

Optimizing the Interpretation of Capsule Endoscopic Images: Shortsighted or Taking the Long View?

Anastasios Koulaouzidis · Ervin Toth

Received: 16 February 2015 / Accepted: 18 February 2015 / Published online: 28 February 2015
© Springer Science+Business Media New York 2015

Capsule endoscopy (CE) has become the preferred method of investigating the small intestine [1, 2]. Although capsule administration is a relatively straightforward task in the absence of medical complications, reading of CE videos is laborious, with outcome and reporting dependent not only on reviewer attentiveness and expertise but also on several other specific perceptual and interpretational factors [3, 4]. When viewing a soporific stream of often repetitive non-distinct images in a quiet, dark room, a significant risk of loss of concentration can lead to inaccuracy of reported findings [4, 5]. Nevertheless, practicing gastroenterologists may be considered adequately trained in CE reporting after a short, 1-day training program [6]. Moreover, formal training in CE during gastrointestinal (GI) fellowship, defined only loosely, includes completion of a hands-on course with a minimum of 8 h of continuing medical education (CME) credit, followed by review of CE studies by a credentialed capsule endoscopist [6]. There is currently no standardization of national or international training programs, although guidelines are being developed [7]. Furthermore, only limited evidence-based information on the optimal reading mode of CE review is currently available [3, 4, 8].

The past has taught us a great deal about medical image perception, not only in “classical” image-based specialties such as radiology and pathology, but also in other clinical specialties

that use imaging technology—such as gastroenterology, laparoscopic surgery, or dermatology [9–11]. Medical images and videos represent a significant source of information that aid clinicians with diagnostic and therapeutic decisions [11]. Yet, the correct interpretation of medical images relies on a host of factors, with significant health and medicolegal issues accruing from their inaccurate interpretation, which consists of two basic processes—visual perception (image inspection) and cognition (rendering an interpretation) [10, 11]. The use and development of computer-based models to predict human performance has also been a topic of interest for which a paucity of perception-oriented research exists, yet the opportunities abound.

The American Society of Gastrointestinal Endoscopy (ASGE) recommends a minimum number of 20 supervised procedures to provide adequate experience for those intending to practice CE independently [6]. Commercially available software provides a diverse range of viewing modes (VM) and frame rates (FR), in addition to other image enhancement tools such as digital chromoendoscopy [3, 12, 13]. No consensus has been reached for the latter technique according to a number of studies, its optimal mode of application yet to be determined [2, 3]. The use of differing VM has been, to date, the subject of only two studies [14, 15]. In the most recent large cohort study, Zheng et al. [15] reported that the low lesion detection rates observed were not influenced by increasing CE experience. Detection rates are significantly higher when reading in single VM/FR15 (single screen with FR 15/s) and quad VM/FR20 (four screens with FR 20/s) compared with reading in single View/FR25 (single screen with FR 25/s). Increasing viewing speed in quad VM from FR20 to FR30 appears to have no significant effect on detection ability. Therefore, the investigators suggested that quality control measures to compare and improve lesion detection rates need further study.

A. Koulaouzidis (✉)
Endoscopy Unit, Royal Infirmary of Edinburgh, 51 Little France
Crescent, Edinburgh EH16 4SA, Scotland, UK
e-mail: Tassos.Koulaouzidis@luht.scot.nhs.uk
URL: <http://www.drkoulaouzidis.com>

E. Toth
Endoscopy Unit, Skåne University Hospital, Malmö, Sweden

In this issue of *Digestive Diseases and Sciences*, Nakamura et al. [14] used a standardized, single-type lesion model to explore the relationship between VM and FR on lesion detection, determining the effect of these settings on CE reading time. They randomly selected 10 complete (to cecum) CE videos, obtained with PillCam[®] SB2, recording “real time,” i.e., the actual time the video was playing, without interruptions, from the point of duodenal entry to cecal exit, with 11 different combinations of VM and FR. Thereafter, a single CE video clip of excellent image clarity, comprised 60 positive images of small bowel angioectasias, was selected. To examine the effect of experience, the video was then read by six CE reviewers (three novices and three experienced) using nine combinations of VM and FR. Videos were presented to each reader in randomized order to minimize the risk of lesion recall. Readers were asked to count each positive image when an angioectasia was seen using a manual counter, generating a maximum number of positive images (MPIs). At the same time, the reading time for each combination of VM and FR was recorded. The authors reported that the optimal combination (for a high MPI) was FR10 using dual VM or quad VM. The outcome measure used was the maximum number of frames on which one or more angioectasias was seen. Increasing FR10–FR15 shortened reading times by 33 %, reduced mean MPI of 25–28 %. Altering VM had no effect on the reading time for any given FR.

Naturally, there are certain limitations to the study design; for instance, investigators knew that angioectasias were the only lesions observable on the video clips, which does not accurately reflect clinical practice, since this pre-knowledge likely increased targeted and focused awareness which in turn increased the detection rate. Furthermore, the uninterrupted viewing of the CE video with the concurrent use of a manual counter seems unwieldy. Angioectasias are one of the most recognizable types of GI tract lesion [14]; it is likely that non-angioectatic, more subtle pathology would have been less identifiable, particularly by novices, which could lower detection rates for any given VM and FR than was reported. Interwoven in this study design is the lack of information on the detection ability in real-life circumstances, i.e., when the video clarity is suboptimal or during rapid transit of the capsule within the GI tract. To an extent, the authors attempted to offset this by eliminating access to the stop-and-roll and to-and-fro functions of the reviewing software.

