Published in final edited form as:

Clin Gastroenterol Hepatol. 2017 November ; 15(11): 1708–1714.e3. doi:10.1016/j.cgh.2016.07.024.

A System to Assess the Competency for Interpretation of Esophageal Manometry Identifies Variation in Learning Curves

Rena Yadlapati¹, Rajesh N. Keswani¹, Jody D. Ciolino², David P. Grande¹, Zoe I. Listernick¹, Dustin A. Carlson¹, Donald O. Castell³, Kerry B. Dunbar⁴, Andrew J. Gawron⁵, C. Prakash Gyawali⁶, Philip O. Katz⁷, David Katzka⁸, Brian E. Lacy⁹, Stuart J. Spechler⁴, Roger Tatum¹⁰, Marcelo F. Vela¹¹, and John E. Pandolfino¹

¹Division of Gastroenterology and Hepatology, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author: Rena Yadlapati, MD, Division of Gastroenterology & Hepatology, Northwestern University, 676 North St. Clair St. Suite 1400 Chicago, IL 60611, Office: (312) 695-4065; Fax: (312) 695-3999, rena.yadlapati@northwestern.edu.

Disclosures/Conflicts of Interests: RNK: Consults for Boston Scientific and Cook Endoscopy; Speaks for Covidien; MFV: Consults for Covidien. Was a prior speaker for Sandhill Scientific; JEP: Consults for Covidien, Sandhill Scientific, and Given. The authors have no other reported disclosures or conflicts of interest.

Writing Assistance: None

Author Contributions: *Rena Yadlapati:* Study concept & design; acquisition of data; analysis and interpretation of data; drafting of manuscript; critical revision of the manuscript for important intellectual content; approval of final manuscript

Rajesh N. Keswani: Study concept & design; acquisition of data; drafting of manuscript; critical revision of the manuscript for important intellectual content; approval of final manuscript

Jody D. Ciolino: Data analysis; drafting of manuscript; critical revision of the manuscript for important intellectual content; approval of final manuscript

David P. Grande: Study concept & design; acquisition of data; critical revision of the manuscript for important intellectual content; approval of final manuscript

Zoe I. Listernick: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Dustin A. Carlson: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Donald O. Castell: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Kerry B. Dunbar: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Andrew J. Gawron: Study concept & design; acquisition of data; analysis and interpretation of data; drafting of manuscript; critical revision of the manuscript for important intellectual content

Chandra P. Gyawali: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Philip Katz: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

David Katzka: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Brian Lacy: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Stuart Spechler: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Roger Tatum: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

Marcelo F. Vela: Study concept & design; critical revision of the manuscript for important intellectual content; approval of final manuscript

John E. Pandolfino: Study concept & design; acquisition of data; drafting of manuscript; critical revision of the manuscript for important intellectual content

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

²Department of Preventive Medicine-Division of Biostatistics, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

³Division of Gastroenterology & Hepatology, Medical University of South Carolina, Charleston, SC, USA

⁴University of Texas Southwestern Medical Center and the Dallas VA Medical Center, Department of Internal Medicine, Division of Gastroenterology and Hepatology, Dallas, TX, USA

⁵Division of Gastroenterology, University of Utah, Salt Lake City, IL, USA

⁶Division of Gastroenterology, Washington University School of Medicine, Saint Louis, Missouri, USA

⁷Thomas Jefferson University, Philadelphia, Pennsylvania; Division of Gastroenterology, Albert Einstein Medical Center, Philadelphia, PN, USA

⁸Division of Gastroenterology, Mayo Clinic, Rochester, MN, USA

⁹Division of Gastroenterology, Dartmouth Hitchcock Medical Center, Lebanon, NH, USA

¹⁰Department of Surgery, University of Washington, Seattle, WA, USA

¹¹Division of Gastroenterology, Mayo Clinic Arizona, Scottsdale, AZ, USA

Abstract

Background & Aims—Quality esophageal high-resolution manometry (HRM) studies require competent interpretation of data. However, we have little understanding of learning curves, training requirements, or measures of competency for HRM. We aimed to develop and use a competency assessment system to examine learning curves for interpretation of HRM data.

