The Journal of Laryngology & Otology

cambridge.org/jlo

Assessment of progress and competence in tonsillectomy surgery using cumulative sum analysis

A K Abou-Foul[1], A Taghi[1], N S Tolley[1,2] and Z Awad[1,2]

[1]Department of Otolaryngology, Head and Neck Surgery, Imperial College Healthcare NHS Trust, London, and [2]Department of Surgery, Imperial College London, UK

Main Article

Dr A K Abou-Foul takes responsibility for the integrity of the content of the paper

Data from this study were presented at the 177th American Academy of Otolaryngology, Head and Neck Surgery Foundation annual meeting, 29 September – 2 October 2013, Vancouver, Canada.

Cite this article: Abou-Foul AK, Taghi A, Tolley NS, Awad Z. Assessment of progress and competence in tonsillectomy surgery using cumulative sum analysis. *J Laryngol Otol* 2018; **132**:734–738. https://doi.org/10.1017/S0022215118001081

Accepted: 23 February 2018
First published online: 12 July 2018

Key words:
Tonsillectomy; Surgical Training; Learning Curve; Clinical Competency; Outcomes Assessment; Otolaryngology

Author for correspondence:
Dr Ahmad K Abou-Foul,
Department of Otolaryngology,
Head and Neck Surgery,
Imperial College Healthcare NHS Trust,
St Mary's Hospital,
Praed Street,
London W2 1NY, UK
E-mail: Dr.akaboufoul@hotmail.com
Fax: +44 203 312 1847

Abstract

Objective. To assess the feasibility of using cumulative sum analysis to show trainees' performance curves and highlight concerns in tonsillectomy surgery.
Methods. In this prospective study, the performance of eight otolaryngology trainees (in their first 6–12 months in the specialty) was compared to that of experts (over 150 tonsillectomies performed) in terms of operative time and post-operative complications. Cumulative sum analysis curves were generated to highlight trainees' performance, and were updated after each performance.
Results. The average operative time was 23 minutes (standard deviation = 11) for experts and 38 minutes (standard deviation = 16) for trainees ($p < 0.0001$). Cumulative sum analysis charts for trainees' operative time initially rose and then started to plateau after a number of cases (range, 25–30), while that of experts remained low. Cumulative sum analysis charts for complications can be used to monitor performance. In this study, complications were combined (because of low incidence) to allow use of this outcome measure.
Conclusion. The flexibility of the cumulative sum analysis makes it adaptable to any outcome. It is a strong adjunct in surgical training to monitor progress and competence. Its sensitivity also allows early detection of poor performance, to instigate intervention.

Introduction

Tonsillectomy is one of the most common operations performed by otolaryngologists. In the USA, 399 000 tonsillectomies were carried out in 2010 and 27 500 were conducted in the UK in 2014.[1,2] The National Prospective Tonsillectomy Audit, carried out in England and Northern Ireland between 2003 and 2004, found that around one-third of all tonsillectomies included were performed by surgical trainees.[3] The audit also showed that trainees had statistically higher complication rates than experts.[3]

Tonsillectomy is one of the first operative procedures that otolaryngology trainees are exposed to and expected to perform independently. It requires knowledge of anatomy and various generic technical skills such as hand–eye co-ordination, two-handedness, tying at depth, and sound tissue handling. All UK otorhinolaryngology training programmes require trainees to be competent in performing tonsillectomy operations independently, prior to completion of training. Complications such as primary or secondary haemorrhage arising from tonsillectomy can be potentially disastrous. Therefore, quality assurance of the trainee performing tonsillectomy is critically important, but this can be very difficult to assess objectively.

Cumulative sum ('CUSUM') analysis is a statistical technique, which was initially proposed in 1954 as a quality control method.[4] Cumulative sum analysis tables are used to generate a control chart to monitor outcomes of a procedure performed over time by an individual operator. This allows a judgement to be made as to whether a variation in performance is significant or the result of a random variation.[5] It has flexibility in allowing the analysis of a dynamic sample size. In addition, the chart can be updated after every performance, facilitating 'real-time' feedback.

