical Trials Registry Platform Search Portal apps.who.int/trialsearch/	Searched Condition field, Recruitment status = ALL	Wireless motility capsule OR Smartpill

Appendix B. List of Ongoing Studies

Title/ Identifier(s)	Study Dates	Description	Sponsor or Principal Investigator Collaborator(s)	Source
Title: Lubiprostone Effect on Gastrointestinal (GI) Tract Transit Times Measured by Smartpill in Patients With Chronic Constipation **Identifier(s):** NCT01469819	**Start date:** November 2011 **Estimated study completion date:** December 2012 **Estimated primary completion date:** December 2012 (Final data collection date for primary outcome measure)	**Purpose:** To measure the time difference in the duration of transit of the FDA approved SmartPill capsule in all segments of gastrointestinal (GI) tract before and after exposure to lubiprostone. **Study design:** Endpoint Classification: Efficacy Study Intervention Model: Single Group Assignment Masking: Open Label Primary Purpose: Treatment **Condition(s):** Chronic Idiopathic Constipation **Intervention(s):** Drug: Lubiprostone **Estimated enrollment:** 15	Texas Tech University Health Sciences Center Takeda Pharmaceuticals North America, Inc.	ClinicalTrials.gov **Accessed at:** clinicaltrials.gov/ct2/show/NCT01469819

Title/ Identifier(s)	Study Dates	Description	Sponsor or Principal Investigator Collaborator(s)	Source
Title: SmartPill Monitoring for Assessment of GI Function in SCI **Identifier(s):** NCT00856648	**Start date:** April 2009 **Estimated study completion date:** December 2011 **Estimated primary completion date:** April 2011 (Final data collection date for primary outcome measure)	**Purpose:** To evaluate the relationship between the level of SCI and the impairment of Colonic transit time (CTT) and Total transit time (TTT) by using the SmartPill device. **Study design:** Observational Model: Case Control Time Perspective: Prospective **Condition(s):** SCI **Intervention(s):** Device: SmartPill ingestion and monitoring **Estimated enrollment:** 40	Department of Veterans Affairs	ClinicalTrials.gov **Accessed at:** clinicaltrials.gov/ct2/show/NCT00856648

Title/ Identifier(s)	Study Dates	Description	Sponsor or Principal Investigator Collaborator(s)	Source
Title: A New Method for Determining Gastric Acid Output Using a Wireless Capsule **Identifier(s):** NCT00702533	**Start date:** June 2008 **Estimated study completion date:** Ongoing (Currently recruiting participants) **Estimated primary completion date:** Ongoing (currently recruiting participants)	**Purpose:** To evaluate the usefulness and accuracy of the SmartPill for gastric analysis, compared with current procedures. **Study design:** Observational, Time Perspective: Prospecti ve **Condition(s):** Gastrointestinal Diseases **Intervention(s):** Wireless Capsule SmartPill **Estimated enrollment:** 80	National Institute of Diabetes and Digestive and Kidney Diseases	ClinicalTrials.gov **Accessed at:** clinicaltrials.gov/ct2/show/NCT00702533

Title/ Identifier(s)	Study Dates	Description	Sponsor or Principal Investigator Collaborator(s)	Source
Title: Evaluation of Gastrointestinal Motility With SmartPill **Identifier(s):** NCT01159002	**Start date:** November 2007 **Estimated study completion date:** January 2009 **Estimated primary completion date:** May 2008 (Final data collection date for primary outcome measure)	**Purpose:** To evaluate the clinical usefulness of a capsule (SmartPill~) measuring pH, pressure and temperature from within the entire GI tract to determine gastric emptying time, combined small and large bowel transit time and total transit time. **Study design:** Observational Model: Cohort Time Perspective: Prospective **Condition(s):** Gastrointestinal Motility **Intervention(s):** Device: SmartPill **Estimated enrollment:** 8	University of Louisville	ClinicalTrials.gov **Accessed at:** clinicaltrials.gov/ct2/show/NCT01159002

