Published in final edited form as:

J Surg Res. 2014 November ; 192(1): 62–67. doi:10.1016/j.jss.2014.05.056.

Cumulative Sum: a proficiency metric for basic endoscopic training

Yinin Hu, M.D.¹, Joshua S. Jolissaint, B.A.², Adriana Ramirez, M.D.¹, Ryan Gordon, B.S.², Zequan Yang, M.D.¹, and Robert G. Sawyer, M.D.¹

¹University of Virginia, Department of Surgery, Charlottesville, Virginia

²University of Virginia School of Medicine, Charlottesville, Virginia

Abstract

Background—As work hour restrictions increasingly limit some operative experiences, personalized evaluative methods are needed. We prospectively assessed the value of cumulative sum (Cusum) to measure proficiency with percutaneous endoscopic gastrostomy (PEG) among surgical trainees.

Materials and Methods—Nine post-graduate year one surgery residents each underwent a one-month rotation dedicated to endoscopy. Procedure durations for all PEG insertions were recorded prospectively. Criteria for task failure included need for attending take-over or procedure duration greater than 10 minutes. Cusum parameters were defined *a priori*, with acceptable and unacceptable failure rates of 5% and 15%, respectively. Concurrently, expert endoscopists blinded to Cusum results evaluated trainee proficiency weekly using a multi-category, five-point Likert scale survey.

Results—Nine surgical residents performed an average of 21 PEG's each. Expert evaluations and Cusum analyses identified 8 and 7 participants who attained proficiency after a median of 11.5 and 12 cases, respectively. For 4 of the residents who achieved proficiency by Cusum criteria, eventual relapses to inadequate performance were identified. These relapses were not detected by expert evaluation. Six participants who attained proficiency by both metrics performed a combined 32 superfluous cases, which could have been redistributed to poor-performing trainees.

Conclusions—Although lacking the granular insight of expert evaluations, Cusum analysis is more sensitive to relapses of sub-proficient performance. Adding Cusum analysis to expert

© 2014 Elsevier Inc. All rights reserved.

Correspondence Author: Robert G. Sawyer, MD, P.O. Box 800709, Charlottesville, VA 22908-0709, Fax: (434) 924-2260
rws2k@virginia.edu Phone: (434) 982-1632.

Author Contributions: Hu, Y: study conception/design, data analysis/interpretation, manuscript writing and approval Jolissaint, J: data collection/analysis, manuscript writing, critical revision and approval Ramirez, A: data collection/analysis, manuscript critical revision and approval Gordn, R: data collection, manuscript critical revision and approval Yang, Z: study design, data collection, manuscript critical revision and approval Sawyer, R: study conception/design, data collection, manuscript critical revision and approval

Disclosures: The authors report no proprietary or commercial interest in any product mentioned or concept discussed in this article.

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.

evaluations can provide longitudinal, formative feedback and promote efficient redistribution of operative experiences.

Keywords

cumulative sum; learning curve; upper endoscopy; percutaneous gastrostomy; resident training; evaluation

Introduction

Because flexible endoscopy is often an indispensable tool within the practicing general surgeon's armamentarium [1], endoscopic training during surgical residency is invaluable [2]. Accordingly, in 2006 the Residency Review Committee (RRC) elevated resident training criteria to include a minimum of 35 upper endoscopies and 50 colonoscopies [3]. Although most institutions can accommodate this volume, some rural training programs may struggle to match these criteria [4]. Fundamentally, volume-based criteria underappreciate variable learning rates and provide no assurance that competency persists after initial training [5, 6]. Over recent years, attention has shifted toward the development of competency-based metrics for endoscopic training [7], which commonly take the form of periodic expert-conducted surveys or checklists. To date, there exists no simple, objective measure of proficiency that can be easily applied on a case-by-case basis.

Cumulative sum (Cusum) is a mathematical inspection scheme first described by E.S. Page in 1954 as a method to monitor performance in the manufacturing industry [8]. It has since been implemented to assess technical training in a variety of procedures [5, 9, 10]. Percutaneous endoscopic gastrostomy (PEG) placement is one basic endoscopic intervention that lends well to Cusum analysis due to its comparatively limited case variability. Our objective was to pilot a proficiency metric for PEG placement that augments objective evaluations with Cusum analyses. We hypothesize that Cusum will capture variable learning speeds, identify performance relapses, and promote case redistribution based on skill level.