Cumulative sum analysis is being increasingly used as a strong adjunct to surgical training. The performance of a trainee can be compared to that of experts as a surrogate marker of competence.[6] The analysis has been applied to training in several surgical procedures, including thyroidectomy,[7] parotidectomy,[8] oesophagectomy,[9] robotic-assisted laparoscopic colorectal surgery[10] and sentinel lymph node biopsy.[11]

Our primary objective was to assess the applicability of cumulative sum analysis control charts in demonstrating trainees' progress and competence with regard to tonsillectomy operations, in addition to highlighting any concerns relating to performance.

Materials and methods

Study design and participants

Eight junior otolaryngology surgical trainees in their first 6–12 months of training within the specialty were selected. The trainees were observed in operating theatres by their

https://doi.org/10.1017/S0022215118001081 Published online by Cambridge University Press

supervising trainers. Two outcome measures were used: operative time and post-operative complications. The results were compared to those of experienced surgeons who had performed more than 150 tonsillectomies.

Generating cumulative sum analysis curves

The mean and standard deviation (SD) of the experts' performance was used to draw the cumulative sum analysis curve, applying the standard formula ($Cn = 0$, $Cn - 1 + Xn - k$) on consecutive procedures, where C = trainee's performance, n = number of procedures in a chronological order, Xn = outcome measure for the nth procedure and k = target value (pre-specified standard of performance).

A positive performance (success) was recorded for completion within 1 SD of the experts' time, and denoted as $Xn = 0$. If operative time was more than 1 SD of the experts' completion time, then $Xn = 1$ (failure). When charted on the y-axis (against the number of cases on the x-axis), this then allows interpretation of training progress by the direction and trajectory of the curve. Performance in line with the experts will move the curve downwards by accumulation of scores of $Xn = 0$. Conversely, an upwards curve denotes performance moving away from the experts' level. Acceptable failure or permitted error is corrected for by $k = 0.2$; as such, prolonged performance time in one of every five cases is considered allowable based on departmental data.

Post-operative complication data obtained from medical records, and by contacting the patients or parents, were also recorded. Complications included bleeding, delayed hospital discharge, delayed return to school or work, hospital re-attendance, or visiting a doctor for pain or infection. Complications were combined because of their low incidence and the usual overlap between them. Tonsillectomy patients presenting with post-operative bleeding will usually require additional hospital stay, and those who return to hospital or visit doctors often do so because of excessive pain or infection, and tend to require antibiotic prescriptions.

Each complication was regarded as a negative episode ($Xn = 1$). Charting negative episodes over a number of cases produced a similarly interpretable curve, whereby a downward slope indicates performance approaching that of the experts' level. Allowable failure was set at $k = 0.1$ based on the experts' data, which showed 10 per cent incidence of such complications.

Ethical considerations

Local research and development department approval was obtained for the purpose of data collection and protection. All trainees consented to be enrolled in the study. No new interventions were instigated and no changes to clinical practice were made.

Statistical analysis

The statistical analysis was performed with SPSS software, version 23.0 (IBM, Armonk, New York, USA), with statistical significance set at $p < 0.05$. The Mann–Whitney U test was used to compare the means of non-parametric data. Diagramming of the cumulative sum analysis charts for the learning curves was performed in Microsoft Office Excel 2007 spreadsheet software (Microsoft, Redmond, Washington, USA).