Methods—We conducted a prospective multicenter study of 20 gastroenterology trainees with no experience in HRM, from 8 centers, over an 8 month period (May through December 2015). We designed a web-based HRM training and competency assessment system. After reviewing the training module, participants interpreted 50 HRM studies and received answer keys at the fifth and then at every second interpretation. A cumulative sum procedure produced individual learning curves with preset acceptable failure rates of 10%; we classified competency status as: competency not achieved, competency achieved, or competency likely achieved.

Results—Five (25%) participants achieved competence, 4 (20%) likely achieved competence, and 11 (55%) failed to achieve competence. A minimum case volume to achieve competency was not identified. There was no significant agreement between diagnostic accuracy and accuracy for individual HRM skills.

Conclusion—We developed a competency assessment system for HRM interpretation; using this system, we found significant variation in learning curves for HRM diagnosis and individual skills. Our system effectively distinguished trainee competency levels for HRM interpretation and contrary to current recommendations, found that competency for HRM is not case-volume specific.

Keywords

CUSUM; competency-based medical education; high-resolution manometry

Introduction

Heartburn and dysphagia account for over two million annual outpatient ambulatory visits and evaluation often requires esophageal manometry.^{1, 2} All gastroenterologists should have a strong understanding of esophageal motility disorders, and experts must be proficient in interpreting esophageal manometry.²⁻⁴ However, a minority of trainees receive formal esophageal manometry training; in these rare instances, training is primarily apprenticeship based, relying on case volume, time-based training periods and subjective evaluations as surrogates for competency.³ Research in medical learning, including work in gastrointestinal endoscopy, repeatedly demonstrates that specific case volumes do not ensure competency, and robust competency assessment methods are imperative to producing clinicians that provide high-quality care in unsupervised practice.⁵⁻⁹

The critical deficiency in competency-based training translates to marked variability in the interpretation of esophageal manometry. We previously found that agreement in diagnostic impressions for esophageal high-resolution manometry (HRM) amongst trainees and physicians in practice is suboptimal.¹⁰ Diagnostic inaccuracies adversely impact patient outcomes, require repeated studies and contribute to avoidable healthcare costs, and experts recognize that poor interpreter competency remains a major impediment to a quality HRM study.² However, an understanding of learning curves and measure of competency for HRM does not exist, in part, due to a deficiency in standardized training systems and established competency benchmarks for HRM.^{2, 3}

In this study, we aimed to develop and use a competency assessment system for HRM to examine learning curves for HRM interpretation amongst HRM-naïve trainees. We hypothesized that an integrated web-based system will effectively measure trainee competency, and that learning curves for HRM will vary amongst trainees.

Methods

Study Design & Participants

We performed a prospective multicenter trial at 10 ACGME-accredited gastroenterology training programs across the United States over an eight-month period (May 2015 to December 2015). Gastroenterology trainees were included as study participants if they had no to minimal (read less than five HRM studies) prior exposure to HRM. Northwestern University's Institutional Review Board (IRB; STU00201267) approved this study, and all participants provided written informed consent.

Over a two-month development phase (May 2015 to July 2015) the study team designed a web-based training module and competency assessment system for HRM. The training system consisted of a handbook including a step-by-step educational tutorial on ManoView version 3.0.1™ (Given Imaging, Los Angeles, CA, USA) software use and interpretation, a

review of HRM metrics and definitions according to Chicago Classification version 3.0¹¹, and an embedded video lecture delivered by an expert in HRM (JEP). The lecture consisted of an introduction to HRM and a recommended systematic approach to interpreting HRM. The competency assessment tool consisted of 50 de-identified HRM cases seen at Northwestern University and included Achalasia Type I (5), Type II (7), Type III (4), Esophagogastric Junction (EGJ) Outflow Obstruction (6), Distal Esophageal Spasm (3), Absent Contractility (3), Hypercontractile (Jackhammer) Esophagus (4), Fragmented Peristalsis (3), Ineffective Esophageal Motility (6), Normal Esophageal Motility (6), and Technically Limited studies (3), in random order. Four study team members (an esophageal expert (JEP), two advanced esophageal fellows (RY, DC), and an esophageal motility technician (ZL)) individually reviewed each study and developed a consensus answer key. In cases of minor disorders, either a diagnosis of Ineffective Esophageal Motility or Fragmented Peristalsis was acceptable.

