Learning Curves for Bag-and-mask Ventilation and Orotracheal Intubation

An Application of the Cumulative Sum Method

Ryu Komatsu, M.D.,* Yusuke Kasuya, M.D.,* Hisanori Yogo, M.D.,* Daniel I. Sessler, M.D.,†
Edward Mascha, Ph.D.,‡ Dongsheng Yang, M.S.,§ Makoto Ozaki, M.D.‖

ABSTRACT

Background: In this study, the authors determined the success and failure rates for interns learning bag-and-mask ventilation and orotracheal intubation. Their goal was to determine the amount of experience needed to perform these procedures correctly.

Methods: The authors recorded 695 bag-and-mask ventilations and 679 orotracheal intubations performed by 15 inexperienced interns during their 3 month-long anesthesia rotations. Learning curves for each procedure for each intern were constructed with both the standard and risk-adjusted cumulative sum methods. The average number of procedures required to attain a failure rate of 20% was estimated for each technique.

Results: Fourteen of 15 interns attained acceptable failure rates at bag-and-mask ventilation after 27 ± 13 procedures, with a median (95% confidence interval) of 25 (15–32) procedures to cross the decision limit when considering all 15 interns. Nine of 15 interns attained acceptable failure rates at orotracheal intubation after 26 ± 8 procedures, with a median of 29 (22–not estimable) procedures to cross the limit when considering all interns. The proportion of interns who attained acceptable failure rates for mask ventilation was greater than for tracheal intubation (93% *vs.* 60%, *P =*

0.025). Overall, our interns achieved a bag-and-mask ventilation failure rate of 20% or better after a median of 25 procedures; approximately 80% of interns achieved the goal after 35 procedures or less.

Conclusions: Participating interns developed mask ventilation skills faster than orotracheal intubation skills, and there was more variability in the rate at which intubation skills developed. A median of 29 procedures was required to achieve an 80% orotracheal intubation success rate.

What We Already Know about This Topic

❖ Learning curves for endotracheal intubation have been examined, but not so much for bag-and-mask ventilation and using modern, cumulative sum statistical methods

What This Article Tells Us That Is New

❖ In 15 inexperienced interns in specialties other than anesthesia, 14 achieved an 80% success rate after a median of 25 bag-and-mask ventilation procedures

❖ Only 9 of 15 interns achieved an 80% success rate for endotracheal intubation, and these did so after a median of 35 procedures

* Clinical Instructor, ‖ Professor and Chair, Department of Anesthesiology, Tokyo Women's Medical University, Tokyo, Japan. † Professor and Chair, Department of OUTCOMES RESEARCH, ‡ Staff Biostatistician, § Statistical Programmer, Departments of OUTCOMES RESEASRCH and Quantitative Health Sciences, The Cleveland Clinic, Cleveland, Ohio.

Address correspondence to Dr. Sessler: Department of OUTCOMES RESEARCH, The Cleveland Clinic, 9500 Euclid Avenue, P-77, Cleveland, Ohio 44195. ds@or.org. Information on purchasing reprints may be found at www.anesthesiology.org or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

MAINTENANCE of the patient airway is a primary responsibility of anesthesiologists and all physicians required to provide patient care in emergency medicine or critical care settings. Interruption of gas exchange, even for a few minutes, can result in catastrophic outcomes including brain damage or death. In the closed-claims analysis, 28% of anesthesia-related deaths or brain damage were because of respiratory-related damage events.[1] Commonly used airway management techniques include bag-and-mask ventilation and laryngoscopic orotracheal intubation. Learning curves for various anesthesia procedures have been evaluated using various statistical approaches,[2–8] but few studies have investigated learning curves for bag-and-mask ventilation.

Control charts are statistical tools for analyzing data during production or research in industry; the values of the quality characteristics are plotted in sequence. The distribution of the plotted values in relation to the control limits

provides statistical information on the process under study. In cumulative sum (cusum) control charts, the cumulative differences of the quality characteristic from a target level are plotted in sequence, leading to a tighter control of a given process and allowing detection of deviations from preestablished standards.[9] Cusum charts have been used for surgical skill assessment.[10] The control chart is especially sensitive to short runs of failures and identifies them quickly. Hence, the trainee who has difficulty in a particular technique can be detected, and immediate corrective measures can be taken.

The purpose of this study was to construct learning curves for two basic airway management techniques: bag-and-mask ventilation and laryngoscopic orotracheal intubation. Specifically, we sought to determine the amount of experience required for each technique and to determine whether one airway management strategy was significantly easier to learn than the other.