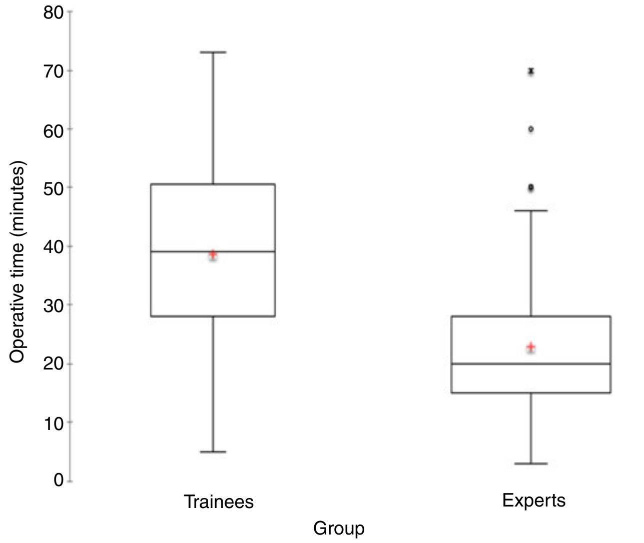

Fig. 1. Box-and-whisker plot of time taken to perform tonsillectomy, comparing trainees to experts (Mann–Whitney U test $p < 0.0001$).

Results

All of the 8 trainees performed 14–50 tonsillectomies each. The mean (± SD) operative time for experts, who completed over 150 procedures, was 23 ± 11 minutes. Trainees' operative time was significantly longer, at 38 ± 16 minutes ($p < 0.0001$), as demonstrated in Figure 1.

Operative time failure was judged by the criteria described above, where $Xn = 0$ if performance was within 1 SD of the experts' completion time (34 minutes) and $Xn = 1$ if longer than 34 minutes. The presence of complications gave a score of $Xn = 1$, and the absence of complications was scored as $Xn = 0$.

For trainees who had never performed a tonsillectomy before, their operative time and complications cumulative sum analysis curves rose as they accumulated failures, and then started to plateau after a variable number of cases (range, 25–30) (Figures 2 and 3). Some trainees had performed tonsillectomies prior to this study and this is presented as a delayed start in their curves. By contrast, the operative time and complications cumulative sum analysis curves of experts stayed flat, as demonstrated in Figures 2 and 3.

Discussion

It has long been known that training and practice are key factors for improving clinical skills and achieving mastery in surgical procedures.[5] Recent research has shown that practice significantly reduces operative time and improves surgical outcomes.[12,13] However, the time available for training and 'on-the-job' learning has reduced considerably, and there is an emerging need for alternative methods of learning, and, more importantly, monitoring for progress and competence.

In the UK, assessing competency in surgical practice has been performed by a number of tools, such as workplace-based assessments and objective structured clinical examinations.[14] These assessment tools are considered essential for training progression and are widely adopted by surgical trainers. However, these tools have been criticised for being subjective, and they require an adequate allocation of time and resources.[15,16] Cumulative sum analysis is a new emerging assessment tool focused on technical competency that permits a dynamic, objective and reliable assessment of a trainee's

https://doi.org/10.1017/S0022215118001081 Published online by Cambridge University Press

