Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Cumulative sum: a proficiency metric for basic endoscopic training

Yinin Hu, MD,[a] Joshua S. Jolissaint, BA,[b] Adriana Ramirez, MD,[a] Ryan Gordon, BS,[b] Zequan Yang, MD,[a] and Robert G. Sawyer, MD[a,*]

[a] Department of Surgery, University of Virginia, Charlottesville, Virginia
[b] University of Virginia School of Medicine, Charlottesville, Virginia

ARTICLE INFO

Article history:
Received 27 March 2014
Received in revised form
29 April 2014
Accepted 16 May 2014
Available online 23 May 2014

Keywords:
Cumulative sum
Learning curve
Upper endoscopy
Percutaneous gastrostomy
Resident training
Evaluation

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 postgraduate year 1 surgery residents each underwent a 1-month rotation dedicated to endoscopy. Procedure durations for all PEG insertions were recorded prospectively. Criteria for task failure included need for attending takeover or procedure duration >10 min. 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 multicategory, five-point Likert-scale survey.

Results: Nine surgical residents performed an average of 21 PEGs each. Expert evaluations and Cusum analyses identified eight and seven participants who attained proficiency after a median of 11.5 and 12 cases, respectively. For four 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 subproficient performance. Adding Cusum analysis to expert evaluations can provide longitudinal, formative feedback and promote efficient redistribution of operative experiences.

1. 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 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].

* *Corresponding author*. Department of Surgery, University of Virginia, P.O. Box 800709, Charlottesville, VA 22908-0709. Tel.: +1 434 982 1632; fax: +1 434 924 2260.
E-mail address: rws2k@virginia.edu (R.G. Sawyer).
0022-4804/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jss.2014.05.056

Fundamentally, volume-based criteria under-appreciate variable learning rates and provide no assurance that competency persists after initial training [5,6]. Over the 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 because of 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.

2. Materials and methods

We enrolled consenting postgraduate year 1 surgery residents who each subsequently completed a 1-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 used a guidewire-assisted push-placement gastrostomy kit (Bard Access Systems, Salt Lake City, UT), 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 takeover were prospectively recorded. Procedure time started on 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 *versus* failure), and sums this data to provide a longitudinal representation of the training process. Cusum uses several parameters defined by evaluators. The acceptable task failure rate (p_0) is the inherent failure rate of a proficient practitioner, whereas 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 subpar (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, (h_0) are determined by the following equations:

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

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

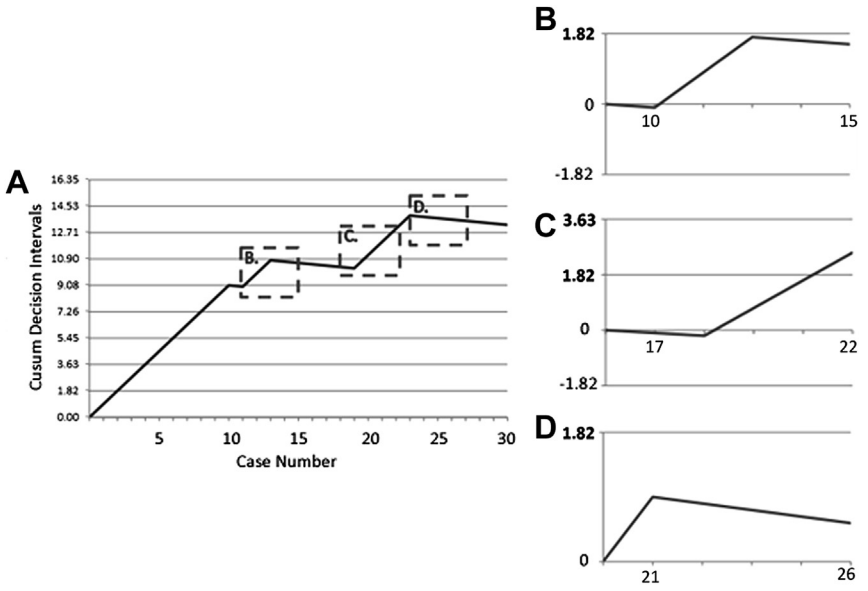

Fig. 1 – Cusum analysis of participant number 2. Overall Cusum curve (A). Windowed analyses demonstrate proficiency at case 15 (B) with relapse at case 22 (C), at which point the 1.82 unacceptable decision interval is crossed. Proficiency is regained at case 26 (D).