Materials and methods

We enrolled consenting post-graduate year one surgery residents who each subsequently completed a one-month rotation dedicated to endoscopy. Both preliminary and categorical surgery residents were eligible for participation. No participant had prior formal exposure to upper endoscopy, either in simulation settings or at the bedside. To ensure sufficient data for Cusum analysis, participants who performed fewer than ten PEGs during their rotation were excluded. Each PEG procedure employed a guidewire-assisted push-placement gastrostomy kit (Bard Access Systems), and involved upper endoscopy using a standard 9.4 mm flexible video gastroscope. Each procedure was concluded with a pyloric cannulation for training purposes. All procedures were directly supervised by surgical faculty. Cases with significantly aberrant oropharyngeal anatomy, equipment malfunction requiring replacement, and combined endoscopic procedures were excluded. Procedure duration and need for attending take-over were prospectively recorded. Procedure time started upon passing the endoscope beyond the incisors and finished after transection of the PEG's external port tubing.

Cusum analysis for technical training is described in detail by Bolsin and Colson [11]. Briefly, this method assigns point values to repetitive task attempts based on a binary set of outcomes (success vs failure), and sums this data to provide a longitudinal representation of the training process. Cusum employs several parameters defined by evaluators. The acceptable task failure rate (p_0) is the inherent failure rate of a proficient practitioner, while the unacceptable failure rate (p_1) is the failure rate above which a practitioner should not be allowed to perform a procedure. Thus, $p_1 - p_0$ represents the maximum acceptable level of human error [12]. Evaluators also define the allowable risks of falsely labeling a proficient practitioner as sub-par (type 1, α) and falsely certifying an inadequate practitioner (type 2, β). Using these parameters, point values for successful (-s) and unsuccessful (1-s) attempts and the corresponding unacceptable decision intervals (UDI, h_0) are determined by the following equation:

$$s = \ln[(1 - p_0)/(1 - p_1)] / \{ \ln(p_1/p_0) + \ln[(1 - p_0)/(1 - p_1)] \}$$

$$h_0 = \ln[(1 - \alpha)/\beta] / \{ \ln(p_1/p_0) + \ln[(1 - p_0)/(1 - p_1)] \}$$

A Cusum curve that trends upward and crosses a UDI indicates a non-proficient trainee, whereas a down-trending or level Cusum curve indicates proficiency.

For the present study, Cusum parameters were defined *a priori* by expert consensus, with criteria for failure defined as need for attending takeover or procedure duration greater than 10 minutes. A two-minute allowance was given for cases involving a jejunal-arm extension. The acceptable and unacceptable failure rates were set at 5% and 15%, respectively. Type 1 and type 2 error rates were both set to 0.1. To capture a running measure of performance, a moving window of five task attempts provided the data for Cusum calculations. Thus, each participant underwent a series of consecutive, “windowed” Cusum analyses beginning from the fifth PEG attempt (Figure 1), with each analysis incorporating performances from the five most recent cases.

Within our department, PEGs are predominantly performed by one of three expert endoscopists, all of whom use the same procedural methodology. Each week, the expert endoscopist assigned to oversee PEG procedures during the preceding week completed a multi-category Likert-scale objective proficiency assessment (Table 1). Scoring was based on the Global Assessment of Gastrointestinal Endoscopic Skills (GAGES-UE), modified for PEG insertion. Although experienced endoscopists on average score 18.5 on the GAGES-UE [13], due to the lower complexity of the PEG procedure, we defined proficiency by a minimum score of 20. With each evaluation, expert endoscopists were aware of the number of weeks that a given participant has been active within the endoscopy rotation. However, evaluators were blinded to participants' Cusum results at the time of evaluation, while participants were blinded to both Cusum and expert survey results. Cusum and expert evaluation metrics were considered to be in agreement if more than half of the Cusum analyses during the week preceding an expert survey were in concordance with that survey's result (proficient vs not proficient).

To gain insight into how a proficiency-driven protocol may affect operative assignments, we sought to determine the number of superfluous, “redistributable” cases that were performed by participants who met proficiency criteria for both metrics prior to concluding their endoscopy rotation. In order for a case to be considered redistributable, participants must be proficient by expert evaluation and concurrently possess three consecutive case-specific windowed Cusum analyses that show proficient status. The University of Virginia Institutional Review Board approved this study (IRB-SBS protocol #2012-0387).

