Can J Anesth/J Can Anesth (2013) 60:771–779
DOI 10.1007/s12630-013-9974-y

REPORTS OF ORIGINAL INVESTIGATIONS

Anesthesiologists' learning curves for bedside qualitative ultrasound assessment of gastric content: a cohort study

Les courbes d'apprentissage des anesthésiologistes pour l'évaluation du contenu gastrique par échographie qualitative au chevet: une étude de cohorte

Cristian Arzola, MD · Jose C. A. Carvalho, MD, PhD · Javier Cubillos, MD · Xiang Y. Ye, MSc · Anahi Perlas, MD

Received: 24 January 2013 / Accepted: 13 May 2013 / Published online: 24 May 2013

Abstract

Purpose *Focused assessment of the gastric antrum by ultrasound is a feasible tool to evaluate the quality of the stomach content. We aimed to determine the amount of training an anesthesiologist would need to achieve competence in the bedside ultrasound technique for qualitative assessment of gastric content.*

Methods *Six anesthesiologists underwent a teaching intervention followed by a formative assessment; then learning curves were constructed. Participants received didactic teaching (reading material, picture library, and lecture) and an interactive hands-on workshop on live models directed by an expert sonographer. The participants were instructed on how to perform a systematic qualitative assessment to diagnose one of three distinct categories of gastric content (empty, clear fluid, solid) in healthy volunteers. Individual learning curves were constructed using the cumulative sum method, and competence was defined as a 90% success rate in a series of ultrasound examinations. A predictive model was further developed based on the entire cohort performance to determine the number of cases required to achieve a 95% success rate.*

Results *Each anesthesiologist performed 30 ultrasound examinations (a total of 180 assessments), and three of the six participants achieved competence. The average number of cases required to achieve 90% and 95% success rates was estimated to be 24 and 33, respectively.*

Conclusion *With appropriate training and supervision, it is estimated that anesthesiologists will achieve a 95% success rate in bedside qualitative ultrasound assessment after performing approximately 33 examinations.*

Résumé

Objectif *L'évaluation ciblée de l'antre pylorique par échographie est un outil utile pour estimer la qualité du contenu de l'estomac. Notre objectif était de déterminer la quantité de formation nécessaire à un anesthésiologiste afin d'acquérir des compétences en échographie au chevet pour l'évaluation qualitative du contenu gastrique.*

Méthode *Six anesthésiologistes ont suivi une formation, puis une évaluation formative; des courbes d'apprentissage ont ensuite été élaborées. Les participants ont reçu un enseignement didactique (matériel de lecture, bibliothèque d'images, et cours) et suivi un atelier pratique interactif sur*

Author contributions *Cristian Arzola, Jose C.A. Carvalho, Javier Cubillos, Xiang Y. Ye,* and *Anahi Perlas* helped design the study. *Cristian Arzola, Jose C.A. Carvalho, Javier Cubillos,* and *Anahi Perlas* helped conduct the study. *Cristian Arzola* and *Xiang Y. Ye* helped analyze the data. *Jose C.A. Carvalho, Javier Cubillos,* and *Anahi Perlas* helped with the interpretation of data. *Cristian Arzola* and *Jose C.A. Carvalho* helped write the manuscript. *Javier Cubillos, Xiang Y. Ye,* and *Anahi Perlas* helped revise the manuscript. *Cristian Arzola, Jose C.A. Carvalho, Javier Cubillos, Xiang Y. Ye,* and *Anahi Perlas* have seen the original study data and reviewed the analysis of the data. *Cristian Arzola* is the author responsible for archiving the study files.

C. Arzola, MD (✉) · J. C. A. Carvalho, MD, PhD ·
J. Cubillos, MD
Department of Anesthesia and Pain Management, Mount Sinai
Hospital and University of Toronto, 600 University Avenue,
Room 19-104, Toronto, ON M5G 1X5, Canada
e-mail: carzola@mtsinai.on.ca

X. Y. Ye, MSc
Maternal-Infant Care Research Centre, Mount Sinai Hospital,
Toronto, ON, Canada

A. Perlas, MD
Department of Anesthesia, Toronto Western Hospital, University
Health Network and University of Toronto, Toronto ON, Canada

des modèles vivants dirigé par un échographiste expert. On a enseigné aux participants la façon de réaliser une évaluation qualitative systématique afin de diagnostiquer l'une de trois catégories distinctes de contenu gastrique (vide, liquide clair, solide) chez des volontaires sains. Les courbes d'apprentissage individuelles ont été construites sur la base de la méthode des sommes cumulées. La compétence était définie en tant que taux de réussite de 90 % dans une série d'examens échographiques. De plus, un modèle de prévision a été créé sur la base de la performance de la cohorte entière.