Participants reviewed the web-based training module and completed a baseline questionnaire. Participants then completed 50 HRM studies within a five-month period (July 2015 to November 2015). Participants used locally installed ManoView software to interpret all de-identified HRM studies and submitted their study interpretations through REDCap (Research Electronic Data Capture), a secure web-based data collection application, housed at Northwestern University.¹² Interpretations were based on quality measures previously identified by the experts in a RAND UCLA Appropriateness Method process and included locations of the: upper esophageal sphincter (UES) and pressure inversion point (PIP), measurement of: basal EGJ pressure, integrated relaxation pressure (IRP), distal contractile integral (DCI), and distal latency (DL), determination of: EGJ morphology, presence of hiatal hernia, peristaltic integrity, pressurization pattern, and overall diagnosis according to Chicago Classification Scheme version 3.0 (Figure 1).² Participants received structured feedback via e-mail after the fifth and then every second submitted interpretation. This feedback consisted of the correct HRM interpretation and an associated explanation. Correct answers were based on number ranges for: UES location (± 1 cm), basal EGJ pressure (± 5 mmHg), PIP location (± 1.5 cm), IRP (± 2.5 mmHg), and multiple choice options for: EGJ morphology (type I, II, or III), presence of hiatal hernia (yes or no), DCI (<450, 450-8000, or >8000 mmHg-s-cm), DL (<4.5, 4.5-5.5, or > 5.5 sec), peristaltic integrity (normal or fragmented), pressurization pattern (none, panesophageal, or compartmentalized), and overall diagnosis; additionally confidence level in interpretation (scale of 1-10, with 10 being the highest) were reported. At the end of the study, participants received information on their overall performance and competency level, and completed a post-study questionnaire. Following successful completion of the study, participants were remunerated.

Outcomes

The primary outcome was the cumulative learning curve score for overall diagnostic accuracy for each participant. The secondary outcomes were the cumulative learning curve scores for diagnostic accuracy in individual skills (e.g., basal EGJ pressure) for each participant and diagnostic accuracy by diagnosis type.

Statistical Analysis

Evaluation of competency status employed a modified cumulative sum (CUSUM) procedure to assess the learning curve for each participant. CUSUM control charts continuously assess the performance of an individual against a predetermined standard. The CUSUM procedure for learning curves of this nature requires pre-specification of four parameters: (1) p_0 , the acceptable failure rate, (2) p_1 , the unacceptable failure rate, (3) type I error (α), (4) type II error (β). In this case, we set $p_0=0.1$ *a priori* as 90% diagnostic accuracy is the expert performance standard, and we set $p_1=0.2$ as standard practice suggests a reasonable threshold is $p_1 = 2(p_0)$.^{13,14} We set type I and type II errors at 10% which are typical values in the context of performance monitoring.¹⁵ Each diagnostic accuracy outcome was binary (success = correct diagnosis, failure = incorrect diagnosis). We detail the CUSUM score calculation in supplemental material (Supplemental Text). Briefly, we set CUSUM to zero for the first case. For each subsequent assessment of accuracy we added to the CUSUM score in case of an incorrect diagnosis and subtracted from the CUSUM score in case of a correct diagnosis. If the score crossed a pre-specified threshold (indicative of an out of control process), CUSUM score was reset to zero. An upward projection of the CUSUM score or one that requires a reset to zero according to the algorithm outlined above suggests unstable competency for diagnostic accuracy. A stable or downward projection indicates anticipated or increasing competency for diagnostic accuracy. Upon completion of the 50 HRM studies, we defined participants' competency as Competency Not Achieved if a reset occurred within the last 10 studies and/or if the final CUSUM score exceeded 1.0, Competency Achieved if the final CUSUM score fell below 0.0, and Competency Likely Achieved if the final CUSUM score fell between 0.0 to 1.0. A subsequent sensitivity analysis considered 80% diagnostic accuracy as the performance standard ($p_0 = 0.2$, $p_1 = 0.3$).

In order to evaluate association between underlying disorder and diagnostic accuracy for any given case study we employed a generalized linear model with logit link for accuracy using diagnosis as the sole predictor. We used generalized estimating equations (GEE) with the robust sandwich estimator to account for within case/participant associations. We present model-estimated (least squares) means and compared pairwise differences across diagnoses using the Tukey-Kramer adjustment for multiple pairwise comparisons. CUSUM plots were generated in R (Version 3.2.2, Copyright 2015 by The R Foundation for Statistical Computing), and additional analyses were performed in SAS (Version 9.4, Copyright 2012 by SAS Institute, Inc.).