Results

Between December 2012 and November 2013, 12 study participants were enrolled in the study. Three participants performed fewer than 10 PEGs each due to fluctuations in operative volume. The 9 remaining participants performed a total of 207 PEG insertions satisfying inclusion criteria. Procedure duration was not recorded for 9 PEGs, and 7 involved extended interruptions for sedation titration. The remaining 191 cases were incorporated into individualized windowed Cusum analyses. Participants performed a median of 21 cases (Interquartile range, IQR, 15-27). By expert evaluation, 89% (8/9) of participants achieved proficiency after a median of 11.5 cases (IQR 9-15), and each proficient participant maintained proficiency for all subsequent evaluations (Figure 2). Windowed Cusum analyses were completed for all participants beginning from the fifth PEG attempt. Seventy-eight percent (7/9) of participants achieved proficient status by Cusum analysis after a median of 12 cases (IQR 10-12.5). Sixty-seven percent (6/9) of participants were proficient by both metrics upon completion of their respective endoscopy rotations.

Expert evaluations and windowed Cusum analyses correlated well during the first two weeks of each rotation, then diverged as the rotations progressed (Table 2). Windowed Cusum analyses detected periods of inadequate performance after initial proficiency in 44% (4/9) of participants (relapses). These relapses all occurred on or after the third week of each endoscopy rotation. For example, Participant #2 initially attained proficiency by Cusum criteria after 15 cases. However, he then crossed a decision interval after 22 cases (relapse), before regaining proficient status following case 26 (Figure 1). Six participants met both Cusum and expert evaluation proficiency criteria before completing their respective endoscopy rotations. This threshold was attained after between 18 and 21 procedures. A total of 32 redistributable cases were performed, ranging from 0 to 12 per participant (Figure 3). Thus, implementing a proficiency-based case redistribution policy would have augmented the endoscopic experience of the three participants that failed to achieve proficiency by 68%.

Discussion

The present study applies a case-by-case metric to track incremental changes in technical proficiency for an upper endoscopic procedure. Using the two performance-driven metrics of windowed Cusum analyses and periodic expert evaluations, we demonstrate substantial variability in rates of endoscopic skill acquisition among surgical trainees. While the two metrics correlate well early in the training experience, Cusum appears to be more sensitive

to sub-par performance as case volume accumulated. Combining both metrics, we identified 32 PEG cases performed by proficient trainees during the course of one year that could have been redistributed to enhance collective competency with this procedure.

Current opinions regarding the adequacy of endoscopic training during surgical residency are inconsistent. Despite an average case volume of 47 gastroscopies among graduating residents, 18% of surgery program directors continue to feel that endoscopy exposure is inadequate [4]. This viewpoint is seconded by the American Society for Gastrointestinal Endoscopy, who recommends 130 gastroscopies among gastroenterology fellows, but also emphasizes a shift toward validated competency thresholds [7, 14]. Toward this end, the Society of American Gastrointestinal and Endoscopic Surgeons piloted the Fundamentals of Endoscopic Surgery (FES), which includes a validated simulation-based assessment [15]. However, transferability of simulation to performance on live patients is incompletely proven in alimentary endoscopy. Although several studies report benefit from endoscopic simulation [16-18], these gains may diminish with clinical experience [19, 20]. The FES program further necessitates flexible endoscopy simulators, for which the fixed cost of \$50-100,000 may be prohibitive among smaller programs. Finally, just as the Fundamentals of Laparoscopic Surgery is not designed to displace clinical case volume during training, principles of credentialing suggest that simulation-based endoscopic training is most useful as a supplement to clinical experience [21].

Although attending takeover is a common criterion for task failure, procedure speed is rarely reported as a surrogate for technical ability. However, in our experience, factors contributing most toward prolonged PEG procedure time are inadequate grasp of gastroscope controls, unfamiliarity with oropharyngeal anatomy, and inadequate visual-spatial awareness. Moreover, prolonged operative time is one of the major roadblocks to resident instruction [22], providing incentive for adopting a metric that could target this handicap. By defining specific start and stop time points, we reduced the impact of variable contributors to procedure duration such as sedation time and patient positioning.