Résultats *Chaque anesthésiologiste a réalisé 30 examens échographiques (soit un total de 180 évaluations), et trois des six participants ont atteint le niveau de compétence prédéfini. Le nombre moyen de cas requis pour atteindre des taux de réussite de 90 % et 95 % a été estimé à 24 et 33, respectivement.*

Conclusion *Avec une formation et une supervision adaptées, on estime que les anesthésiologistes atteindront un taux de réussite de 95 % des évaluations échographiques qualitatives au chevet après avoir réalisé environ 33 examens.*

Pulmonary aspiration of gastric content is considered a major complication in anesthesia practice with resulting morbidity and mortality.[1-4] Certain factors have been linked to the severity of patient outcomes, including the volume, nature, and acidity of the aspirate.[5-8] The preoperative assessment of the risk of pulmonary aspiration relies essentially on the patient's history, and the clinical management typically adheres to fasting recommendations of current guidelines.[9] The ultimate assessment of the nature and volume of gastric content at bedside remains inaccessible to the anesthesiologist.

Recent studies have shown that bedside ultrasonography can provide reliable information about the nature (clear fluid, solid, or none) and volume of gastric content.[10-14] In these studies, either a certified sonographer or a single anesthesiologist with no specified previous training performed all the examinations.[11-13,15-17] The process of knowledge translation demands further research to confirm generalizability of these findings. The amount of training required to achieve competence in the performance of gastric ultrasonography as a bedside clinical tool remains unknown. The aim of this study was to determine the amount of training an anesthesiologist would need to achieve competence in the bedside ultrasound technique for qualitative assessment of gastric content in healthy volunteers.

Methods

Following approval by the Research Ethics Board (ref: 11-0237-E) of Mount Sinai Hospital (Toronto, Canada) on October 14, 2011, we conducted this observational cohort study of a group of anesthesiologists learning qualitative ultrasound assessment of gastric content in healthy volunteers along with the reference standard of an expert sonographer. The study was carried out from October 2011 to December 2011. Written informed consent was obtained from all participants (anesthesiologists, volunteers, and expert sonographer). We followed the guidelines for reporting and the applicable checklist from the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) Statement.[18]

Eligible participants were advanced level trainees (anesthesia fellows during one year of subspecialty training) or staff anesthesiologists. Participants had previous experience in ultrasound-assisted/guided procedures, but not in gastric ultrasound.

Inclusion criteria for volunteers were: age 18 yr or older, weight 50-120 kg, and height 150 cm or greater. Volunteers with a pre-existing abnormal anatomy of the upper gastrointestinal tract (previous esophageal, gastric, or upper abdomen surgery and hiatus hernia) were excluded. A certified sonographer experienced in diagnostic abdominal ultrasonography with more than 500 examinations of gastric content was considered the reference standard for the ultrasonographic assessments. The cohort of anesthesiologists underwent a teaching intervention which was followed by a formative assessment process in order to construct learning curves.

Teaching intervention

Initially, the cohort of anesthesiologists received *didactic teaching* in the form of reading material, educational video/picture library, and a 30-min lecture on the topic. The reading material included research studies[11-14] and one educational article with a description of the technique.[19] Subsequently, the teaching proceeded with an *interactive* three-hour session consisting of a hands-on workshop and a live demonstration performed by the expert sonographer. Five volunteers were scanned after fasting for at least eight hours, and they were later scanned after ingesting either clear fluids (300 mL of apple juice) or a standardized solid meal (muffin and apple juice). Each anesthesiologist received one-on-one feedback on ten exams that included every prandial status (fasted, clear fluids, and solid meal). During the workshop, the anesthesiologists were instructed on how to perform a systematic ultrasound assessment in order to make a qualitative diagnosis of three distinct gastric content categories (empty, clear fluid, solid).

The standardized technique was performed with a portable ultrasound system equipped with a 5-2 MHz curved array transducer (M-Turbo™, SonoSite® Canada Inc., Markham, ON, Canada). Volunteers were first placed

