Characterizing the Learning Curve and Accuracy of Wireless Capsule Endoscopy Interpretation amongst Digestive Healthcare Practitioners

Vinay Chandrasekhara¹, Russell T Shinohara², Kerry B Dunbar³, Monica Riegert⁴, Eric F Tomakin⁴ and Gerard E Mullin⁴

¹ University of Pennsylvania, Perelman School of Medicine, Gastroenterology Division Philadelphia, PA, USA
² University of Pennsylvania, Perelman School of Medicine, Department of Biostatistics & Epidemiology Philadelphia, PA, USA
³ University of Texas, Southwestern Medical School, Division of Gastroenterology & Hepatology Dallas, TX, USA
⁴ Johns Hopkins University, Division of Gastroenterology & Hepatology Baltimore, MD, United States

Abstract

Goals: To prospectively evaluate the accuracy and learning curve of trainees to interpret wireless capsule endoscopy (WCE).

Background: Trainees and mid-level providers are often used as first-line readers to earmark WCE images that are subsequently reviewed by an attending gastroenterologist. This practice has been demonstrated to be cost-effective and highly accurate, but the learning curve of trainees and physician extenders has not been well described.

Study: Each WCE trainee individually reviewed and interpreted 48 full-length de-identified WCE studies and recorded significant findings, time required to interpret each study, and confidence of the interpretation. After every 10 studies, each trainee reviewed WCE findings with the expert. The first 10 studies were considered the training period with the next 19 studies considered study period 1 and last 19 studies considered study period 2. Reading times were compared using Wilcoxon rank-sum tests.

Results: The median time for WCE interpretation improved significantly from the training period to the first study period for each trainee; GI fellow (41 min vs. 21 min, p<0.002), CRNP fellow (67 min vs. 33 min, p<0.002), RN (82 min vs. 30 min, p<0.001). Median reading times appeared to plateau for the GI fellow and CRNP fellow after 10 and 20-25 studies, respectively. This finding suggests that additional studies are unlikely to improve reading efficiency. The GI nurse never plateaus during the study. Confidence improved for all trainees during the study; however, in the final study period, the GI fellow reported higher confidence than the CRNP fellow, who reported higher confidence than the GI nurse. After the training period of 10 capsule studies, the MD fellow was found to have a higher sensitivity (89%), specificity (95%) and accuracy (92%) compared to other trainees.

Conclusion: Reading time, self-reported confidence, and accuracy appear to be good markers for assessing learning. Twenty-five capsule studies appear to be sufficient for trainees enrolled in a GI fellowship program who are actively performing endoscopy; however physician extenders who do not perform endoscopy appear to require interpretation of more than 25 WCE studies before adequate proficiency is achieved.

Keywords: Wireless capsule endoscopy; Endoscopic training; Learning curve

Received: November 03, 2015; Accepted: January 27 2016; Published: February 05, 2016
Introduction

Wireless capsule endoscopy (WCE) has been embraced as an important diagnostic tool for small bowel disorders since it was first introduced and approved by the FDA in 2001 [1, 2]. This single use, pill-sized device has provided a novel method of directly visualizing the small bowel, without the risks typically associated with endoscopy or the radiation risks associated with imaging studies. Consequently, capsule endoscopy has gained popularity amongst both clinicians and patients, and has continued to expand both technologically and clinically since its introduction.

Several studies have demonstrated that capsule endoscopy has a higher diagnostic yield in identifying small bowel lesions compared to push enteroscopy and radiology studies including barium x-rays, computed tomography and magnetic resonance imaging [3-7]. The primary indication for WCE has been for evaluation of obscure gastrointestinal bleeding, but the indications for WCE have expanded to include evaluation of suspected Crohn’s disease, suspected or refractory or malabsorptive syndromes such as celiac disease and surveillance or evaluation of patients with small bowel tumors or hereditary polyposis syndromes [8].