Results

Participants

Thirty trainees from 10 centers were invited to participate; 25 trainees enrolled and 20 (80%) trainees from eight centers completed the entire study (i.e., interpreted all 50 HRM studies). Of the five trainees that withdrew, three were first year fellows and dropped out due to time constraints, and four withdrew prior to case 20 without having achieved competence. Mean age was 32.2 ± 2.5 years and 17 (85%) were male. Participants were evenly distributed through fellowship training: six (30%) first year, seven (35%) second year and seven (35%)

third year. The majority of participants were in a clinical training track and in a university training setting. Eighteen (90%) had no prior manometry experience and two (10%) participants previously analyzed 1-5 HRM studies. The median confidence level in reading HRM prior to study onset was 1 (range: 1-5) on a scale of 1-10 (1 indicating the lowest level of confidence) (Table 1). One HRM study was excluded from final analysis given later disagreement in diagnosis amongst the study team; thus each participant contributed data for 49 studies, resulting in a total of 980 studies for analyses.

Primary Outcome

Nine (45%) participants demonstrated acceptable performance at the end of the study: five (25%) definitely achieved competence and four (20%) likely achieved competence. Eleven (55%) participants failed to achieve competence by study conclusion (Figure 2).

Competence was achieved by the five participants at variable case volumes (Case number 1, 22, 43, 44, and 50). There were no significant differences in year of training ($p = 0.2$), number of prior manometry studies interpreted ($p = 0.2$), or university versus non-university training setting ($p = 0.5$) with regard to likelihood of achieving competency.

In the sensitivity analysis using 80% diagnostic accuracy as acceptable, 11 (55%) definitely achieved competence, three (15%) likely achieved competence, and six (30%) did not achieve competence (Supplemental Figure).

Secondary Outcomes

Individual HRM Skills—The majority of participants achieved or likely achieved competency for UES location (95%), DL (70%), peristaltic integrity (60%), and pressurization pattern (60%). In contrast, no participants reached competency for EGJ morphology and presence of hiatal hernia, and only a minority achieved or likely achieved competency for basal EGJ pressure (5%), PIP location (35%) and IRP (25%). Data suggest lack of significant agreement between diagnostic accuracy and accuracy for individual HRM skills, as depicted in Table 2.

Diagnostic Accuracy—Diagnostic accuracy was higher for Type I Achalasia (87%), Normal Esophageal Motility (86%), Type II Achalasia (77%), Hypercontractile Esophagus (76%), Absent Contractility (75%), and EGJ Outflow Obstruction (69%). Alternatively, diagnostic accuracy was lower for Technically Limited studies (30%), Distal Esophageal Spasm (53%), Type III Achalasia (58%), and Fragmented Peristalsis (60%). GEE model results suggest that diagnosis significantly predicts trainee accuracy ($p = 0.04$; Supplemental Table). In addition, increasing trainee confidence (in identifying the correct diagnosis) corresponded to an increased odds of diagnostic accuracy (OR 1.4, 95% CI 1.3, 1.4; $p < 0.01$).

Post-Study Feedback

In post-study surveys, 95% of participants responded that study participation increased their likelihood of independently reading HRM in practice. Overall mean confidence level (Post – Pre) increased across all groups (Competence achieved: 5.6 ± 0.9 ; Competence likely achieved: 5.5 ± 1.3 ; Competence not achieved: 5.8 ± 1.5 ; $p = 0.9$). Trainee comfort with

independently reading HRM in unsupervised practice correlated with competence achieved (Spearman Rho 0.45; $p = 0.05$). All participants who achieved or likely achieved competence indicated that they periodically reviewed the baseline educational material throughout the study (at least three time points). Many participants suggested that more detailed visual feedback following studies would be helpful and reported ongoing difficulty in understanding EGJ morphology.