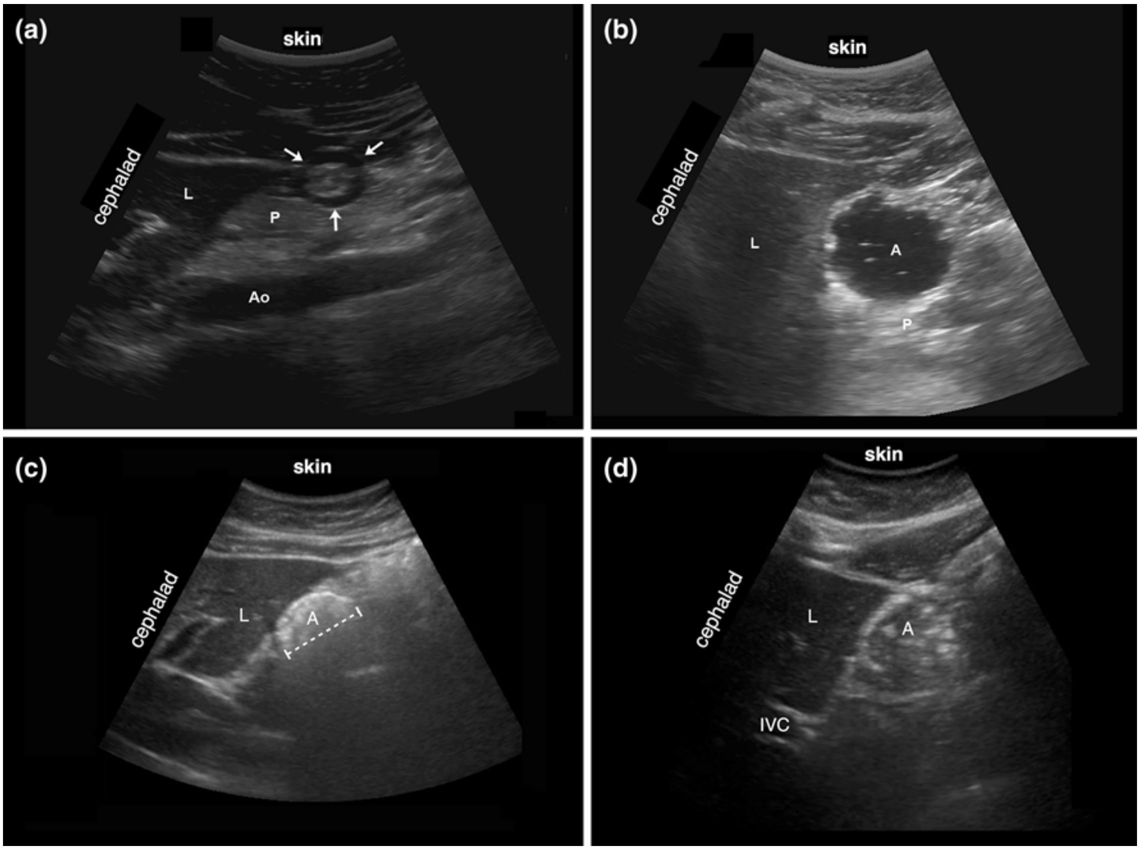

Fig. 1 Sonographic images of the epigastric area (parasagittal plane). L = liver; A = antrum; P = pancreas; IVC = inferior vena cava; Ao = aorta. (a) Empty gastric antrum. The arrowheads denote an empty gastric antrum. The "bull's eye" appearance is typical if the antrum is contracted. (b) Gastric antrum after 300 mL of oral fluid intake. The "starry night appearance" is caused by air mixed with fluid. (c) Gastric antrum with a "frosted glass appearance" secondary to air after recent ingestion of a solid meal. (d) Gastric antrum with heterogeneous granular appearance of ongoing digestion

in a 45° semi-recumbent position and then in a right lateral decubitus position. In either of these positions, fluid or semi-fluid content gravitates preferentially to the antrum, and air or gas is displaced proximally towards the body or fundus, thus facilitating antral sonography.[11] Both positions are part of the same continuous diagnostic process, and we did not consider them two different diagnostic modalities (Fig. 1). A detailed description of the ultrasound technique and ultrasonography characteristics of gastric antrum content has been reported recently.[19]

Assessment process

The formative assessment process consisted of a series of ultrasound examinations performed by each anesthesiologist followed by confirmation and feedback from the expert sonographer. We scheduled six sessions with five volunteers per session, with 30 separate sets of data per anesthesiologist. For each evaluation session, five study volunteers were randomly allocated to one prandial status similar to that previously described in the hands-on workshop (eight-hour fasting, clear fluids, and solid meal). Randomization allocation was in two blocks of 15. Each block of 15 included the three gastric content categories: five empty, five clear fluids, and five solid. The allocation concealment comprised of sequentially numbered and sealed opaque envelopes prepared by an independent research assistant. At each session, the ultrasound examinations were started five minutes after the ingestion of liquid or solid, and the assessment session was paused every 15 min to provide a top-up of clear fluids (100 mL) to the fluid group, which was also blinded. The purpose of the top-up was to attempt to maintain equal examination conditions due to possible rapid gastric emptying of clear fluids in a fasted volunteer. The expert sonographer and the anesthesiologists were blinded to the volunteer prandial status, and they examined all five volunteers in a random sequence. Additionally, each of the six sessions was scheduled one week apart (six-week assessment period), and each volunteer was scanned only once per session by each anesthesiologist. The prandial status was disclosed only when the examination of the five volunteers was completed by all anesthesiologists and the sonographer. The expert sonographer then conducted a 50-min feedback session, including review of images and discussion and reinforcement of the technical aspects. In order to maintain uniform learning progress for the entire group, we did not allow participants to practice the ultrasound technique between sessions; however, they were encouraged to review the didactic teaching material.