Objective expert evaluations are unsurpassed in granularity and clinical relevancy [23]. However, performing validated, multi-category surveys such as GAGES-UE on a case-by-case basis is resource-intensive. For this reason, expert evaluations are most commonly adopted as periodic feedback [7, 24]. Our results highlight a second, underappreciated weakness of objective expert evaluations: they may not be consistently objective. According to Cusum results, nearly half of our participants experienced periods of sub-par performance after initially acquiring proficiency. These periods were missed by expert surveys. While human evaluators retain a predilection to believe that increasing experience implies increasing skill, a quantitative metric like Cusum can act as a quality check on the objectivity of expert assessments. Further, Cusum analyses are cheap, have negligible impact on clinical proceedings, and offer real-time feedback for every attempted task. The moving window modification focuses analysis on current performance, and is an adaptation of the exponentially weighted moving average method previously employed to monitor performance in congenital cardiac surgery [25]. Due to these advantages, Cusum has been implemented for numerous quality-assurance and training scenarios in surgery [5, 9, 10, 12, 26, 27]. We did not formally evaluate reasons for technical relapse within our study.

However, expert evaluators have noted that, while some trainees appear to use a combination of visual and tactile cues to direct scope navigation, others rely more heavily on visualization alone. These predominantly sight-dependent trainees tended to strive for more meticulous visualization with increasing experience, leading to prolonged procedure durations. To more clearly delineate these intriguing patterns of learning will require larger, multi-institutional studies adopting Cusum-style monitoring tools.

We demonstrated that combining windowed Cusum analyses and expert evaluations could help redistribute cases from proficient trainees to those in need of further practice. Far from suggesting that trainees proficient in PEG insertion have mastered upper endoscopy, we instead recommend assigning these trainees to more complex endoscopic cases. Adopting performance-driven metrics encourages fast learners to be continuously challenged and allows slower learners to acquire additional case volume. Using the combined metrics, we identified 32 superlative PEG procedures performed by proficient trainees which could have augmented the practice volume of those participants who never acquired proficiency. Meanwhile, proficient trainees could have been assigned to more colonoscopies and diagnostic upper endoscopies. Of note, our data suggest that relapses to sub-par skill can occur following initial proficiency; thus, proficient trainees should intermittently undergo Cusum “follow-up” with additional PEG insertions at regular intervals during the course of residency training. While our institution benefits from a dedicated endoscopy rotation, this framework of case redistribution could prove even more valuable to programs that currently lack such a rotation [4, 28].

There are several important limitations within our study. As a comparator to the Cusum metric, we adopted a 25-point expert evaluation metric, modified from the GAGES-UE. Through this modification, the validity of our metric may be weakened. There are also limitations inherent to the Cusum tool. In order for this binary metric to be successful at measuring proficiency, the target procedure must have consistent criteria for success. Percutaneous endoscopic gastrostomy lends well to Cusum analysis because it has relatively few variable components after patient sedation and positioning. In its present form, our Cusum metric does not account for clinical outcomes such as reoperation, infection, and other complications. Accounting for these outcomes in a post hoc manner may be useful, but may also redirect Cusum's role toward quality improvement rather than immediate feedback. More importantly, Cusum requires the specification of parameters *a priori*. These parameters, including criteria for success and failure and values for acceptable and unacceptable failure rates, require careful validation. Our goal is not to disseminate institutional Cusum parameters for the PEG procedure, but to demonstrate the potential utility for a case-specific proficiency metric. In light of this, our results highlight the need for multi-institutional research to further develop Cusum as a validated endoscopic training tool.

Conclusion

While some endoscopic trainees attain proficiency with PEG insertion after 18-21 cases, others require additional experience. Combining periodic expert evaluations with windowed Cusum analyses creates a composite proficiency metric that is objective, pragmatic and

clinically-relevant. Construct validation of Cusum parameters through multi-institutional investigation is necessary to effectively promote the redistribution of endoscopic cases among surgical trainees.