With images captured at a rate of 2 images per second and a battery life of 8 hours, a single capsule study generates more than 50,000 images. Reading and interpretation of WCE images from a single study can vary between 30 and 120 minutes [9, 10]. As the interpretation of WCE studies can be time-consuming, GI fellows and mid-level providers are routinely used as first-line readers to interpret WCE studies. Consequently, capsule endoscopy has gained popularity amongst both clinicians and patients, and has continued to expand both technologically and clinically since its introduction.

Materials and Method

Procedure details

All patients undergoing WCE for various indications were instructed to take a bowel preparation followed by a 12 hour overnight fast prior to the exam. All wireless capsule endoscopy studies were performed using the PillCam SB 2™ Capsule (Given Imaging Ltd., Yokneam, Israel). Studies were de-identified for subsequent review by the study participants. The study was approved by the Institutional Review Board.

Capsule interpretation

Fifty consecutive capsule studies were individually reviewed and interpreted by study participants. Each reader was responsible for interpreting the entire de-identified study, approximately 8 hours in duration. In this study, three WCE trainees with no capsule endoscopy experience: a second-year GI fellow with endoscopy experience of approximately 100 colonoscopies and 200 upper endoscopies, a GI nurse practitioner fellow with minimal endoscopy experience but undergoing concomitant training with colonoscopy, and a senior GI nurse without personal experience performing endoscopy were compared to a capsule endoscopy expert who has read over 1000 capsule endoscopy studies. Each trainee was blinded to the interpretation of the capsule endoscopy expert and the interpretation of other study participants. WCE trainees were aware of the indication for each study, but were blinded to findings from prior endoscopic interventions. None of the trainees had attended a formal WCE training course. For each study the reader recorded the amount of time required for WCE review or interpretation, were used in this study. Data were analyzed after a set number of studies to determine the learning curve of WCE interpretation with the hope of trying to identify a minimum number of WCE studies that had to be reviewed before one was deemed proficient with interpretation of capsule endoscopy. We prospectively characterized the learning curve of gastroenterology trainees including a GI fellow, GI CRNP fellow as well as a senior GI nurse by measuring the time required for WCE interpretation, analyzing self-reported confidence measurements and assessing the sensitivity, specificity and accuracy of WCE interpretation.

This study prospectively evaluates trainees interpreting WCE studies in order to characterize the learning curve. Practitioners with varying levels of GI training, but no previous experience with WCE review or interpretation, were used in this study. Data were analyzed after a set number of studies to determine the learning curve of WCE interpretation with the hope of trying to identify a minimum number of WCE studies that had to be reviewed before one was deemed proficient with interpretation of capsule endoscopy. We prospectively characterized the learning curve of gastroenterology trainees including a GI fellow, GI CRNP fellow as well as a senior GI nurse by measuring the time required for WCE interpretation, analyzing self-reported confidence measurements and assessing the sensitivity, specificity and accuracy of WCE interpretation.

Despite the increase in use of capsule endoscopy, there is very little data to guide the ideal training method for reading and interpreting WCE studies [18]. Prior studies have demonstrated that medical trainees and non-gastroenterologists can accurately interpret capsule endoscopy studies, yet the number of capsule studies that one must review prior to achieving proficiency is unknown [10-13]. One study of a single endoscopy nurse demonstrated consistent agreement with physician interpreters after interpretation of 80 WCE studies [19]. At present, there are no standardized WCE training requirements or assessment tools for all healthcare practitioners. One society recommends that GI physicians must either receive formal training in capsule endoscopy during GI fellowship or complete a hands-on course with a minimum of 8 hours CME credit endorsed by a national or international GI society followed by review of the first 10 capsule studies by a credentialed capsule endoscopist [20]. Current American guidelines for endoscopic training in routine procedures within a fellowship define 25 capsule endoscopy studies as a threshold for assessing competence; however, this recommendation was not based on prior data [21]. Another guideline acknowledges non-physician personnel are able to locate significant endoscopic abnormalities with WCE and states that intensive training is required, but does not go so far as to state the minimum number of studies required or how training should be conducted [22].
Statistical analysis

Statistical calculations were performed in R (Version 2.15.1, R Foundation for Statistical Computing, Vienna, Austria). Median reading times were compared using Wilcoxon rank-sum tests. To assess changes in reading time due to learning, the difference in reading time between each participant and the WCE expert for each study was modeling by nonlinear least squares fitting of a single-compartment monoexponential decay pharmacokinetic model [23].