Learning curves

The aim of the qualitative ultrasound examination was to diagnose three distinct gastric content categories (empty, clear fluid, or solid). Success was defined as a correct diagnosis, which was recorded over time. Although the diagnosis from the expert sonographer was the reference standard by study design, we also followed her performance, as she was also blinded to the volunteers' prandial status. In this way, we were able to show the construct validity of the technique. We aimed to construct learning curves for a series of ultrasound examinations through the cumulative sum method (CUSUM).[20]

Study outcomes

The primary outcome was the number of ultrasound examinations (cases) an anesthesiologist required to achieve competence in the ultrasonography qualitative assessment of gastric content. Competence was defined as 90% success rate in a series of ultrasound examinations.[20-22] The secondary outcomes included the overall success rate of correct ultrasound diagnoses, the specific success rate per gastric content category, and characteristics of the series of ultrasound examinations.

Statistical analysis

In order to estimate the primary outcome for each anesthesiologist, individual learning curves were constructed using the CUSUM graphical method (Appendix 1). The CUSUM graphs display the cumulative differences plotted in sequence, allowing detection of deviations from a predetermined standard.[20-26] A negative trend of the individual CUSUM line indicates success, whereas a positive trend indicates failure. We considered the anesthesiologist competent at a 10% failure rate if the graph passed through two decision lines from above. In this way, early failures did not penalize the anesthesiologist for later increases in accuracy.[25,26]

We calculated a minimum sample size of 13-18 consecutive cases as necessary to determine competence at an acceptable failure rate of 10% and an unacceptable failure rate of 30%, respectively (Appendix 1). We increased the sample to 30 consecutive cases per anesthesiologist (twice the average of calculated cases) to allow an adequate assessment process.

Furthermore, we used the entire cohort data in the construction of a predicted learning curve based on Bush and Mosteller's mathematical learning model, which estimates the average number of cases to achieve a 95% success rate (Appendix 2).[23,27] In this way, we were able to determine an average CUSUM score and construct a learning curve for the average of all participants, allowing interpretations for the entire cohort rather than for each individual alone.

The assessments were also compared with respect to a correct or incorrect diagnosis. We examined the effect of the gastric content category on the correct ultrasound diagnosis. For the comparisons, we used the generalized regression models with a generalized estimating equation approach accounting for the correlated data. Further evaluation of the effect of the gastric content category and the duration of the ultrasound examination was conducted using multiple logistic regression models for correlated data.

Descriptive statistical methods were used to describe the study population. The statistical analyses were performed using SAS® 9.2 (SAS Institute Inc., Cary, NC, USA), R 10.2 (http://www.r-project.org/), and STATA® for Macintosh, Release 12.1 (StataCorp, College Station, TX, USA). A two-sided significance level of < 0.05 was used without multiple comparison adjustment.

Results

The cohort of anesthesiologists consisted of four anesthesia fellows and two anesthesia staff. All participants completed the teaching and evaluation sessions. The group of healthy male volunteers had a mean (standard deviation = SD) age of 25 (2) yr, weight of 71 (12) kg, height of 173 (4) cm, and body mass index of 25 (3) $\text{kg}\cdot\text{m}^{-2}$.

Learning curves

Each of the six anesthesiologists performed 30 ultrasound examinations for a total of 180 assessments, and the blinded sonographer made the correct diagnosis in every case. The analysis of learning curves through CUSUM showed that three of the six anesthesiologists (two staff and one anesthesia fellow) achieved competence (90% success rate) with a median number of 24 cases (15, 24, and 28 cases) (Fig. 2). The average CUSUM curves were obtained for those who achieved and those who did not achieve competence in the series of 30 cases (Fig. 3A). The Bush and Mosteller's learning model predicted 33 as the average