Materials and Methods

Interns from Tokyo Women's Medical University were recruited as subjects at the beginning of their 3-month anesthesia rotation. Participating interns were committed to a nonanesthesia specialty and presumably had no special interest in airway management. Among the 15 interns included, six specialized in internal medicine, one in family practice, two in general surgery, one in cardiothoracic surgery, one in urology, one in emergency medicine, and three in pediatrics. They were involved in the study during the first 9 months of their internship. With the approval of the Tokyo Women's Medical University Institutional Review Board (Tokyo, Japan) and written informed consent from the interns, we asked the interns to complete a questionnaire to assess their previous experience with training in the performance of mask ventilation and laryngoscopic tracheal intubation. Subjects who had previously performed more than 10 of either procedure were excluded.

Protocol and Measurements

All interns received a formal lecture on performing mask ventilation and laryngoscopic tracheal intubation. They were permitted to practice these procedures on a mannequin before their first attempt in the clinical setting. The period of mannequin practice was not restricted, and the procedures were performed under an instructor's supervision with remarks about the technique allowed. At the beginning of the training period, the interns were instructed about criteria for failure and success for each procedure. The interns completed a data collection form immediately after the procedure that was contemporaneously reviewed and signed by the supervising instructor.

Adult general surgical patients, except for cardiothoracic cases, were included in the study. Patients with anticipated difficult intubation or mask ventilation, with history of previous difficult intubation or mask ventilation, or with significant cervical spine pathology were excluded from the study, as were those in whom awake intubation was indicated for any reason.

Tokyo Women's Medical University institutional review board waived consent from the patients for participation in the study, as it was clearly explained to all patients at admission that trainees were performing the procedures under the supervision of attending physicians in this teaching hospital.

Head position was standardized with a 7-cm-high anesthesia pillow and sniffing position. A head band was not used for mask ventilation. Mask ventilation was assessed after anesthesia induction with or without muscle relaxation. When patients required cricoid pressure and rapid sequence intubation, face mask ventilation was not assessed. Mask ventilation was deemed successful when it produced chest excursion sufficient to maintain capnograph waveforms with plateau formation, and the oxyhemoglobin saturation was maintained at the preinduction level (the value after preoxygenation). Use of an oral airway was allowed. Need for any physical assistance by the supervising instructor was considered a failure.

Laryngoscopic orotracheal intubation was performed with cuffed tracheal intubation under direct laryngoscopy with Macintosh blades. The size of endotracheal tube and laryngoscope blade were chosen by the interns depending on the overall size of the patient. The interns were allowed to ask for external laryngeal pressure to improve the view during laryngoscopy. An intubation attempt was defined by one entry of the blade into the mouth. If one blade was switched for another after the initial attempt, it was considered a second attempt. If the first attempt of intubation failed, the intern was allowed one additional attempt; after two failed attempts, the supervising instructor took over the procedure and secured the airway. Successful intubation was confirmed by chest movement, auscultation, and capnography. If intubation was not successful within two attempts by the participating intern, the effort was considered a failure.

Data Analysis

We calculated both standard and risk-adjusted cusum charts for each intern and each intubation method.

Standard Cusum. Standard cusum charting requires specification of acceptable (p_0) and unacceptable (p_1) failure rates for the process under study and type I and type II errors (α and β, respectively). A type I error is the probability of crossing the acceptable failure rate limit when the true failure rate is not in the acceptable range, whereas type II error is the probability of failing to cross the acceptable limit when the true failure rate is in the acceptable range. Upper and lower decision limits (h_1 and h_0, respectively), corresponding to unacceptable and acceptable failure rates, are functions of p_0, p_1, α, and β and are calculated as follows:

$$h_1 = a/(P + Q), \quad h_0 = b/(P + Q), \text{where}$$

$$a = \ln\left[(1 - \beta)/\alpha\right],$$

$$b = \ln\left[(1 - \alpha)/\beta\right],$$

$$P = \ln(p_1/p_0) \text{ and}$$

$$Q = \ln\left[(1 - p_0)/(1 - p_1)\right].$$

Table 1. Cumulative Sum Calculations for Varying Acceptable and Unacceptable Failure Rates