Results

A total of 50 capsule studies were reviewed by each capsule reader. After all studies were interpreted by all readers with subsequent unblinding, it was determined that the independent technician who uploaded the capsule video images accidentally uploaded two studies twice. Therefore the redundant studies were excluded from the analysis. Therefore a total of 48 capsule studies were included in the analysis with the first 10 studies included in the "training period". The subsequent 38 studies were divided into two groups of 19 studies. Each capsule study was categorized into one of four indications: obscure GI bleeding, suspected Crohn’s disease, evaluation of small bowel polyps/surveillance for polyposis syndromes and suspected malabsorptive disorders. The indications did not appear to vary across the study period (Table 1). The most common indications for WCE were for suspected obscure GI bleeding and evaluation of Crohn’s disease. There was only one WCE performed for small bowel polyps in the entire study.

The median time required for interpretation of each study improved significantly from the training period to the first study period for each trainee; GI fellow (41 min vs. 21 min, p<0.002), CRNP fellow (67 min vs. 33 min, p<0.002), RN (82 min vs. 30 min, p<0.001) (Table 2). The median time for interpretation was not significantly different from study period 1 to study period 2 for each trainee. As expected, the expert reader’s average reading time did not significantly vary throughout the study. Inherent to WCE, some studies are more intensive to review based on a variety of factors including the number of significant findings, quality of the preparation and indication for the study. In an attempt to control for these factors, the reading times for each trainee were compared to the expert reader for each study and are summarized according to the study period in Table 3. When compared to the expert reader’s times, all trainees demonstrated continued improvement in the second study period compared to the first study period, but this difference was statistically different for the GI fellow (median difference 4 min in study period 1 vs. -4 min in study period 2, p=0.03) and for the CRNP fellow (median difference 13 min vs. 7 min, p=0.05) but only marginally significant for the GI nurse (median difference 14 min vs. 6 min, p=0.06). Pharmacokinetic modeling of this data demonstrates that the GI fellow and CRNP fellow reach an asymptote during the study period, suggesting that additional studies would not improve overall reading efficiency (Figure 1). The MD fellow reaches this asymptote at approximately 10 studies, whereas the CRNP fellow reaches this asymptote at roughly 20-25 studies. The RN does not approach a plateau after 48 studies, suggesting that interpreting additional capsule studies would be required in order to improve reading time.

Each WCE interpreter was asked to rate their own confidence level for each study. Confidence was recorded as “not confident”, “somewhat confident”, “moderately confident”, “very confident”, or “extremely confident”. During the study period, each trainee reported higher confidence levels during the second study period compared to the first study period (Figure 2). For the second study period, the GI fellow on average reported higher confidence levels than the CRNP fellow, rating half of the studies rated as “extremely confident” and the other half as “very confident”. The CRNP fellow reported higher confidence levels than the RN, with the majority of the studies in the second study period rated as “very confident”. The RN was still only “moderately confident” for approximately half of the studies in the second study period.

Every WCE reviewer documented all significant findings during each capsule study. The WCE expert identified significant pathology that accounted for the patient’s symptoms in 19 of the 38 studies (50%) during the study period. A positive study was identified in 10 of 14 WCE studies (71%) performed for obscure GI bleed, 8 of 16 studies (50%) performed for suspected Crohn’s disease, 1 of 1 studies performed for suspected small bowel polyps, and in 0 of 7 studies performed for suspected malabsorptive disorders. The presence or absence of significant findings was compared to findings recorded for each trainee in order to determine the overall sensitivity, specificity and accuracy of interpretation. During the study period, there were few WCE studies for suspected small bowel polyps [1] and none of the studies for suspected malabsorptive syndromes were positive; therefore the data were not specifically analyzed for these indications alone. Overall, the MD fellow was found to have a higher sensitivity (89%), specificity (95%), and accuracy (92%) compared to other trainees for the identification of clinically significant findings (Table 4). The GI CRNP fellow was found to have the highest accuracy for studies performed for suspected GI bleed (93%). The CRNP fellow and GI nurse were found to have a higher accuracy for studies performed for suspected Crohn’s disease (88% and 75%, respectively) compared to studies performed for suspected GI bleed (79% and 64%, respectively).