Conclusions

We report the successful development of a competency assessment system for HRM interpretation, and present the results of a prospective multi-center study using this system in the evaluation of trainee HRM learning curves. We found significant variation in learning curves for HRM diagnosis and individual skills, and data suggest that a minimum set case volume does not ensure competency. Specifically, approximately half of participants did not achieve competency after viewing a training module and interpreting 50 HRM studies. In addition, competency for individual metrics of interpretation was not discriminatory for overall competency for HRM. Finally, participants had particular difficulty in identifying Technically Limited studies, Type III Achalasia, and Distal Esophageal Spasm.

In our current training model, the majority of gastroenterologists are not prepared to interpret gastrointestinal motility testing. Although hands-on motility training is offered at select centers of excellence in the US, these programs are unable to meet the training needs of thousands of trainees and gastroenterologists in practice.¹⁶ Deficiencies in formal structured training and assessments of competency results in marked variation in practice quality and compromises patient care. Interpretation of esophageal HRM is complex, and an individualized competency based training approach is needed.¹⁶

Existing literature on competency based training in gastroenterology primarily focuses on endoscopic training, reporting that learning curves vary for colonoscopy, endoscopic ultrasound and endoscopic retrograde cholangiopancreatography and suggests that current training structures and societal recommendations underestimate requirements for procedural competency.^{4-6, 17, 18} In colonoscopy training, assessments of individual skills (i.e., cecal intubation rate or insertion time) are not reliable indicators of overall competency,¹⁹ and current studies are examining comprehensive tools to measure endoscopic skills.²⁰⁻²³ Use of computer-based tools and virtual reality simulation with active feedback seem to be effective training methods for colonoscopy and provide reliable measures of competency.²⁴⁻²⁶ Similar to the work in endoscopy, our study suggests that competency for the interpretation of esophageal HRM should not be determined on the sole basis of case-volume, and that the current recommendation of interpreting 50 esophageal manometry studies to achieve expertise likely underestimates training requirements.⁴

Furthermore, our study provides important information about trainee learning patterns. For instance, our study corroborates findings from prior smaller studies demonstrating suboptimal interrater agreement for spastic motor disorders (Type III Achalasia and Distal Esophageal Spasm).^{10, 27} Additionally, skills to identify technically limited studies need to be emphasized as technical limitations influence manometric readings, and identification

remains a challenge. Conversely, important patterns such as Normal Esophageal Motility and Type I or II Achalasia seem to be more readily recognized, even by those failing to achieve full competence. As expected, EGJ morphology type was challenging for participants; since this skill is not critical to accurately diagnosing a motor disorder, the nuances of EGJ morphology may be more effectively learned later in the training process and simple recognition of a large (>5cm) hernia may be of greater clinical relevance rather than specification of morphology type. This data should be considered in the future development of formal training curricula.

There are important limitations to our study. Future studies should examine the validity of these results amongst a larger, diverse group of participants. Despite our attempts to recruit trainees from a variety of training settings and years, the majority were trainees at motility centers of excellence with potentially greater exposure to hands-on, in person training than the general trainee population, thus limiting the generalizability and potentially overestimating the rate of competence achieved. In addition, future learning curve assessments require examination over a higher case-volume to fully delineate learning patterns since the majority did not achieve competency by the 50th case. Our tool was not entirely automated and did not provide individualized performance feedback. Future designs should use an integrated multimedia interface that provides automated immediate feedback in order to reduce trainee fatigue and prompt meaningful self-assessment. Adaptive learning strategies reinforce competencies achieved and direct continuous learning where gaps are uncovered; this is essential to an effective competency based-training system for diagnostic tools such as HRM.²⁸ The discrepancy between accuracy for diagnosis and individual HRM skills likely reflects the overly stringent criteria our research team pre-defined for acceptable answers for individual skills, and learning curves for individual skills may be better assessed using more clinically applicable definitions of accuracy such as a high or normal IRP rather than a specified IRP range. Finally, as demonstrated in the sensitivity analysis, outcomes differ when considering acceptable diagnostic accuracy as 90% ($p=0.1$) compared to 80% ($p=0.2$).