Parameters	p0 = 5%, p1 = 10%	p0 = 10%, p1 = 20%	p0 = 20%*, p1 = 40%*
Amount subtracted for each success, s = Q/(P + Q)	0.072	0.145	0.293
Acceptable failure rate threshold, h0 = b/(P + Q)	−2.94	−2.71	−2.24
Unacceptable failure rate threshold h1 = a/(P + Q)	2.94	2.71	2.24
a = ln ([1 − β]/α)	2.197	2.197	2.197
b = ln ([1 − α]/β)	2.197	2.197	2.197
P = ln (p1/p0)	0.693	0.693	0.693
Q = ln [(1 − p0)/(1 − p1)]	0.054	0.117	0.288
Expected number of attempts to cross h0 under given failure rate p0: (h0 [1 − α] − αh1)/(s − p0)	105	48	19
Average number of attempts to cross h1 under given failure rate p1: (h1 [1 − β] − βh0)/(p1 − s)	85	40	17

These values (p0 = 0.20 and p1 = 0.40) are used in the main analysis.

α = 0.10 = probability of type 1 error; β = 0.10 = probability of type 2 error; p0 = acceptable failure rate; p1 = unacceptable failure rate; 1 − S = added for each failure; S = amount subtracted from cumulative sum lines for each success.

The resulting chart begins at 0, and for each success, the amount S = Q/(P + Q), a function of p0 and p1, is subtracted from the previous cusum value; for each failure, the amount 1 − S is added to the previous cusum value. A cusum chart is formed by connecting the cusum values over time for each participant. When the chart crosses the upper decision limit (h1) from below, the participant's failure rate is deemed significantly greater than the acceptable failure rate. When the chart crosses the lower decision limit (h0) from above, the true failure rate is deemed to be as low or lower than the acceptable failure rate, with false positive probability α. When the cusum line is maintained between the decision limits, no statistical inference can be made.

The incidence of difficult mask ventilation is reported to be as high as 5%.[11,12] But considering that our interns were nonanesthesiologists rotating in anesthesia for only 3 months and that we did not permit physical assistance from instructors to facilitate mask ventilation, we set the acceptable failure rate (p0) for mask ventilation at 20% (*i.e.*, four times the reported failure rate among experienced anesthesiologists). As is customary for cusum analysis, the unacceptable failure rate (p1) was set at twice p0 or 40%. The probability of type I (α) and II (β) errors were each set to 0.1.

The reported incidence of a suboptimal laryngoscopic view for intubation (*e.g.*, Cormack and Lehane grade 3 and 4) has been reported to be as high as 10%.[11–14] Again considering the inexperience of our interns, we set the acceptable failure rate (p0) for laryngoscopic tracheal intubation at 20% and unacceptable failure rate (p1) at twice p0 (40%) and type I and II errors at 0.1. These parameters were identical to those used in a similar study by de Oliveira Filho.[2]

Cusum calculations were performed for mask and intubation procedures using the aforementioned formulas (and in table 1) and the chosen values of p0, p1, α, and β. In table 1, average sample sizes of runs having acceptable failure rates (p0) of 5%, 10%, and 20% were estimated as 105, 48, and 19 procedures, respectively. For runs having unacceptable failure rates (p1) of 10%, 20%, and 40%, average sample sizes were estimated as 85, 40, and 17 procedures, respectively. Assuming that the interns would experience 40 to 50 procedures for both mask ventilation and intubation during 3 months, the study would provide enough power for acceptable failure rate (p0) of 20% and unacceptable failure rate (p1) of 40%.

Time to crossing h0, the lower decision boundary indicating acceptable failure rate, was summarized in a Kaplan-Meier time-to-event plot for each technique for the prespecified rate of p0 and error rates. Success proportions from before and after crossing the lower decision limit were compared by a generalized estimating equation Z-test to adjust for the within-subject correlation over repeated attempts within-intern. The proportion of interns eventually crossing the lower decision limits for mask ventilation and tracheal intubation were compared using McNemar's test for paired proportions.

Risk-adjusted Cusum. In addition to the standard cusum chart, we report the learning curves adjusted for varying levels of difficulty presented by each patient. We created risk-adjusted charts using the observed minus expected (O − E) cusum method.[15] A risk score for each patient was first calculated as the estimated probability of failure predicted as a function of traditional difficult mask ventilation and intuba-

Table 2. Mask Ventilation: Patient Risk Demographics

	Fail (n = 103)	Success (n = 592)	P Value*
Factor			
Body mass index >30 kg/m^2	3 (2.9)	16 (2.7)	0.75†
Mallampati 3, 4	14 (13.6)	66 (11.2)	0.47
Age 57+ yr	59 (57.3)	250 (42.2)	0.005
Mandibular protrusion (severely limited)	1 (0.97)	5 (0.84)	0.99†

Data are expressed as n (%).