Discussion

The use of wireless capsule endoscopy for the evaluation of small bowel disease has dramatically increased over the past decade as the diagnostic yield of WCE has been demonstrated to be superior to alternate imaging modalities. Interpreting an 8-hour study with over 50,000 images can be time-consuming; therefore, many physicians utilize trainees and mid-level providers to mark images of concern that are later reviewed by an attending physician. Although this approach has been demonstrated to be cost-effective without sacrificing accuracy, the learning curve of trainees and physician extenders has not been well described. American guidelines for endoscopic training in routine procedures within a fellowship define 25 capsule endoscopy studies as a threshold for assessing competence [21]. However, these guidelines acknowledge that the minimal training required to competently perform capsule endoscopy of the small
intestine has not been evaluated. Another guideline suggests that GI physicians not formally trained in WCE during fellowship should receive training in a hands-on WCE course with the only the first 10 WCE studies reviewed by a credentialed capsule endoscopist [20]. A recent document on small bowel endoscopy core curriculum suggested a minimum number of 20 supervised procedures for those intending to practice WCE independently [24]. This issue has been further potentiated by the lack of structured training and varied requirements for credentialing at different institutions as well as ill-defined measurements for the assessment of proficiency. Rajan et al proposed a capsule competency test based on a combination of didactic questions, video clips, still images and one full length WCE study and found that fellows who had completed more than 20 prior WCE interpretations scored similarly to WCE staff, whereas those with less experience had significantly lower scores [25]. Based on these observations, the authors concluded that trainees should complete more than 20 WCE studies before assessing competence. Training and assessment for physician extenders has not been standardized or validated.

<table>
<thead>
<tr>
<th>Indications</th>
<th>Training Period (n=10)</th>
<th>Study Period 1 (n=19)</th>
<th>Study Period 2 (n=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obscure GI bleed</td>
<td>3</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Suspected Crohn’s Disease</td>
<td>5</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Small bowel polyps/Polyposis syndrome</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Suspected Malabsorptive Syndrome</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 1** Capsule study indications according to time period.

<table>
<thead>
<tr>
<th>WCE Interpreter</th>
<th>Training Period</th>
<th>Study Period 1</th>
<th>Study Period 2</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert</td>
<td>17</td>
<td>17</td>
<td>23</td>
<td>0.15</td>
</tr>
<tr>
<td>MD Fellow</td>
<td>41</td>
<td>21</td>
<td>19</td>
<td>0.27</td>
</tr>
<tr>
<td>CRNP Fellow</td>
<td>67</td>
<td>33</td>
<td>29</td>
<td>0.42</td>
</tr>
<tr>
<td>RN</td>
<td>82</td>
<td>30</td>
<td>30</td>
<td>0.05</td>
</tr>
</tbody>
</table>

**Table 2** Median capsule reading times (in minutes) for each individual according to time period. The differences in reading times for WCE trainees were statistically significantly improved from the training period to study period 1, but were not different for the trainees from study period 1 to study period 2.

<table>
<thead>
<tr>
<th>WCE Interpreter</th>
<th>Training Period</th>
<th>Study Period 1</th>
<th>Study Period 2</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD Fellow</td>
<td>22</td>
<td>4</td>
<td>-4</td>
<td>0.03</td>
</tr>
<tr>
<td>CRNP Fellow</td>
<td>42</td>
<td>13</td>
<td>7</td>
<td>0.05</td>
</tr>
<tr>
<td>RN</td>
<td>56.5</td>
<td>14</td>
<td>6</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**Table 3** Median difference in capsule reading time (in minutes) compared to the capsule expert. The median difference in capsule reading time between study period 1 and 2 was statistically significant from the GI fellow, but not the nurse practitioner fellow or the GI nurse.