Acknowledgments

None

References

1. Nimeri AA, Hussein SA, Panzeter E, McNeill J, Gusz J, Chen PM, Yuh JN, Marks JM. The economic impact of incorporating flexible endoscopy into a community general surgery practice. *Surg Endosc.* 2005; 19:702–704. [PubMed: 15759183]
2. Vitale GC, Davis BR, Tran TC. The advancing art and science of endoscopy. *Am J Surg.* 2005; 190:228–233. [PubMed: 16023436]
3. Britt LD, Richardson JD. Residency review committee for surgery: an update. *Arch Surg.* 2007; 142:573–575. [PubMed: 17576896]
4. Subhas G, Gupta A, Mittal VK. Necessity for improvement in endoscopy training during surgical residency. *Am J Surg.* 2010; 199:331–4. discussion 334–5. [PubMed: 20226905]
5. Hu Y, Puri V, Crabtree TD, Kreisel D, Krupnick AS, Patterson AG, Meyers BF. Attaining proficiency with endobronchial ultrasound-guided transbronchial needle aspiration. *J Thorac Cardiovasc Surg.* 2013; 146:1387–1392. e1. [PubMed: 24075565]
6. Jowell PS, Baillie J, Branch MS, Affronti J, Browning CL, Bute BP. Quantitative assessment of procedural competence. A prospective study of training in endoscopic retrograde cholangiopancreatography. *Ann Intern Med.* 1996; 125:983–989. [PubMed: 8967710]
7. ASGE Training Committee. ASGE's assessment of competency in endoscopy evaluation tools for colonoscopy and EGD. *Gastrointest Endosc.* 2013 article in print.
8. Page E. Continuous Inspection Schemes. *Biometrika.* 1954; 41:100–115.
9. East JM, Valentine CS, Kanchev E, Blake GO. Sentinel lymph node biopsy for breast cancer using methylene blue dye manifests a short learning curve among experienced surgeons: a prospective tabular cumulative sum (CUSUM) analysis. *BMC Surg.* 2009; 9:2. [PubMed: 19173714]
10. Naik VN, Devito I, Halpern SH. Cusum analysis is a useful tool to assess resident proficiency at insertion of labour epidurals. *Can J Anaesth.* 2003; 50:694–698. [PubMed: 12944444]
11. Bolsin S, Colson M. The use of the Cusum technique in the assessment of trainee competence in new procedures. *Int J Qual Health Care.* 2000; 12:433–438. [PubMed: 11079224]
12. Blackstone EH. Monitoring surgical performance. *J Thorac Cardiovasc Surg.* 2004; 128:807–810. [PubMed: 15573062]
13. Vassiliou MC, Kaneva PA, Poulou BK, Dunkin BJ, Marks JM, Sadik R, Sroka G, Anvari M, Thaler K, Adrales GL, Hazey JW, Lightdale JR, Velanovich V, Swanstrom LL, Mellinger JD, Fried GM. Global Assessment of Gastrointestinal Endoscopic Skills (GAGES): a valid measurement tool for technical skills in flexible endoscopy. *Surg Endosc.* 2010; 24:1834–1841. [PubMed: 20112113]
14. Cass OW, Freeman ML, Cohen JR, et al. Acquisition of competency in endoscopic skills (ACES) during training: a multicenter study. *Gastrointest Endosc.* 1996; 43:308.
15. Vassiliou MC, Dunkin BJ, Fried GM, Mellinger JD, Trus T, Kaneva P, Lyons C, Korndorffer JR Jr, Ujiki M, Velanovich V, Kochman ML, Tsuda S, Martinez J, Scott DJ, Korus G, Park A, Marks JM. Fundamentals of endoscopic surgery: creation and validation of the hands-on test. *Surg Endosc.* 2013
16. Frezza EE, Halldorsson A, Griswold JA. Future directions in training surgical residents to perform endoscopic examinations. *Am Surg.* 2008; 74:187–188. [PubMed: 18306878]
17. Ende A, Zopf Y, Konturek P, Naegel A, Hahn EG, Matthes K, Maiss J. Strategies for training in diagnostic upper endoscopy: a prospective, randomized trial. *Gastrointest Endosc.* 2012; 75:254–260. [PubMed: 22153875]