* Chi-square test unless noted; † Fisher exact test.

Table 3. Orotracheal Intubation: Patient Risk Demographics

Factor	Fail (n = 151)	Success (n = 528)	P Value*
Body mass index >30 kg/m²	4 (2.7)	14 (2.7)	0.99†
Mallampati 3, 4	20 (13.3)	57 (10.8)	0.40
Age 57+ yr	58 (38.4)	239 (45.3)	0.13
Mandibular protrusion (severely limited)	4 (2.7)	2 (0.38)	0.024†
Mouth opening <3 cm	11 (7.3)	29 (5.5)	0.41
Thyromental distance <6 cm	13 (8.6)	36 (6.8)	0.45
Neck movement <90°	30 (19.9)	60 (11.4)	0.007

Data are expressed as n (%).

* Chi-square test unless noted; † Fisher exact test.

tion risk factors[14,16,17] (tables 2 and 3) using logistic regression. The risk-adjusted cusum chart was then formed by adding 1 minus the individual patient risk score to the cumulative score for each failure and subtracting the risk score for each failed attempt. The cusum at time t is, thus, $c_t = c_{t-1} + (x_t - x_0)$, where c_{t-1} is the cusum through the previous attempt, x_t is 1 for failure and 0 for success (observed), and x_0 is the estimated risk for the patient being intubated or ventilated at time t (expected). Interns whose accumulated observed rate of failure is generally consistent with the expected values for the patients they intubated will have scores near the zero line, whereas those performing better or worse than expected will have scores below and above the zero line, respectively.

Data are presented as mean ± SDs unless otherwise stated. Individual interns are represented by randomly assigned capital letters. SAS statistical software (SAS Institute, Inc., Cary, NC) and the R programming language were used for all analyses.

Results

A total of 706 patients were involved in mask ventilation and/or intubation. Standard and risk-adjusted cusum lines for mask ventilation and orotracheal intubation are presented in figure 1.

Mask Ventilation

The 15 interns attempted a total of 695 mask ventilations, with 592 (85%) being successful (table 4). An oral airway was used on 136 occasions, and muscle relaxant was used on 669 occasions of the attempts. An instructor took over ventilation in 103 (15%) of the cases. The mean number of mask ventilation procedures per intern was 46 ± 15 (range, 27–84). All but one
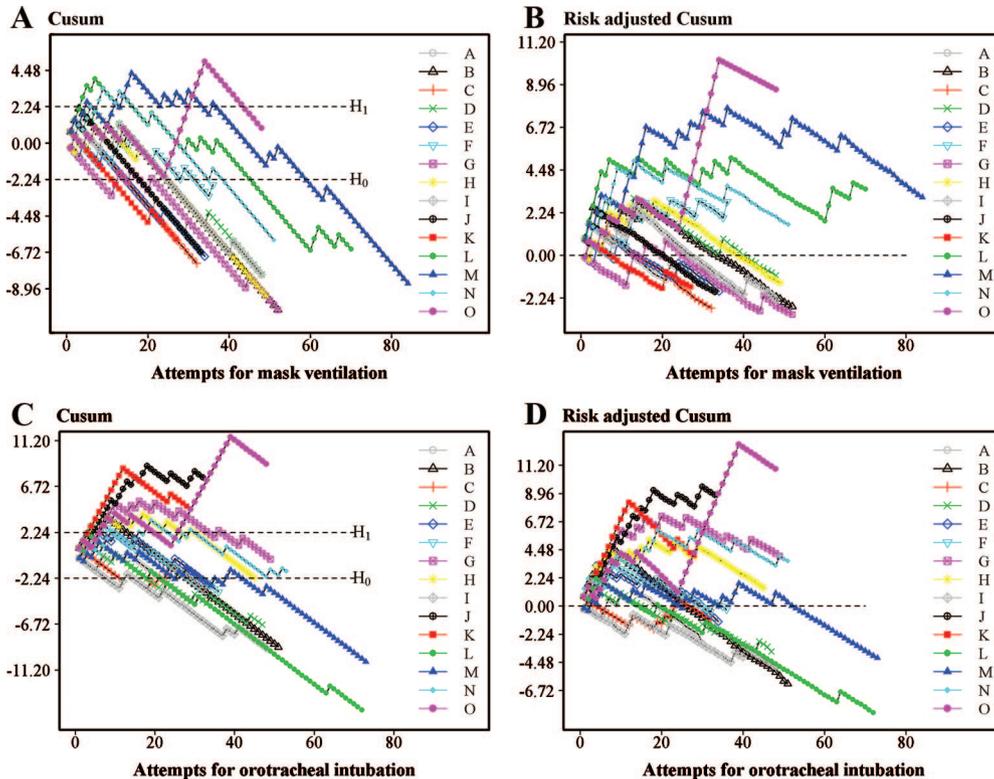
Fig. 1. *A,* Cumulative sum (cusum) chart for bag-and-mask ventilation. Lines A − O represent the learning curves of individual interns. Lines h1 and h0 represent upper and lower control limits of 2.24 and −2.24. Y-axis values represent multiples of h1 and h0. *B,* Risk-adjusted cusum chart for bag-and-mask ventilation. *C,* Cusum chart for orotracheal intubation. Lines A − O represent the learning curves of individual interns. Lines h1 and h0 represent upper and lower control limits 2.24 and −2.24. Y-axis values represent multiples of h1 and h0. *D,* Risk-adjusted cusum chart for orotracheal intubation.