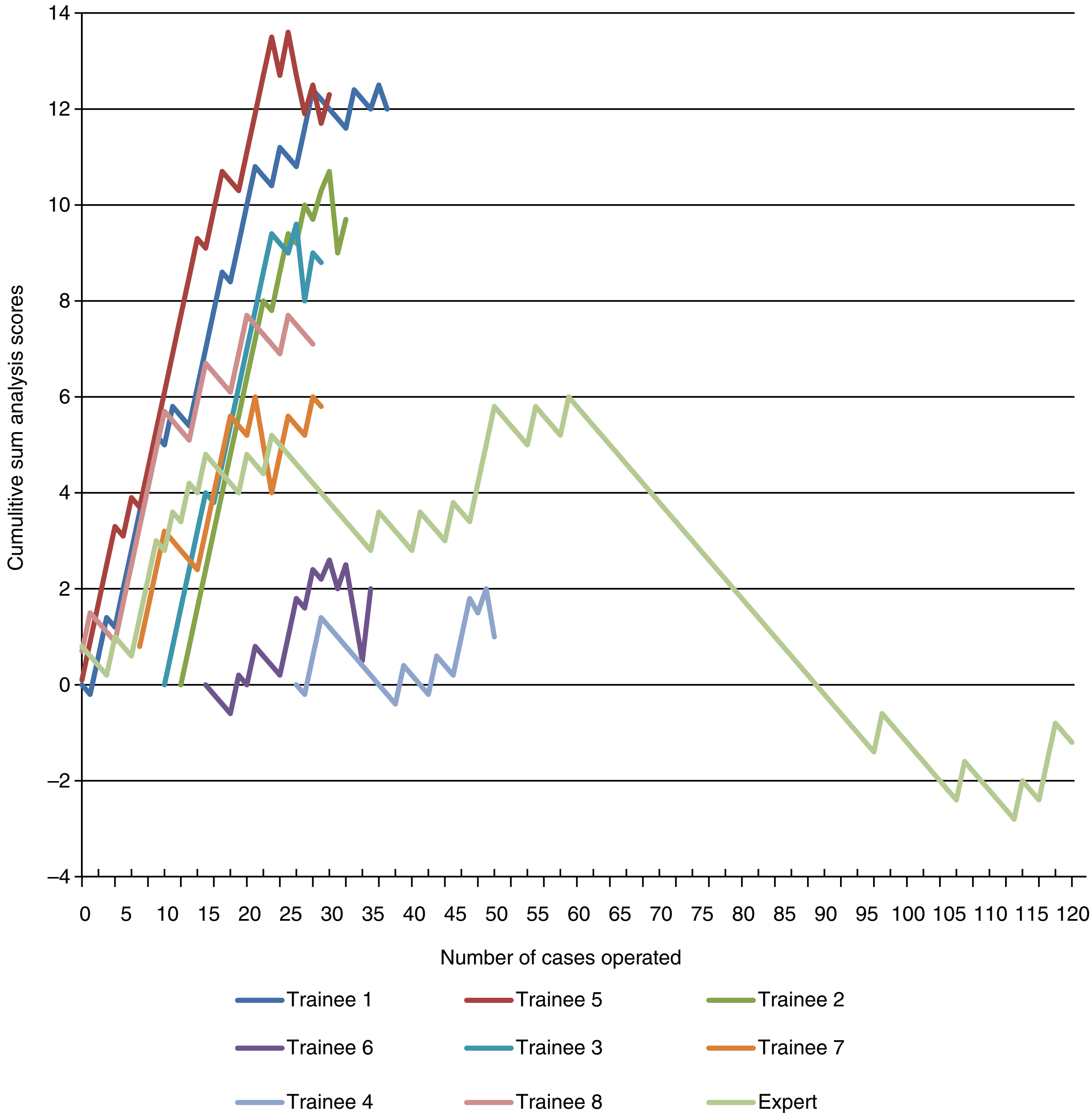

Fig. 2. Diagram showing cumulative sum analysis curves of individual trainees versus expert with regard to operative time.

performance. It can be used to evaluate the quality of surgical skills accumulated over the training period, provide individualised feedback and potentially benchmark training programmes.[17,18]

This study highlights the potential role of cumulative sum analysis curves for mapping the performance of trainees in relation to the number of procedures performed. It addresses three key issues in surgical training. Firstly, the number of procedures alone is not a reliable indicator of competence.[19] Secondly, individual learning curves can give trainees insight and allow direction towards appropriate training opportunities. Thirdly, cumulative sum analysis curves have the potential to create a feasible and reliable measure of competence achievement, which, in time, can be validated.

Our results demonstrate that cumulative sum analysis curves can be drawn to describe trainees' learning curves for performing tonsillectomy operations. For most of the surgical trainees, their learning resulted in a curve that initially deflected steeply upwards, and which then began to plateau as their performance improved with increasing surgical experience. The point at which this plateau was reached varied between trainees, but generally occurred after operating on between 25 and 30 cases. As the cumulative sum analysis curve plateaued, it mirrored the expert's curve, and it is reasonable to extrapolate that this equates with competence in performing tonsillectomy.