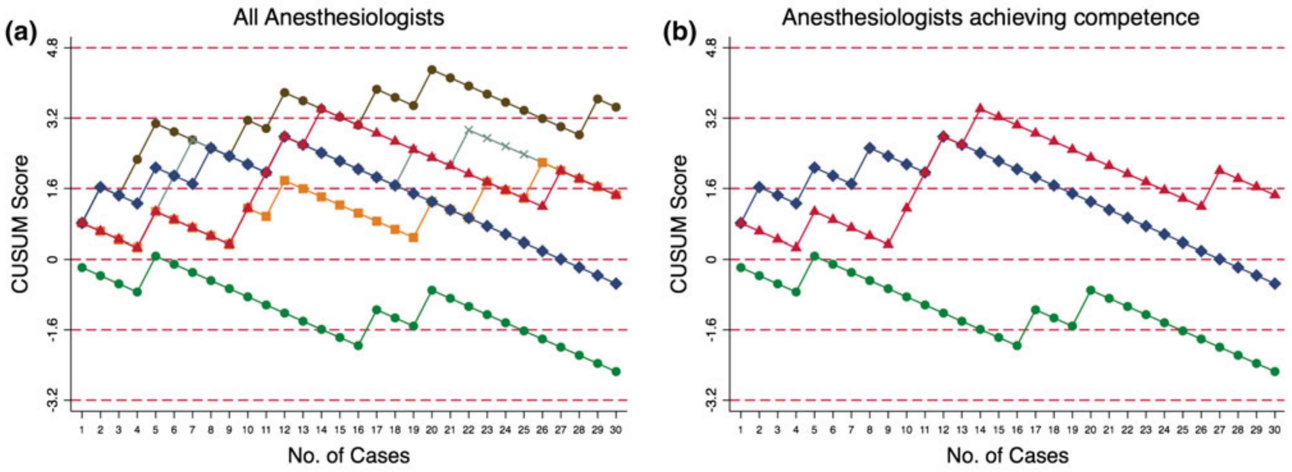

Fig. 2 Learning curves by CUSUM (cumulative sum) graphs. Connected plotted lines represent learning curves for individual anesthesiologists. Upward and downward trend indicates failure and success, respectively. Horizontal red dashed lines represent decision limits. (a) Learning curves of all anesthesiologists. (b) Learning curves of anesthesiologists achieving competence

number of cases required to achieve a 95% success rate (Fig. 3B).

Secondary outcomes

The overall success rate (correct diagnosis) was 79% (142/180) and the failure rate (incorrect diagnosis) was 21% (38/180). The relative proportions of the gastric content categories in the groups of correct diagnosis (success) and incorrect diagnosis (failure) are presented in Table 1. When comparing the three gastric content categories (Table 2), the "clear fluid" group appeared with a significantly lower failure rate than "empty" ($P < 0.05$), while the "solid" group showed a lower failure rate compared with both "empty" and "clear fluid". The duration of the examination for the "clear fluid" was shorter than for the "empty" and "solid" groups ($P < 0.05$).

Furthermore, when comparing success (correct diagnosis) using multivariable analyses, the odds of success for "solid" was 2.6 times that for "empty", and 1.7 times that for "clear fluid". Finally, the odds of success decreased for any one-minute increase in the ultrasound examination (Table 3).

Discussion

This study determined the learning curves for the qualitative ultrasonography assessment of gastric contents performed by a cohort of anesthesiologists. Under the study conditions, CUSUM analysis revealed that 24 consecutive ultrasound examinations were required to achieve competence with a 90% success rate, and at least 33 examinations would be required for a 95% success rate based on a predicted model. The fact that the expert sonographer correctly diagnosed the gastric content in every volunteer confirms the construct validity of the technique. Nevertheless, according to CUSUM analysis, only three of six novice participants achieved competence in the series of 30 examinations. This may indicate a medium level of complexity in learning the technical and cognitive skills necessary to perform this ultrasound diagnostic procedure.

The CUSUM method has been used in many studies investigating the acquisition of competence or assessment of performance in anesthesia technical skills,[22,25,28-31] and it has been used more recently in ultrasound-guided procedures.[21,23,24,26,32] It can be used to show proficiency in a newly learned technical skill or measure quality once a technical skill has been mastered.[33] In the context of medical education, it is crucial to understand CUSUM as a statistical method that looks at the outcome rather than the process of performing procedural skills.[20,28] Although we focused on the *formative assessment* (training, feedback, discussion) aspect in our study, achieving and declaring competence still reflects the goal of *summative assessment*.[34-36] Defining competence is a complex task. Acceptable rates of success can be determined by institutional rates or expert consensus, but they also depend on the definition of success.[20,30] Currently, there is no standard for this new diagnostic skill in the anesthesia setting. In some studies on newly acquired procedural technical skills in anesthesia, 90% success rates have been used based on previous training studies[21] or departmental consensus using a modified Delphi approach.[22] Therefore, for this study, we decided to define competence as a 90% success rate as determined by consensus of opinion among