Table 4. Individual Intern Results

Intern	Mask Ventilation			Orotracheal Intubation		
	Successes/ Attempts	Success % (95% CI)	No. Attempts Needed to Cross H0	Successes/ Attempts	Success % (95% CI)	No. Attempts Needed to Cross H0
A	27/31	0.87 (0.70–0.96)	15	26/30	0.87 (0.69–0.96)	22
B	47/52	0.90 (0.79–0.97)	25	45/51	0.88 (0.76–0.96)	29
C	30/32	0.94 (0.79–0.99)	15	26/32	0.81 (0.64–0.93)	12
D	42/48	0.88 (0.75–0.95)	25	40/47	0.85 (0.72–0.94)	18
E	31/34	0.91 (0.76–0.98)	15	27/34	0.79 (0.62–0.91)	32
F	28/36	0.78 (0.61–0.90)	32	29/36	0.81 (0.64–0.92)	29
G	47/52	0.90 (0.79–0.97)	8	35/49	0.71 (0.57–0.83)	Did not cross
H	44/49	0.90 (0.78–0.97)	25	34/45	0.76 (0.60–0.87)	Did not cross
I	42/48	0.88 (0.75–0.95)	25	43/48	0.90 (0.77–0.97)	8
J	30/33	0.91 (0.76–0.98)	18	16/33	0.48 (0.31–0.66)	Did not cross
K	25/27	0.93 (0.76–0.99)	12	15/28	0.54 (0.34–0.72)	Did not cross
L	56/70	0.80 (0.69–0.89)	46	6672	0.92 (0.83–0.97)	18
M	68/84	0.81 (0.71–0.89)	59	62/73	0.85 (0.75–0.92)	25
N	42/51	0.82 (0.69–0.92)	35	39/53	0.74 (0.60–0.85)	Did not cross
O	33/48	0.69 (0.54–0.81)	Did not cross	25/48	0.52 (0.37–0.67)	Did not cross
Summary		Median (95% CI)†				Median (95% CI)†
	592/695	0.85 (0.82–0.88)*	25 (15–32)	528/679	0.78 (0.71–0.84)*	29 (22, N/A)‡

* Ninety-five percent of confidence interval (CI) from generalized estimating equation intercept model adjusting for within-intern correlation. † Kaplan–Meier estimated number of attempts at which 50% of interns cross acceptable failure rate boundary. ‡ Upper limit for CI not estimable from the given data due to small sample size.
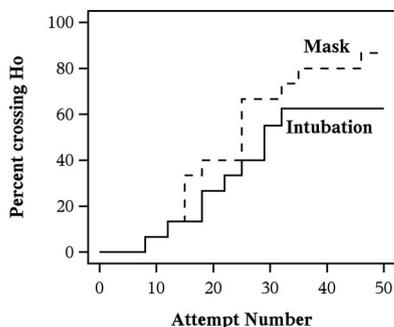
intern (intern "O") crossed the 20% acceptable failure rate line (h0) for mask ventilation (fig. 1A). The 14 successful interns did so after 27 ± 13 (range, 12–59) procedures. The cusum line relative to intern O remained between the decision limit lines after 48 procedures. Using a more stringent 10% specification of p0, no interns crossed the mask ventilation acceptable failure decision limit. Considering all 15 interns, the Kaplan–Meier estimated median (95% confidence interval) time to crossing h0 was 25 (15–32) attempts (fig. 2) using the 20% acceptable failure rate. Pooled success rates before and after reaching h0 under the given criteria were 78% and 94%, respectively ($Z = 5.8$, $P < 0.001$).