It is important to highlight that skills progression and attainment of competence usually varies depending on the individual. Cumulative sum analysis curves can be used to monitor and predict inter-individual variability in performance, and to create individualised learning curves. These curves can be updated after every performance as data increase, and agreed outcomes can be optionally included in the analyses. In addition to the assessment of post-operative complications included in this study, other indirect markers of performance, such as patient-reported outcome measures and post-operative analgesia requirements, can also be included.

https://doi.org/10.1017/S0022215118001081 Published online by Cambridge University Press

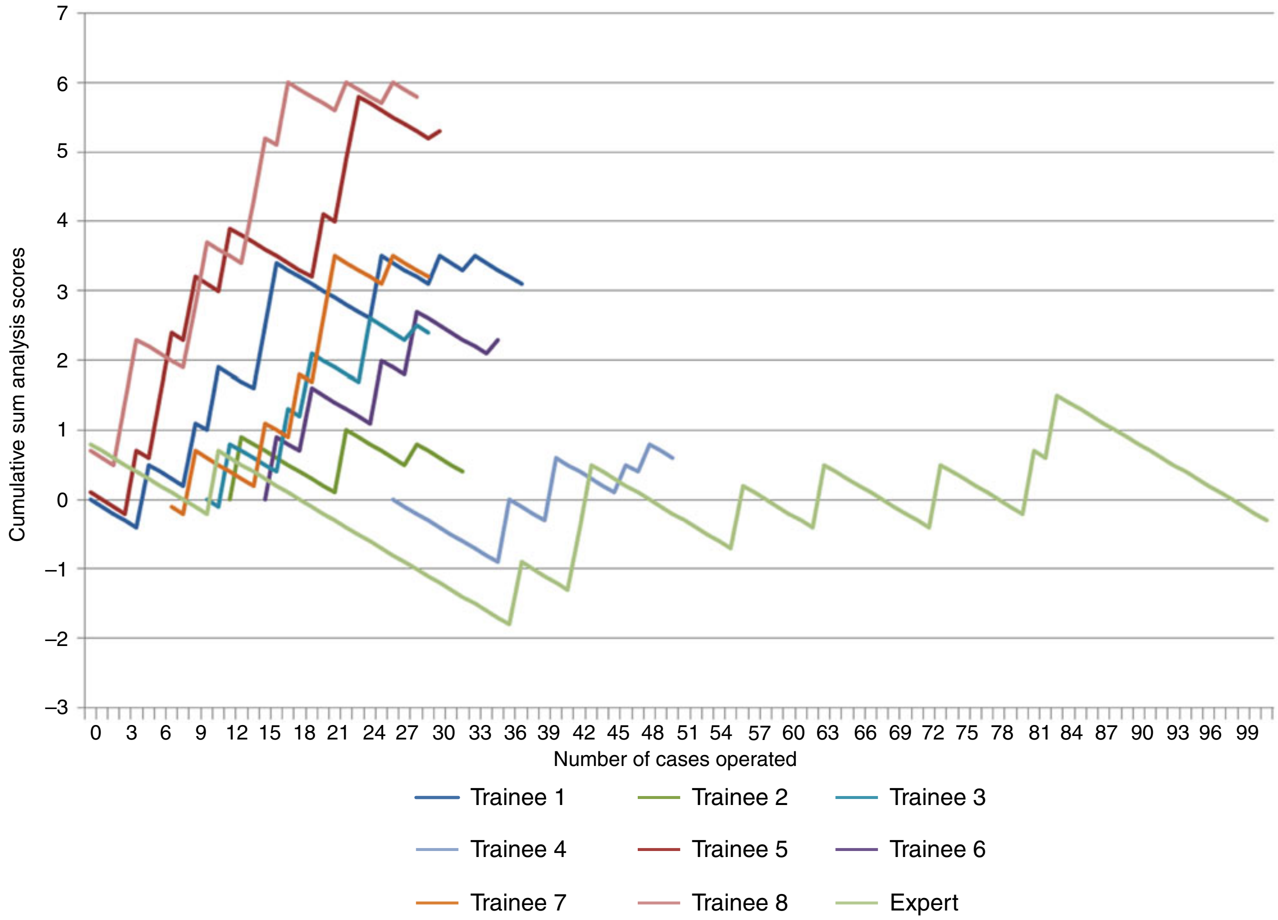

Fig. 3. Diagram showing cumulative sum analysis curves of individual trainees versus expert with regard to post-operative complications.

Understanding and interpreting the cumulative sum analysis curves in a mutually agreed way between the trainer and any trainee is important for their effectiveness. It is mandatory to maintain close supervision on new trainees when their curves are deflecting upwards. This close monitoring should aim to provide immediate feedback on the technique, to improve fluency and reduce complications. A plateau in the curve should be an anticipated milestone, and should be ideally associated with a reduction in the degree of supervision and guidance needed. A consistent plateau for a predetermined number of cases over a reasonable period of time can be safely interpreted as competency. Achieving competency is a benchmark by which trainees are permitted to perform the procedure independently without supervision, with occasional acceptable failure expected, comparable to the experts' level.[20]

Trainees should never be allowed to operate without supervision if their cumulative sum analysis curves have not yet plateaued, as patients' safety can be jeopardised. If no plateau is achieved after a reasonable period of time and number of cases, the supervisor and the trainee have a responsibility to identify areas of weakness, revise the technique and construct a mutually agreed action plan. The same process should be instigated if cumulative sum analysis curves deteriorate after a period of plateau, to ensure consistent patient safety.