Appendix C. Individual Stakeholder Priority Scores for Future Research Questions

Topic	Stake-holder 1	Stake-holder 2	Stake-holder 3	Stake-holder 4	Stake-holder 5	Stake-holder 6	Stake-holder 7	Stake-holder 8	Stake-holder 9	Sum (highest sum = highest priority)	Priority Rank (1 = highest priority)
List 1: Clinical Research The individual stakeholder ratings have been inverted so that the highest value (6) is the highest priority.											
Top Tier											
1. Comparative effectiveness of managing with WMC (gastric dysmotility/gastroparesis)	6	4		6				4	6	26	1
2. Comparative diagnostic accuracy and safety (gastric dysmotility/gastroparesis)	4		6	4	4		5			23	2
3. Gold standard (gastric dysmotility/gastroparesis)				1	6	6	6			19	3
4. Comparison of WMC and scintigraphy results (gastroparesis)		6	4					3	5	18	4

5. Gold standard (colonic dysmotility/slow-transit constipation)					5	5	4			14	5
6. Reproducibility of WMC results (gastroparesis)		5		2		4			3	14	5
7. Comparative diagnostic accuracy and safety (colonic dysmotility/slow-transit constipation)			5	5	1					11	6
8. Incremental diagnostic accuracy (gastric dysmotility/gastroparesis)					3		2	5		10	7
9. Incremental diagnostic accuracy (colonic dysmotility/slow-transit constipation)					2		1	6		9	8
Other Gaps											
10. Distinction of subtypes of gastroparesis (gastric dysmotility/gastroparesis)		1				2			4	7	9
11. Best test or combination for predicting outcomes after colectomy (colonic dysmotility/slow-transit constipation)	5								2	7	9

12. Comparative accuracy in predicting response to treatment (colonic dysmotility/slow-transit constipation)	3	3				1				7	9
13. Utility of colonic pressure (colonic dysmotility)	1		1					2	1	5	10
14. Differentiating IBS-C from idiopathic constipation (colonic dysmotility/slow-transit constipation)		2				3				5	10
15. Comparisons of WMC pressure profiles (gastroparesis)			2	3						5	10
16. Best test or combination for predicting outcomes without colectomy (colonic dysmotility/slow-transit constipation)			3					1		4	11
17. Comparative accuracy in predicting response to treatment (gastric dysmotility/gastroparesis)	2									2	12
18. Comparative value in monitoring response to treatment (gastric dysmotility/gastroparesis)										0	13

19. Comparative value in monitoring response to treatment (colonic dysmotility/slow-transit constipation)										0	13
20. Colonic pressure patterns after waking (colonic dysmotility)										0	13
21. Evaluation of high-amplitude contractions (colonic dysmotility)										0	13
List 2: Questions Involving Methodological Research											
Top Tier											
1. WMC data in patients with gastric and colonic dysmotility*	3	6	5	7	6	7	6	6	1	47	1
2. WMC data in non-diseased populations*	2	7	4		7	4	7	7	2	40	2
3. Rates of test failures and need for additional tests*	6		6	5	5		4	3	5	34	3
4. WMC diagnostic accuracy and safety in gastroparesis subtypes*	7	4		3	4	5	1	5		29	4
5. Thresholds*		5	7		3	6	5	1		27	5
6. Correlation between WMC values and histopathology*				6		1		4	7	18	6
Other Gaps											

7. Colonic pressure patterns following meal ingestion.	5	2		4	1				3	15	7
8. Relation of baseline WMC values with long-term followup values	1	3					3		6	13	8
9. WMC data compared with other tests in non-diseased populations	4		3		2	3				12	9
10. Concurrent tests			1	2		2			4	9	10
11. Masking interpreters of WMC data		1					2	2		5	11
12. Use of historical controls in WMC research			2							2	12

C = clinical knowledge gap; M = methodological knowledge gap; WMC = wireless motility capsule.