Patient risk scores for mask ventilation were calculated from a logistic regression model predicting success from four baseline variables: body mass index more than 30 kg/m², Mallampati 3 or 4 (*vs.* 1, 2), age more than 57 yr, and Mandibular protrusion (severely limited), which have been shown to be associated with difficult mask ventilation. In the

Fig. 2. Mask and intubation ventilation Kaplan-Meier curve for attempt number to crossing h0 for given choices of p0 (0.20), p1 (0.40), type 1 (0.05), and type 2 (0.10) errors.

risk-adjusted observed − expected learning curves for mask ventilation (fig. 1B), 10 interns (66%) finished their attempts below the zero line and, thus, on average performed better than expected, given the level of difficulty of the patients they encountered. This figure is quite different in shape to the standard cusum in figure 1A for most interns, reflecting the distinct cusum formulae (and interpretations) and the varying levels of patient difficulty experienced across interns.

Orotracheal Intubation

Tracheal intubation was attempted on 679 patients, and 528 (78%) of these were successful within two attempts (table 4). An instructor took over 151 (22%) procedures. Among all patients, 16 had rapid sequence inductions. For these patients, intubation data were included in the cusum calculation. There were 45 ± 13 (range, 28–72) procedures per intern. Nine of 15 interns (60%) crossed the 20% acceptable failure rate line (h0). These nine did so after 26 ± 8 (range, 15 to 42) procedures. Cusum lines relative to interns G, H, and N remained between decision limit lines after 49, 45, and 53 procedures, respectively, whereas lines for interns J, K, and O remained greater than the 40% unacceptable failure rate limit (h1) after 33, 28, and 48 procedures, respectively (fig. 1C). Considering all 15 interns, the Kaplan–Meier estimated median (95% confidence interval) time to crossing h0 with the 20% acceptable failure rate was 29 (22, upper limit inestimable) attempts (fig. 2). Pooled success rates before and after reaching h0 were 71% and 90%, respectively ($Z = 4.3$, $P < 0.001$). Under a more stringent specification of p0 = 10%, only one intern crossed the intubation acceptable failure rate decision limit.

Komatsu *et al.*

Patient risk scores for orotracheal intubation were calculated from a logistic regression model predicting success from seven baseline variables, which have been shown to be associated with difficulty, including the four used for mask ventilation plus mouth opening less than 3 cm, thyromental distance less than 6 cm and neck movement less than 90°. Using the risk-adjusted O − E cusum, nine interns (60%) remained below the zero line at the end of their intubation attempts, indicating better-than-expected performance given the level of difficulty assigned to their patients (fig. 1D).

The proportion of interns who attained an acceptable failure rate for mask ventilation (93%) was greater than those of tracheal intubation (60%), $P = 0.025$ (McNemar's test).

Discussion

By using a simple graphical method for detecting the number of successes and failures on a sequential basis, Lawler *et al.*[6] suggested that 20 consecutive, successful tracheal intubations might be appropriate for a student to perform solo anesthesia. However, this approach does not allow statistical inference. Three additional studies examined the learning process for tracheal intubation using statistical approaches. Kopacz *et al.*,[5] using the pooled cumulative success rate at groups of five attempts, demonstrated that a 91% tracheal intubation success rate was achieved after 45 attempts. Konrad *et al.*,[4] using a least-square fit model and Monte Carlo procedures, demonstrated that 90% success rates for tracheal intubation were achieved after 57 attempts. However, these studies did not report the number of attempts at procedures corresponding to certain percentiles of the subjects attaining acceptable failure rate on a time-to-event curve. de Oliveira Filho,[2] using the same statistical approach as our study, demonstrated that four of seven participants attained a 20% acceptable failure rate at intubation, and they did so after mean of 43 attempts.