The time has come to shift our focus towards competency-based training in esophageal motility disorders. This study importantly highlights that learning curves for HRM vary and that a minimum case volume of 50 studies does not ensure competency. Furthermore, our data provides the foundation for future direction in designing an automated digital platform which fosters personalized adaptive learning and simultaneously monitors competency. This mechanism has the potential to fill a critical gap in current HRM training and assessment, and may have additional implications on training and certification requirements.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to acknowledge the grant contribution from the Alumnae of Northwestern University, and the gastroenterology trainees for participating in this study: Jeffrey M. Adler, Saad Alghamdi, Muhammad Ali, Omair Atiq, Jason Bill, Jeffrey W. Brown, Amanda Guentner, Brian Horner, Jennifer L. Horsley-Silva, Patrick

Laing, Donald J. Martin, Jason Martin, Tuan Pham, Anna Platovsky, Joseph Schowalter, Bryan Silon, Gurswhan Singh, Brett Slesman, Joshua A. Sloan, Ralph Vance

Grant Support: RY: Supported by T32 DK101363-02 grant.

References

1. Peery AF, Crockett SD, Barritt AS, et al. Burden of Gastrointestinal, Liver, and Pancreatic Diseases in the United States. *Gastroenterology*. 2015; 149:1731–1741 e3. [PubMed: 26327134]
2. Yadlapati R, Gawron AJ, Keswani RN, et al. Identification of Quality Measures for Performance of and Interpretation of Data From Esophageal Manometry. *Clin Gastroenterol Hepatol*. 2015
3. Rao SS, Parkman HP. Advanced training in neurogastroenterology and gastrointestinal motility. *Gastroenterology*. 2015; 148:881–5. [PubMed: 25805422]
4. American Association for the Study of Liver D, American College of G, American Gastroenterological Association I. et al. The Gastroenterology Core Curriculum, Third Edition. *Gastroenterology*. 2007; 132:2012–8. [PubMed: 17484892]
5. Wani S, Cote GA, Keswani R, et al. Learning curves for EUS by using cumulative sum analysis: implications for American Society for Gastrointestinal Endoscopy recommendations for training. *Gastrointest Endosc*. 2013; 77:558–65. [PubMed: 23260317]
6. Wani S, Hall M, Keswani RN, et al. Variation in Aptitude of Trainees in Endoscopic Ultrasonography, Based on Cumulative Sum Analysis. *Clin Gastroenterol Hepatol*. 2015; 13:1318–1325 e2. [PubMed: 25460557]
7. Barton JR, Corbett S, van der Vleuten CP, et al. The validity and reliability of a Direct Observation of Procedural Skills assessment tool: assessing colonoscopic skills of senior endoscopists. *Gastrointest Endosc*. 2012; 75:591–7. [PubMed: 22227035]
8. Sedlack RE. Training to competency in colonoscopy: assessing and defining competency standards. *Gastrointest Endosc*. 2011; 74:355–366. e1–2. [PubMed: 21514931]
9. Gupta S, Bassett P, Man R, et al. Validation of a novel method for assessing competency in polypectomy. *Gastrointest Endosc*. 2012; 75:568–75. [PubMed: 22154412]
10. Carlson DA, Ravi K, Kahrilas PJ, et al. Diagnosis of Esophageal Motility Disorders: Esophageal Pressure Topography vs. Conventional Line Tracing. *Am J Gastroenterol*. 2015; 110:967–77. quiz 978. [PubMed: 26032151]
11. Yadlapati R, Gawron AJ, Bilimoria K, et al. Development of quality measures for the care of patients with gastroesophageal reflux disease. *Clin Gastroenterol Hepatol*. 2015; 13:874–83 e2. [PubMed: 25460560]
12. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009; 42:377–81. [PubMed: 18929686]
13. Salowi MA, Choong YF, Goh PP, et al. CUSUM: a dynamic tool for monitoring competency in cataract surgery performance. *Br J Ophthalmol*. 2010; 94:445–9. [PubMed: 19951939]
14. Reynolds MR Jr, Stoumbos Zachary G. A CUSUM chart for monitoring a proportion when inspecting continuously. *Journal of Quality Technology*. 1999; 31
15. Kemp SV, El Batrawy SH, Harrison RN, et al. Learning curves for endobronchial ultrasound using cusum analysis. *Thorax*. 2010; 65:534–8. [PubMed: 20522852]
16. Yadlapati R, Keswani RN, Pandolfino JE. Competency based medical education in gastrointestinal motility. *Neurogastroenterol Motil*. 2016
17. Wani S, Hall M, Wang AY, et al. Variation in learning curves and competence for ERCP among advanced endoscopy trainees by using cumulative sum analysis. *Gastrointest Endosc*. 2015
18. Shahidi N, Ou G, Telford J, et al. When trainees reach competency in performing ERCP: a systematic review. *Gastrointest Endosc*. 2015; 81:1337–42. [PubMed: 25841579]
19. Patwardhan VR, Feuerstein JD, Sengupta N, et al. Fellowship Colonoscopy Training and Preparedness for Independent Gastroenterology Practice. *J Clin Gastroenterol*. 2016; 50:45–51. [PubMed: 26125461]