**Figure 1** Pharmacokinetic modeling of the median difference in reading time (in minutes) for each WCE trainee compared to the expert over the duration of the study. The GI fellow appears to plateau at 10 studies, the CRNP fellow plateaus at 20-25 studies, while the GI RN continues to improve reading time throughout the study.
This study was designed to characterize the learning curve of WCE trainees including a GI fellow, CRNP fellow and a senior GI nurse. Unlike other studies that have compared inter-observer variability of a GI nurse, physician extender or a trainee compared to an attending physician for WCE interpretation after a training period of 5-15 capsules, the intent of this study was to better understand how long that training period should be [11-13, 16, 17]. Measurements used to characterize the learning curve included time required for WCE interpretation, self-reported confidence levels and accuracy with the interpretation compared to the expert reader. The GI fellow appeared to reach a plateau for the time required for WCE interpretation after 10 studies, although self-reported confidence continued to improve with subsequent studies. After 25 WCE studies, all subsequent interpretations were rated as “very confident” or “extremely confident” by the GI fellow. After a training period of 10 WCE studies, the GI fellow had a high sensitivity (89%), specificity (95%), and accuracy (92%). The CRNP fellow reached a plateau for the time required for WCE interpretation after 20-25 WCE studies and self-reported confidence continued to improve with subsequent studies. After 25 WCE studies, all subsequent interpretations were rated as “very confident” or “extremely confident” by the CRNP fellow. The GI RN did not reach a plateau during the study period for the time required for WCE interpretation and 50% of the study interpretations were only “moderately confident” in the second study period. The GI nurse’s accuracy for WCE interpretation was only 68% after 10 training studies. This study validates that 25 capsule studies is sufficient for individuals enrolled in a GI fellowship prior to assessing proficiency. Some trainees enrolled in a GI fellowship can achieve proficiency in as few as 10 training capsule studies. Furthermore, physician extenders and GI nurses who do not perform endoscopy likely require interpretation of additional studies before proficiency can be achieved. The GI nurse had not reached a plateau with regard to reading time and confidence by the end of 48 studies, which suggests that additional studies would continue to improve these variables. Prior endoscopic experience appears to enable trainees to more reliably interpret imaging findings and improves the learning curve for WCE interpretation. These conclusions are similar to those drawn from another study that compared gastroenterology trainees with endoscopy experience versus medical students; however, in that study individuals were required to interpret only 10 capsule studies, which is far too few to ascertain a true learning curve [10].

While many studies have proven that physician extenders and GI nurses are equally capable for interpreting and highlighting significant abnormalities with WCE studies for subsequent physician review, the data has been comparing these individuals after a non-standardized and often ill-defined training process. Few studies address the ideal method of training in WCE and

<table>
<thead>
<tr>
<th>SENSITIVITY (%)</th>
<th>All</th>
<th>GI Bleed</th>
<th>Suspected Crohn’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>GI Fellow</td>
<td>89 (76-100)</td>
<td>90 (71-100)</td>
<td>88 (65-100)</td>
</tr>
<tr>
<td>CRNP Fellow</td>
<td>79 (61-97)</td>
<td>70 (42-98)</td>
<td>88 (65-100)</td>
</tr>
<tr>
<td>RN</td>
<td>47 (25-70)</td>
<td>50 (19-81)</td>
<td>50 (19-81)</td>
</tr>
<tr>
<td>SPECIFICITY (%)</td>
<td>All</td>
<td>GI Bleed</td>
<td>Suspected Crohn’s</td>
</tr>
<tr>
<td>GI Fellow</td>
<td>95 (85-100)</td>
<td>100 (NA)</td>
<td>88 (65-100)</td>
</tr>
<tr>
<td>CRNP Fellow</td>
<td>84 (68-100)</td>
<td>100 (NA)</td>
<td>88 (65-100)</td>
</tr>
<tr>
<td>RN</td>
<td>89 (76-100)</td>
<td>100 (NA)</td>
<td>100 (NA)</td>
</tr>
<tr>
<td>ACCURACY (%)</td>
<td>All</td>
<td>GI Bleed</td>
<td>Suspected Crohn’s</td>
</tr>
<tr>
<td>GI Fellow</td>
<td>92 (84-100)</td>
<td>93 (79-100)</td>
<td>88 (71-100)</td>
</tr>
<tr>
<td>CRNP Fellow</td>
<td>82 (69-94)</td>
<td>79 (57-100)</td>
<td>88 (71-100)</td>
</tr>
<tr>
<td>RN</td>
<td>68 (54-83)</td>
<td>64 (39-89)</td>
<td>75 (54-96)</td>
</tr>
</tbody>
</table>