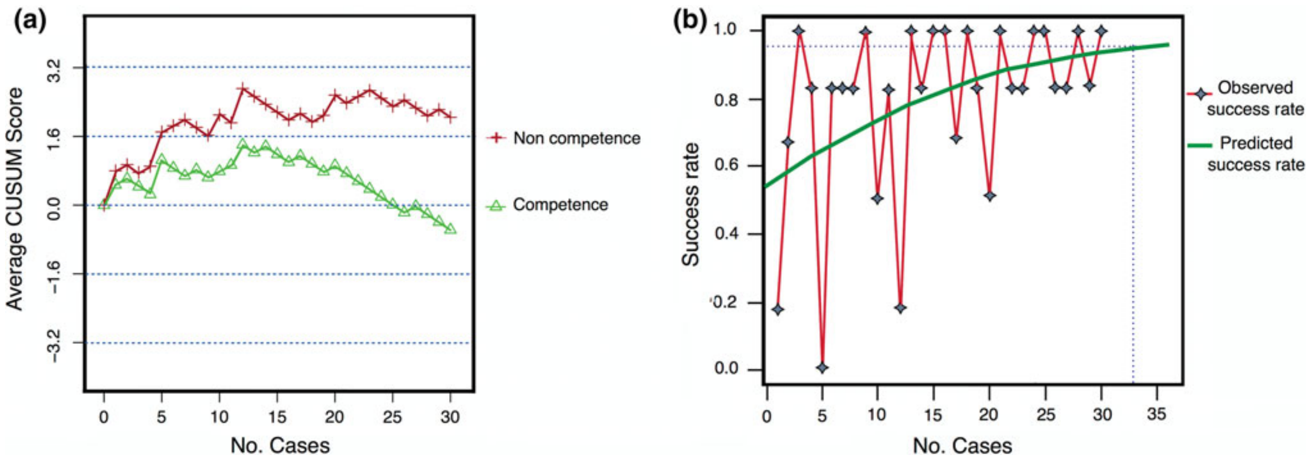

Fig. 3 (a) Average CUSUM (cumulative sum) learning curves for those achieving competence (red cross, connected-plotted line) and those not achieving competence (green triangle, connected-plotted line). (b) The observed success rates (red connected line) and the predicted success rates (green line) at each case using the Bush and Mosteller's learning model. The predicted 95% (blue dotted line) success rate is attained after an average of 33 cases

Table 1 Incidence of correct (Success) and incorrect diagnosis (Failure) according to the gastric content category and the duration of the ultrasound examination

	Success $n = 142$	Failure $n = 38$	*P* value
Gastric content category *n* (%)			0.062
Empty	42 (30)	18 (47)	
Clear fluid	49 (34)	11 (29)	
Solid	51 (36)	9 (24)	
Duration (min)	2.3 (1.4 to 3.1)*	3.9 (2.1 to 5.7)*	0.085

n = number of attempts; Duration = duration of entire ultrasound examination; *Mean (95% confidence intervals); *P* value was based on a generalized regression model using generalized estimating equation methods accounting for the correlated data

Table 2 Diagnostic failure among the three gastric content categories and duration of the ultrasound examination

	Gastric content category		
	Empty $n = 60$	Clear Fluid $n = 60$	Solid $n = 60$
Failure *n* (%)	18 (30)†	11 (18)†	9 (15)
Duration, min*	3.3 (2.1 to 4.5)†	1.3 (0.3 to 2.3)†‡	3.3 (1.8 to 4.8)‡

The values in two groups sharing same symbol (‡,†) are significantly different ($P < 0.05$). Duration = duration of entire ultrasound examination. *Mean (95% confidence intervals)

Table 3 Multivariable analysis: odds of success according to gastric content categories

	Adjusted OR	(95% CI)
Solid *vs* Empty*	2.6*	(1.04 to 6.6)
Solid *vs* Clear Fluid	1.7	(0.6 to 4.8)
Clear Fluid *vs* Empty	1.5	(0.6 to 3.8)
Duration†	0.9	(0.8 to 1.1)

A multiple conditional logistic regression model for correlated data was used. OR = odds ratio; CI = confidence interval; *$P < 0.05$; †One minute increase

the investigators who had performed previous research on ultrasound assessment of gastric content.