20. Sedlack RE, Coyle WJ, et al. Committee AT. ASGE's assessment of competency in endoscopy evaluation tools for colonoscopy and EGD. *Gastrointest Endosc.* 2014; 79:1–7. [PubMed: 24239255]
21. Coyle WJ. Developing tools for the assessment of the learning colonoscopist. *Gastrointest Endosc.* 2014; 79:808–10. [PubMed: 24721628]
22. Fried GM, Marks JM, Mellinger JD, et al. ASGE's assessment of competency in endoscopy evaluation tools for colonoscopy and EGD. *Gastrointest Endosc.* 2014; 80:366–7. [PubMed: 25034851]
23. Walsh CM, Ling SC, Khanna N, et al. Gastrointestinal Endoscopy Competency Assessment Tool: development of a procedure-specific assessment tool for colonoscopy. *Gastrointest Endosc.* 2014; 79:798–807 e5. [PubMed: 24321390]
24. Patel SG, Rastogi A, Austin G, et al. Gastroenterology trainees can easily learn histologic characterization of diminutive colorectal polyps with narrow band imaging. *Clin Gastroenterol Hepatol.* 2013; 11:997–1003 e1. [PubMed: 23466714]
25. Preisler L, Svendsen MB, Nerup N, et al. Simulation-based training for colonoscopy: establishing criteria for competency. *Medicine (Baltimore).* 2015; 94:e440. [PubMed: 25634177]
26. Walsh CM, Sherlock ME, Ling SC, et al. Virtual reality simulation training for health professions trainees in gastrointestinal endoscopy. *Cochrane Database Syst Rev.* 2012; 6:CD008237.
27. Fox MR, Pandolfino JE, Sweis R, et al. Inter-observer agreement for diagnostic classification of esophageal motility disorders defined in high-resolution manometry. *Dis Esophagus.* 2015; 28:711–9. [PubMed: 25185507]
28. Rose S, Shah BJ, Onken J, et al. Introducing the Gastroenterologist-Accountable Professionalism in Practice (G-APP) pathway: bridging the G-APP-replacing MOC with a model for lifelong learning and accountability. *Gastroenterology.* 2015; 149:1609–26. [PubMed: 26824093]
29. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 2005; 37:360–3. [PubMed: 15883903]

Abbreviations

CUSUM	Cumulative sum
DCI	Distal contractile integral
DL	Distal latency
EGJ	Esophagogastric junction
GEE	Generalized estimating equations
HRM	High-resolution manometry
IRB	Institutional Review Board
IRP	Integrated relaxation pressure
PIP	Pressure inversion point
UES	Upper esophageal sphincter

Landmark Phase (Questions 1-5):

1. Location of upper esophageal sphincter _____ (cm)
 2. EGJ morphology ☐ I ☐ II ☐ III ☐ Not measurable
 3. Hiatal hernia present ☐ Yes ☐ Not ☐ Not measurable
 4. Basal EGJ pressure _____ (mmHg) (specify 'not measurable' if not able to interpret)
 5. Location of pressure inversion point _____ (cm) (specify 'not measurable' if not able to interpret)

Findings on Swallow #1 (Questions 6-10):

6. Integrated relaxation pressure _____ (mmHg) (specify 'not measurable' if not able to interpret)
 7. Distal contractile integral ☐ <450 ☐ 450-8000 ☐ >8000 ☐ Not measurable (mmHg-s-cm)
 8. Distal latency ☐ <4.5 ☐ 4.5-5.5 ☐ >5.5 ☐ Not measurable (seconds)
 9. Peristaltic integrity ☐ Normal ☐ Fragmented ☐ Not measurable
 10. Pressurization pattern ☐ None ☐ Compartmentalized ☐ Panesophageal

Based on overall study (10 swallows):

11. Diagnosis according to CCv3.0 ☐ Normal ☐ Achalasia I ☐ Achalasia II ☐ Achalasia III
☐ EGJ Outflow Obstruction ☐ Fragmented Peristalsis ☐ Absent Contractility
☐ Distal Esophageal Spasm ☐ Ineffective Esophageal Motility
☐ Hypercontractile (Jackhammer) Esophagus ☐ Unable to Determine

REDCap

Figure 1.