18. Phillips MS, Marks JM. Overview of methods for flexible endoscopic training and description of a simple explant model. *Asian J Endosc Surg.* 2011; 4:45–52. [PubMed: 22776220]
19. Sedlack RE, Kolars JC. Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: results of a pilot study. *Am J Gastroenterol.* 2004; 99:33–37. [PubMed: 14687137]
20. Cohen J, Cohen SA, Vora KC, Xue X, Burdick JS, Bank S, Bini EJ, Bodenheimer H, Cerulli M, Gerdes H, Greenwald D, Gress F, Grosman I, Hawes R, Mullin G, Schnoll-Sussman F, Starnoli A, Stevens P, Tenner S, Villanueva G. Multicenter, randomized, controlled trial of virtual-reality simulator training in acquisition of competency in colonoscopy. *Gastrointest Endosc.* 2006; 64:361–368. [PubMed: 16923483]
21. Wexner SD, Eisen GM, Simmang C. Principles of privileging and credentialing for endoscopy and colonoscopy. *Surg Endosc.* 2002; 16:367–369. [PubMed: 11967713]
22. Bridges M, Diamond DL. The financial impact of teaching surgical residents in the operating room. *Am J Surg.* 1999; 177:28–32. [PubMed: 10037304]
23. Hope WW, Hooks WB 3rd, Kilbourne SN, Adams A, Kotwall CA, Clancy TV. Assessing resident performance and training of colonoscopy in a general surgery training program. *Surg Endosc.* 2013; 27:1706–1710. [PubMed: 23247738]
24. Vassiliou MC, Kaneva PA, Poulouse BK, Dunkin BJ, Marks JM, Sadik R, Sroka G, Anvari M, Thaler K, Adrales GL, Hazey JW, Lightdale JR, Velanovich V, Swanstrom LL, Mellinger JD, Fried GM. Global Assessment of Gastrointestinal Endoscopic Skills (GAGES): a valid measurement tool for technical skills in flexible endoscopy. *Surg Endosc.* 2010; 24:1834–1841. [PubMed: 20112113]
25. de Leval MR, Francois K, Bull C, Brawn W, Spiegelhalter D. Analysis of a cluster of surgical failures. Application to a series of neonatal arterial switch operations. *J Thorac Cardiovasc Surg.* 1994; 107:914–23. discussion 923-4. [PubMed: 8127123]
26. Kemp SV, El Batrawy SH, Harrison RN, Skwarski K, Munavvar M, Rosell A, Cusworth K, Shah PL. Learning curves for endobronchial ultrasound using cusum analysis. *Thorax.* 2010; 65:534–538. [PubMed: 20522852]
27. Kestin IG. A statistical approach to measuring the competence of anaesthetic trainees at practical procedures. *Br J Anaesth.* 1995; 75:805–809. [PubMed: 8672339]
28. Morales MP, Mancini GJ, Miedema BW, Rangnekar NJ, Koivunen DG, Ramshaw BJ, Eubanks WS, Stephenson HE. Integrated flexible endoscopy training during surgical residency. *Surg Endosc.* 2008; 22:2013–2017. [PubMed: 18297358]

Figure 1.

Cusum analysis of Participant #2. Overall Cusum curve (A). Windowed analyses demonstrate proficiency at case 15 (B) with relapse at case 22 (C), at which point the 1.82 unacceptable decision interval is crossed. Proficiency is regained at case 26 (D).

Figure 2. Trend in weekly expert evaluations. Participant scores are shown as median (circle), IQR₂₅₋₇₅ (box), and range (line). Significant improvement from initial performance is evidenced by Wilcoxon rank-sum test (p-values).

Figure 3.

Redistributable cases. Three participants (#4, 5, 6) failed to attain proficiency by either expert evaluation or Cusum criteria. Six participants (#1, 2, 3, 8, 9) performed a total of 32 cases after becoming proficient by both metrics.

Table 1

Expert evaluation criteria

Evaluative Criteria	Poor				Expert			
Intubation of the Esophagus	1	2	3	4	5			
Pyloric cannulation (scope navigation)	1	2	3	4	5			
Instrumentation	1	2	3	4	5			
Anatomic familiarity	1	2	3	4	5			
Procedural independence	1	2	3	4	5			
Attending Take-Over	Yes							No

Table 2
Rate of agreement between expert evaluations and Cusum analyses

	Rotation Progression				
	Week 1	Week 2	Week 3	Week 4	Week 5
Participants evaluated	9	9	9	8	4
Cusum proficient	1	2	4	5	3
Evaluation proficient	0	2	6	7	4
Agreement rate	89%	100%	78%	50%	75%