If we had followed the method of pooled cumulative success rate described by Kopacz *et al.*,[37] all participants could have been assessed by their actual overall success rate, which was greater than 70% for all six anesthesiologists and corresponds to the predefined cut-off point of 30% unacceptable failure rate. Nevertheless, the CUSUM method tends to benefit consistency over a sequence of cases, and its value comes from providing statistical inference while testing the failure rate against a minimum standard. Furthermore, if we assume that all anesthesiologists had the same baseline experience, we can use the summated changing success rates to develop a learning curve for the average of all participants until a certain target rate is achieved in this outcome-based task. In this way, the predicted model facilitates making interpretations for a population rather than for each individual alone, providing potential application for curriculum development, practice rotations, and mentor evaluations.[38]

We planned a study protocol with very distinct gastric content categories to enhance the early stages of learning. Among the three gastric content categories, study results suggest that "solid" content appears to be more easily diagnosed by novice sonographers, followed by "clear fluid", and then the "empty" category. Additionally, the "clear fluid" is more readily appreciated. We managed to optimize the clear fluid content group in order to avoid the possible confounding factor of rapid gastric emptying. The antrum filled with clear fluid is more obvious to the observer, appearing round and filled with hypoechoic/anechoic content (Fig. 1B). Nevertheless, the “solid” content group appeared to have been diagnosed more easily than both the “clear fluid” and the “empty” groups. This may be related to the relatively larger size of the antrum and the characteristic hyperechoic mucosal-air interface in cases of recent ingestion of solid food that depicts a “frosted glass” appearance, impairing visualization of deeper structures (Fig. 1C).[19] Further research is warranted to determine the particular challenges and difficulty in the diagnosis of various gastric contents.

Our study has some limitations. First, the quantitative dimension (volume assessment) was beyond the scope of this study. This aspect of the ultrasound assessment is not required in the diagnosis of “empty stomach”. Nevertheless, if the gastric sonogram should show clear fluid content, a quantitative volume assessment can help differentiate a low-volume status (similar to baseline physiologic gastric content) from a high-volume status (higher than baseline gastric volume).[11-13] We concentrated instead on using a focused qualitative exam as an approach to a screening test. Second, although we assessed quite distinct categories of gastric content and performed the ultrasound assessments at fairly constant time points, scanning conditions for the individual participants could have varied during the course of the session. Third, our five volunteers were healthy young males with a normal body mass index; while we believe that this was essential for reliable and valid training, results may not be fully applicable to other populations. Although we had a limited number of volunteers and bias was a concern, we planned to minimize such bias by carrying out each of the six sessions one week apart (six-week assessment period), and each volunteer was scanned only once per session by each anesthesiologist. Fourth, some of the predefined statistical variables of CUSUM in the study design may have limited participants’ achievement of competence. Specifically, the learning curves were defined by the level to declare competence, the number of participants, and the individual attempts per participant. Finally, this cohort of anesthesiologists had previous experience in other applications of ultrasound for diagnostic and interventional procedures in anesthesia. This baseline condition could have facilitated the development of their learning curves and could also make conclusions more suitable for this particular population.

The use of ultrasound by anesthesiologists as a perioperative point-of-care resource has dramatically increased over the last ten years, and many current best practices entail ultrasound to guide clinical decision-making and procedures. A growing body of evidence shows the benefits of this change in practice.[39] In this context, bedside ultrasound could potentially become a clinically useful noninvasive tool for accurate determination of gastric content and volume. It could help to establish the risk of perioperative aspiration more precisely, both in a given individual and in different patient populations.[9,12,14,40] Practice guidelines are subject to revision as warranted by the evolution of medical knowledge, technology, and practice.[9,41] Factors such as current literature, expert and practitioner opinion, open forum commentary, and clinical feasibility data support these recommendations. Presently, it may be early and premature to anticipate how this new technology could affect practice, although we believe it is a promising clinical skill to pursue.

In conclusion, we have described the feasibility of appropriate training of anesthesiologists to reach a level of competence in performing bedside qualitative ultrasound assessment of gastric content. Individual learning curves and a predictive model suggest that 24 and 33 are the average numbers of cases required to achieve 90% and 95% success rates, respectively.

Acknowledgements The authors thank Liisa Davis B.Sc., Sonographer, Regional Anesthesia Research Group, Toronto Western Hospital, University Health Network, Toronto, Ontario, Canada, for her invaluable technical assistance in training and subject examinations. The authors also thank Kristi Downey MSc, Perinatal Research Coordinator, Department of Anesthesia and Pain Management, Mount Sinai Hospital, for her invaluable contribution to all stages of this project.

Funding This work was supported by departmental resources.