Example of an HRM case. First, participants were instructed to thermally compensate and hide impedance. Next participants assessed the: A) Baseline phase referenced to gastric pressure. Numbers denote individual skills: (1) – Upper esophageal sphincter location, (2) & (3) – Esophagogastric Junction Morphology (EGJ) and presence of hiatal hernia, (4) – Baseline Expiratory EGJ Pressure, (5) – Pressure Inversion Point. Next participants assessed: B) Swallow #1 referenced to atmospheric pressure. (6) – Integrated Relaxation Pressure, (7) – Distal Contractile Integral, (8) – Distal Latency, (9) – Peristaltic Integrity, (10) – Pressurization Pattern. (8-10) not depicted in this window. Next, participants assessed all 10 supine wet swallows to determine a final diagnosis (11). Finally, they provided their confidence level for this individual case (12). C) Participants submitted their interpretations through REDCap.

Figure 2. CUSUM Learning Curves in Primary Analysis (Considering Acceptable Diagnostic Accuracy as 90% Correct)

A) Learning curves of all trainees. B) Five (25%) trainees achieved competency (final CUSUM score < 0.03 and not reset within the last 10 cases). C) Four (20%) trainees likely achieved competency (final CUSUM score between 0.03 and 1.03, and not reset within the last 10 cases). D) Eleven (55%) trainees did not achieve competency (final CUSUM score > 1.03 and/or reset within the last 10 cases).

Table 1
Trainee Characteristics

Baseline Variables (N=20)	n (%)
Age, Mean (SD), <i>years</i>	32.2 (2.5)
Male Sex	17 (85%)
Year in fellowship training:	
First	6 (30%)
Second	7 (35%)
Third	7 (35%)
Clinical training track	19 (95%)
University training setting	18 (90%)
Community or County training setting	2 (10%)
No previous manometry experience	18 (90%)
Previously analyzed 1-5 manometry studies	2 (10%)
Confidence level prior to study onset (Scale of 1-10), Median (range)	1 (1,5)

Table 2
Competency By Individual HRM Skill and Agreement in Competency Between Skill and Overall Diagnostic Accuracy

HRM Skill	Competency Achieved	Competency Likely Achieved	Competency Not Achieved	Kappa with Diagnostic Accuracy (95% CI)*
UES Location	19 (95%)	0 (0%)	1 (5%)	0.1 (-0.1–0.2)
EGJ Morphology	0 (0%)	0 (0%)	20 (100%)	NA
Presence of Hiatal Hernia	0 (0%)	0 (0%)	20 (100%)	NA
Basal EGJ Pressure	1 (5%)	0 (0%)	19 (95%)	0.1 (-0.1-0.3)
PIP Location	4 (20%)	3 (15%)	13 (65%)	0.0 (-0.5-0.4)
IRP	3 (15%)	2 (10%)	15 (75%)	0.2 (-0.2-0.6)
DCI	6 (30%)	3 (15%)	11 (55%)	0.4 (0.0-0.8)
DL	9 (45%)	5 (25%)	6 (30%)	0.3 (0.0-0.7)
Peristaltic Integrity	8 (40%)	4 (20%)	8 (40%)	0.1 (-0.3-0.5)
Pressurization Pattern	8 (40%)	4 (20%)	8 (40%)	0.1 (-0.3-0.5)
Overall Diagnosis	5 (25%)	4 (20%)	11 (55%)	NA

* Agreement assessed in groups of competency achieved or likely achieved versus competency not achieved. Kappa statistics do not indicate good overall agreement in general.²⁹ High-resolution manometry (HRM); Confidence Interval (CI); Upper esophageal sphincter (UES); Esophagogastric junction (EGJ); Pressure inversion point (PIP); Integrated relaxation pressure (IRP); Distal contractile integral (DCI); Distal latency (DL); Chicago Classification version 3.0 (CCv3.0)