For those of us who use CE regularly in our practice, it comes as no surprise that Nakamura et al. [14] confirmed the findings from Zheng et al. [15] study, i.e., that experience contributes little to lesion detection. In endoscopy, as in life, detection skills are related to attentiveness and awareness. It is in the interpretation of findings that

expertise comes into play. Therefore, the single lesion model finds us in agreement. Still, the use of only one video clip, no matter how randomized the review, can be a source of bias. Perhaps further studies should investigate the use of digital chromoendoscopy under optimal reviewing conditions, with a larger number of video clips and lesion types.

Perhaps one of the main reasons that standardization of review has not been formally adopted is the aforementioned wide variability of trainee exposure combined with the multiplicity of reading modes and instrumentation. Again, the purpose is not to standardize reviewer-based protocols but to develop advanced and controllable CE platforms and software algorithms that can reliably detect and characterize lesions and automatically provide diagnosis [5], not unlike the automated interpretational systems that have revolutionized cardiac arrhythmia detection [16]. Although the latter may seem a tad futuristic and unattainable, we should not forget the origins of wireless endoscopy, which was developed in the same manner.

Conflict of interest A.K. has received research support from Given Imaging and SynMed UK, lecture honoraria from Dr. Falk Pharma UK, and travel support from Abbott, Dr. Falk Pharma UK, Almirall, and MSD. E.T. declares no competing interests.

References

1. Van de Bruaene C, De Looze D, Hindryckx P. Small bowel capsule endoscopy: where are we after almost 15 years of use? *World J Gastrointest Endosc.* 2015;7:13–36.
2. Koulaouzidis A, Rondonotti E, Karargyris A. Small-bowel capsule endoscopy: a ten-point contemporary review. *World J Gastroenterol.* 2013;19:3726–3746.
3. Koulaouzidis A, Iakovidis DK, Karargyris A, Plevris JN. Optimizing lesion detection in small-bowel capsule endoscopy: from present problems to future solutions. *Expert Rev Gastroenterol Hepatol.* 2015;9:217–235.
4. Lo SK. How should we do capsule reading? *Tech Gastrointest Endosc.* 2006;8:146–148.
5. Iakovidis DK, Koulaouzidis A. Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software. *Gastrointest Endosc.* 2014;80:877–883.
6. ASGE Training Committee 2011–2012, Rajan EA, Pais SA, et al. Small-bowel endoscopy core curriculum. *Gastrointest Endosc.* 2013;77:1–6.
7. Rajan E, et al. Training in small-bowel capsule endoscopy: assessing and defining competency. *Gastrointest Endosc.* 2013; 78:617–622.
8. Lewis BS. How to read wireless capsule endoscopic images: tips of the trade. *Gastrointest Endosc Clin N Am.* 2004;14:11–16.
9. Krupinski EA, Chao J, Hofmann-Wellenhof R, Morrison L, Curiel-Lewandrowski C. Understanding visual search patterns of dermatologists assessing pigmented skin lesions before and after online training. *J Digit Imaging.* 2014;27:779–785.
10. Kundel HL, Nodine CF, Krupinski EA, Mello-Thoms C. Using gaze-tracking data and mixture distribution analysis to support a

- holistic model for the detection of cancers on mammograms. *Acad Radiol*. 2008;15:881–886.
11. Krupinski EA, Graham AR, Weinstein RS. Characterizing the development of visual search expertise in pathology residents viewing whole slide images. *Hum Pathol*. 2013;44:357–364.
 12. Iakovidis DK, Koulaouzidis A. Software for enhanced video capsule endoscopy: challenges for essential progress. *Nat Rev Gastroenterol Hepatol*. 2015. doi:[10.1038/nrgastro.2015.13](https://doi.org/10.1038/nrgastro.2015.13).
 13. Krystallis C, Koulaouzidis A, Douglas S, Plevris JN. Chromoendoscopy in small bowel capsule endoscopy: blue mode or fuji intelligent colour enhancement? *Dig Liver Dis*. 2011;43:953–957.
 14. Nakamura M, Murino A, O'Rourke A, Fraser C. A critical analysis of the effect of view mode and frame rate on reading time and lesion detection during capsule endoscopy. *Dig Dis Sci*. (Epub ahead of print). doi:[10.1007/s10620-014-3496-5](https://doi.org/10.1007/s10620-014-3496-5).
 15. Zheng Y, Hawkins L, Wolff J, Goloubeva O, Goldberg E. Detection of lesions during capsule endoscopy: physician performance is disappointing. *Am J Gastroenterol*. 2012;107:554–560.
 16. Liu SH, Cheng DC, Lin CM. Arrhythmia identification with two-lead electrocardiograms using artificial neural networks and support vector machines for a portable ECG monitor system. *Sensors (Basel)*. 2013;13:813–828.