A further application of cumulative sum analysis is as a quality control for trained ENT surgeons. There is much current debate around surgeon-specific outcomes and revalidation of the quality of care provided after completion of training. Cumulative sum analysis could provide one way of objectively demonstrating performance within agreed expert parameters, enabling consultants to map their outcomes against the control curve throughout their careers.

Study limitations

Potential limitations of this study include the inability to correct for complex cases using current parameters. As surgical trainees progress, they are exposed to cases that are innately more difficult or time consuming (e.g. procedures conducted for histological examination or on a background of coagulopathy). At present, there is no stratification system that quantifies the complexity of tonsillectomy cases. If, in the future, such a system was to be developed, it would allow the standard cumulative sum analysis to be converted to a risk-adjusted cumulative sum analysis, with more nuanced evaluation.

- Tonsillectomy is one of the most common operations performed by otolaryngology trainees
- Tonsillectomy performed by trainees results in more complications than that performed by experts
- Cumulative sum analysis is an emerging statistical method used to monitor outcomes of a procedure performed over time by an individual operator
- This study is the first to show the potential of cumulative sum analysis curves for mapping trainees' performance in tonsillectomy operations
- The analysis curves can also be used to monitor and predict inter-individual variability in performance, and individualised learning curves can be created

https://doi.org/10.1017/S0022215118001081 Published online by Cambridge University Press

Another study limitation is that only 8 trainees were included, who performed less than 50 cases each. A larger number of trainees and cases over a longer timeframe would improve the accuracy of the time taken to achieve a plateau in the cumulative sum analysis curve and allow for the degree of variability between trainees. In addition, complications such as bleeding are rare after tonsillectomy, and hence larger numbers are important to confirm improvement.

Conclusion

Cumulative sum analysis can be used as an evaluation tool to monitor performance and draw learning curves for tonsillectomy. The flexibility of the cumulative sum analysis makes it adaptable to any outcome, which gives it an advantage over existing assessment techniques. Moreover, cumulative sum analysis curves can be utilised to give trainees real-time feedback of their performance using an objective measure, and can alert trainers to a potential trainee in difficulty.

Competing interests. None declared.