The cusum method consists of relatively simple calculations, and statistical inference can be made from the observed successes and failures. The method nonetheless requires several considerations. For example, the criteria for success and failure must be clearly defined and represented by a binary variable. Our criteria for successful tracheal intubation (allowing two laryngoscopy attempts and permitting the instructor to provide external laryngeal pressure) was less stringent than those of de Oliveira Filho,[2] who required successful intubation after a single laryngoscopy without any physical assistance, or Konrad *et al.*,[4] who required a successful procedure within two attempts without physical assistance. However, we consider it appropriate for interns to request external laryngeal pressure to optimize the laryngoscopic view and regard it as a reflection of their understanding of upper airway anatomy. Thus, we consider our criteria of successful intubation to be clinically relevant. Furthermore, even experienced intubators sometimes require a second attempt; we, thus, do not consider a second attempt to be an overall failure to intubate.

The standard cusum method also requires specification of an acceptable failure rate (p0). The acceptable rate may be taken from actual institutional rates, published studies, or expert consensus.[3,4,7,18] In this study, acceptable failure rate for tracheal intubation was set at twice the maximum reported incidence of suboptimal laryngoscopic view for intubation (*i.e.*, Cormack and Lehane grades 3 and 4).[11–14] This rather generous acceptable failure rate was chosen because our interns were nonanesthesiologists and presumably would only occasionally be required to manage airways. This acceptable failure rate is also used in a previously performed similar study by de Oliveira Filho,[2] making it possible to directly compare our interns' performances with anesthesia residents from another site.

As seen in table 1, the recommended number of attempts per individual required to adequately assess an individual's learning curve using the cusum method depends on the specified parameters: namely type 1 and 2 errors and the acceptable and unacceptable failure rates. Our study provided acceptable statistical power for our *a priori* chosen 20% limit, but would have been underpowered for a stricter 10% limit.

Using our chosen parameters, an average of 19 attempts per intern would normally be required to reach the acceptable failure rate threshold. We far surpassed this with a mean (range) of 46 (27–84) for mask ventilation and 45 (28–72) for intubation procedures. An acceptable failure rate of 10% would have required an average of 48 attempts per intern, which is close to our average. Many of our interns had less than 48 attempts, and we would have been underpowered to assess the 10% rate.

Given the data, we observed that only one intern would have crossed the h0 decision limit under a 10% acceptable failure rate for intubation and none for mask ventilation.

At tracheal intubation, 60% of the interns crossed the acceptable (lower) decision limit after 26 ± 8 procedures. Although we used the same statistical method and same acceptable and unacceptable failure rates, our interns crossed the lower decision limit after a smaller number of procedures than those in the study by de Oliveira Filho.[2] This may be explained by the difference in the criteria of successful intubation between the two studies. De Oliveira Filho[2] allowed only one attempt at laryngoscopy with no physical assistance for intubation to be considered successful; in contrast, we allowed two laryngoscopy attempts and physical assistance in the form of external laryngeal pressure, which perhaps better reflects how intubation is normally approached.

We did not have an adequate number of interns to directly compare the number of attempts to cross the acceptable failure rate decision limit (*i.e.*, h0) with mask and intubation. We nonetheless observed a difference in the proportion of interns who crossed the decision limit for the respective methods: with mask ventilation, 93% of our interns crossed the lower decision limit after 27 ± 13 procedures, which was significantly greater than the proportion for intubation, which was 60%. Median time to cross h0 with a 20% acceptable failure rate was similar between mask venti-

lation and intubation, but intubation can be a technique associated with more interindividual variability in initial skill acquisition than mask ventilation. Although the higher proportion of interns crossed the acceptable failure rate decision limit for mask ventilation than intubation, for those infrequently managing the patient's airway, not only ease of skill acquisition but also ease of skill maintenance is of clinical importance. A previous report showed occasional performance of intubation does not ensure skill maintenance.[19] No comparable reports for mask ventilation are available, and the evaluation of skill maintenance in mask ventilation is warranted. In the current study, many patients were paralyzed, which facilitates mask ventilation. Per protocol, patients who had difficult airways were excluded because they would have been designated for awake intubation. Therefore, caution needs to be used to apply these data to the urgent floor airway situation where patients are not fasting, bed position is suboptimal, and an efficient respiratory circuit of an anesthesia machine is not available.