Table 4 Sensitivity, specificity, and accuracy of WCE interpretation for trainees after a training period of 10 WCE studies.

![Figure 2](image-url) Confidence level of each WCE interpreter during the first study period compared to the second study period.
assessment tools for evaluating the learning curve. Selectively reviewing edited video clips containing significant pathology can help improve lesion recognition skills, but reviewing a clip lasting several minutes is much different that reviewing an entire 8-hour long study [26]. Our study has several strengths. We characterized the learning curve by not only assessing accuracy of interpretation, but also measuring the time required for interpretation of each study and the self-reported confidence level of each reader. Every individual interprets studies at his or her own rate. To better characterize each individual’s learning curve, we analyzed how many cases an individual had to review before the trainee plateaued with reading speed. This study also reflected a “real-world” practice with consecutive studies including normal studies or those with subtle findings. The main weakness of our study was that there was only one individual in each WCE trainee category and the findings may not be generalizable to all individuals at that training level. In addition, the gold-standard comparison was the interpretation of the attending physician and not definitive pathology or findings validated with additional imaging modalities. While our expert reader has experience with over 1000 WCE studies, it is possible the expert may fail to correctly identify significant lesions [27]. One study suggested that experts with experience of more than 400 WCE tended to be more precise, suggesting continued learning even after interpreting many WCE studies [28]. Although not formally studied, the interpretation of multiple readers appeared to have a higher accuracy than any individual interpreter alone, including the expert reader. This study is only intended to address proficiency with WCE interpretation. For one to achieve true competency with WCE one must not only be skilled with the interpretation of the study, but must be knowledgeable of the indications for the study as well as the risks, benefits, alternative diagnostic modalities, and subsequent management recommendations based on WCE findings. Physicians are educated on these aspects during fellowship. Physician extenders and nurses not enrolled in a formal GI training program can best address these issues by attending a hands-on course with CME credit endorsed by a national or international GI society. These courses also highlight pathology that may not be commonly encountered or easily missed.

Wireless capsule endoscopy has proven to be a useful diagnostic modality for visualization of small bowel diseases. As the use of WCE continues to expand, it is important to properly train individuals and identify measurements that can assess one’s proficiency with WCE interpretation. Trainees including fellows, physician extenders and nurses can achieve proficiency; however, the learning curve for these individuals appears to be different. A minimum of 25 capsule studies, with each study read in its entirety, appears to be sufficient for trainees enrolled in a GI fellowship program who are actively performing endoscopy. Some GI trainees are able to achieve proficiency in as few as 10 studies. Physician extenders and nurses not enrolled in a formal GI training program who do not perform other endoscopic procedures and wish to pre-read WCE studies along with a GI physician should be encouraged to attend a capsule endoscopy course held by a GI society. Physician extenders and nurses appear to require interpretation of more than 25 WCE studies before adequate proficiency is achieved. For these individuals, it remains to be seen if reviewing additional selectively edited video clips or participating in a computer-based training module has a positive impact on the learning curve.
References


© Under License of Creative Commons Attribution 3.0 License