References

1 Hall MJ, Schwartzman A, Zhang J, Liu X. Ambulatory surgery data from hospitals and ambulatory surgery centers: United States, 2010. *Natl Health Stat Report* 2017;**102**:1–15
2 ENT UK. Tonsillectomy commissioning guide 2016. In: https://www.rcseng.ac.uk/-/media/files/rcs/standards-and-research/commissioning/ent-uk-tonsillectomy-revised-commissioning-guide-2016-published.pdf [04 June 2018]
3 Lowe D, van der Meulen J, Cromwell D, Lewsey J, Copley L, Browne J *et al.* Key messages from the National Prospective Tonsillectomy Audit. *Laryngoscope* 2007;**117**:717–24
4 Page ES. Continuous inspection schemes. *Biometrika* 1954;**41**:100–15
5 Chen T, Vamos AC, Dailey SH, Jiang JJ. CUSUM analysis of learning curves for the head-mounted microscope in phonomicrosurgery. *Laryngoscope* 2016;**126**:2295–300
6 Steiner SH, Woodall WH. Debate: what is the best method to monitor surgical performance? *BMC Surg* 2016;**16**:15
7 Duclos A, Carty MJ, Peix J-L, Colin C, Lipsitz SR, Kraimps JL *et al.* Development of a charting method to monitor the individual performance of surgeons at the beginning of their career. *PLoS One* 2012;**7**:e41944
8 Sharp JF, Cozens N, Robinson I. Assessment of surgical competence in parotid surgery using a CUSUM assessment tool. *Clin Otolaryngol Allied Sci* 2003;**28**:248–51
9 Roberts G, Tang CB, Harvey M, Kadirkamanathan S. Real-time outcome monitoring following oesophagectomy using cumulative sum techniques. *World J Gastrointest Surg* 2012;**4**:234–7
10 Bokhari MB, Patel CB, Ramos-Valadez DI, Ragupathi M, Haas EM. Learning curve for robotic-assisted laparoscopic colorectal surgery. *Surg Endosc* 2011;**25**:855–60
11 Lerch L, Donald JC, Olivotto IA, Lesperance M, van der Westhuizen N, Rusnak C *et al.* Measuring surgeon performance of sentinel lymph node biopsy in breast cancer treatment by cumulative sum analysis. *Am J Surg* 2007;**193**:556–60
12 Rogers DA, Elstein AS, Bordage G. Improving continuing medical education for surgical techniques: applying the lessons learned in the first decade of minimal access surgery. *Ann Surg* 2001;**233**:159–66
13 Awad Z. *A Comprehensive Evaluation of Work and Simulation Based Assessment in Otolaryngology Training*. London: Imperial College London, 2014
14 Memon MA, Brigden D, Subramanya MS, Memon B. Assessing the surgeon's technical skills: analysis of the available tools. *Acad Med* 2010;**85**:869–80
15 Massie J, Ali JM. Workplace-based assessment: a review of user perceptions and strategies to address the identified shortcomings. *Adv Health Sci Educ Theory Pract* 2016;**21**:455–73
16 Wilkinson JR, Crossley JG, Wragg A, Mills P, Cowan G, Wade W. Implementing workplace-based assessment across the medical specialties in the United Kingdom. *Med Educ* 2008;**42**:364–73
17 Hu Y, Jolissaint JS, Ramirez A, Gordon R, Yang Z, Sawyer RG. Cumulative sum: a proficiency metric for basic endoscopic training. *J Surg Res* 2014;**192**:62–7
18 Campbell RD, Hecker KG, Biau DJ, Pang DS. Student attainment of proficiency in a clinical skill: the assessment of individual learning curves. *PLoS One* 2014;**9**:e88526
19 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**:983–9
20 Li X, Wang J, Ferguson MK. Competence versus mastery: the time course for developing proficiency in video-assisted thoracoscopic lobectomy. *J Thorac Cardiovasc Surg* 2014;**147**:1150–4

https://doi.org/10.1017/S0022215118001081 Published online by Cambridge University Press