The standard cusum method is limited in that it does not allow weighting of the cusum score according to the expected difficulty at each procedure. Therefore, we also calculated risk-adjusted learning curves in the form of observed − expected (O − E) cusum charts with patient variables shown in the literature to be related to difficult mask ventilation and orotracheal intubation. For the O − E cusum method, interns whose performance reflects the average across interns for patients of similar difficulty will have charts with values close to the zero line. This risk adjustment uses the average success rate for interns across the entire learning curve as the "expected" success rate for a patient. Ideally, the expected scores would be derived from data external to the study at hand. Such external data might also make attractive more sophisticated forms of risk adjustment, such as the log-likelihood method,[20] where a baseline probability of failure is specified for particular patients or groups of patients, and the curve monitors *a priori* reduction or increase from baseline.

In summary, participating interns achieved 80% mask ventilation success rate faster than 80% orotracheal intubation success rate, and there was more variability in the rate at which intubation skills were developed. The achievement of 80% orotracheal intubation success rate required a median of 29 procedures.

References

1. Cheney FW, Posner KL, Lee LA, Caplan RA, Domino KB: Trends in anesthesia-related death and brain damage: A closed claims analysis. ANESTHESIOLOGY 2006; 105:1081–6
2. de Oliveira Filho GR: The construction of learning curves for basic skills in anesthetic procedures: An application for the cumulative sum method. Anesth Analg 2002; 95:411–6
3. Kestin IG: A statistical approach to measuring the competence of anaesthetic trainees at practical procedures. Br J Anaesth 1995; 75:805–9
4. Konrad C, Schupfer G, Wietlisbach M, Gerber H: Learning manual skills in anesthesiology: Is there a recommended number of cases for anesthetic procedures? Anesth Analg 1998; 86:635–9
5. Kopacz DJ, Neal JM, Pollock JE: The regional anesthesia "learning curve." What is the minimum number of epidural and spinal blocks to reach consistency? Reg Anesth 1996; 21:182–90
6. Lawler PG, Patla VR, Garcia E, Puttick N: Assessment of training in anaesthesia and related skills. Anaesthesia 1991; 46:597
7. Schuepfer G, Johr M: Generating a learning curve for penile block in neonates, infants and children: An empirical evaluation of technical skills in novice and experienced anaesthetists. Paediatr Anaesth 2004; 14:574–8
8. Wilson ME: Assessing intravenous cannulation and tracheal intubation training. Anaesthesia 1991; 46:578–9
9. Davies O, Goldsmith PL: Statistical Methods in Research and Production. London, Longman, 1976
10. Van Rij AM, McDonald JR, Pettigrew RA, Putterill MJ, Reddy CK, Wright JJ: Cusum as an aid to early assessment of the surgical trainee. Br J Surg 1995; 82:1500–3
11. Asai T, Koga K, Vaughan RS: Respiratory complications associated with tracheal intubation and extubation. Br J Anaesth 1998; 80:767–75
12. Langeron O, Masso E, Huraux C, Guggiari M, Bianchi A, Coriat P, Riou B: Prediction of difficult mask ventilation. ANESTHESIOLOGY 2000; 92:1229–36
13. Koh LK, Kong CE, Ip-Yam PC: The modified Cormack-Lehane score for the grading of direct laryngoscopy: Evaluation in the Asian population. Anaesth Intensive Care 2002; 30:48–51
14. el-Ganzouri AR, McCarthy RJ, Tuman KJ, Tanck EN, Ivankovich AD: Preoperative airway assessment: Predictive value of a multivariate risk index. Anesth Analg 1996; 82:1197–204
15. Sibanda T, Sibanda N: The CUSUM chart method as a tool for continuous monitoring of clinical outcomes using routinely collected data. BMC Med Res Methodol 2007; 7:46
16. Kheterpal S, Han R, Tremper KK, Shanks A, Tait AR, O'Reilly M, Ludwig TA: Incidence and predictors of difficult and impossible mask ventilation. ANESTHESIOLOGY 2006; 105:885–91
17. Shiga T, Wajima Z, Inoue T, Sakamoto A: Predicting difficult intubation in apparently normal patients: A meta-analysis of bedside screening test performance. ANESTHESIOLOGY 2005; 103:429–37
18. Williams SM, Parry BR, Schlup MM: Quality control: An application of the cusum. BMJ 1992; 304:1359–61
19. Bishop MJ, Michalowski P, Hussey JD, Massey L, Lakshminarayan S: Recertification of respiratory therapists' intubation skills one year after initial training: An analysis of skill retention and retraining. Respir Care 2001; 46:234–7
20. Steiner SH, Cook RJ, Farewell VT: Risk-adjusted monitoring of binary surgical outcomes. Med Decis Making 2001; 21:163–9