Conflict of interest Dr. Anahi Perlas is Associate Editor of the journal, *Regional Anesthesia and Pain Medicine.* Dr. Anahi Perlas receives support for academic time from a University of Toronto, Department of Anesthesia Merit Award 2011-2013. None of the other authors have financial or personal relationships or affiliations that could influence (or bias) this research.

Appendix 1: Cumulated sum (CUSUM)

Parameters for the construction of a CUSUM graph are the acceptable (*p0*) and the unacceptable (*p1*) failure rates and probabilities of type I and II errors (*a* and *b*). We set the acceptable standard of 10% failure rate, the unacceptable standard of 30% failure rate, the probability of a type I error (*a*) of 0.10, and the probability of a type II error (*b*) of 0.10. From these, two decision limits (*h1* and *h0*) and the variable s are calculated. Since a and b are equal, $h1 = h0$ and will be referred to as "h". We then drew decision lines at h, $2h$, and $3h$, as required, parallel to the X-axis. The graphs start at zero (CUSUM score), and for each successful case, the amount s is subtracted from the previous CUSUM score. For each failed case, the amount $1 - s$ is added to the previous CUSUM score. Thus, a negative trend of the CUSUM line indicates success, whereas a positive trend indicates a failure at the case being analyzed. We considered the anesthesiologist competent at the 10% failure rate, with the probability of a type II error $= b$ if the graph passed through two decision lines from above. In this way, early failures did not penalize the anesthesiologist for later increases in accuracy.[25,26]

Calculation of Sample Size (number or cases):

1. Average number of cases with acceptable failure rate:

$$[(h0(1-\alpha) - \alpha h1)/(s - p0)]$$

2. Average number of cases with unacceptable failure rate:

$$[(h1(1-\beta) - \beta h0)/(p1 - s)]$$

$a = \ln [(1 - \beta) / \alpha]$
$b = \ln [(1 - \alpha) / \beta]$
$P = \ln (p1/ p0)$
$Q = \ln [(1 - p0) / (1 - p1)]$
$s = Q / (P + Q)$
$h0 = -b / (P + Q)$
$h1 = a / (P + Q)$

Appendix 2

The Bush and Mosteller's learning model for symmetric choices:[23,27]

$$V_{1n} = V_{10}/[V_{10} + (1 - V_{10})\exp\{-(\phi_1 - \phi_2)(1 - \alpha_1)n\}],$$

where V_{1n} is the mean probability of success at trial n; $\phi_1 =$ the expected probability of success; $\phi_2 =$ the expected probability of failure; α_1 and V_{10} are the slope parameter and the average success rate at the initial trials to be estimated respectively. T_1 satisfied the following equation is the mean number of successes at the $N - 1$ trials:$\alpha_1 = 1 - [(\phi_1 - V_{10})/(N\phi_1 - T_1)]$.

The model parameters ϕ_1 (expected probability of success) and ϕ_2 (the acceptable expected probability of failure) were set as 90% and 30%, respectively. The estimated slope parameter of the learning model/curve ($\alpha 1$) was 0.86; the initial probability of success was 0.54. The mean number of successes was 27 at case 29 (30 cases − 1).

References

1. *Tiret L*, *Desmonts JM*, *Hatton F*, *Vourc'h G*. Complications associated with anaesthesia - a prospective survey in France. Can Anaesth Soc J 1986; 33: 336-44.
2. *Warner MA*, *Warner ME*, *Weber JG*. Clinical significance of pulmonary aspiration during the perioperative period. Anesthesiology 1993; 78: 56-62.
3. *Lienhart A*, *Auroy Y*, *Pequignot F*, *et al*. Survey of anaesthesia-related mortality in France. Anesthesiology 2006; 105: 1087-97.
4. *Olsson CL*, *Hallen B*, *Hambraeus-Jonzon K*. Aspiration during anaesthesia: a computer-aided study of 185,358 anaesthetics. Acta Anaesthesiol Scand 1986; 30: 84-92.
5. *Raidoo DM*, *Roche DA*, *Brock-Utne JG*, *Marszalek A*, *Engelbrecht HE*. Critical volume for pulmonary acid aspiration: reappraisal in a primate model. Br J Anaesth 1990; 65: 248-50.
6. *James CF*, *Modell JH*, *Gibbs CP*, *Kuck EJ*, *Ruiz BC*. Pulmonary aspiration – effects of volume and pH in the rat. Anesth Analg 1984; 63: 665-8.
7. *Engelhardt T*, *Webster NR*. Pulmonary aspiration of gastric contents in anaesthesia. Br J Anaesth 1999; 83: 453-60.
8. *Landreau B*, *Odin I*, *Nathan N*. Pulmonary aspiration: epidemiology and risk factors (French). Ann Fr Anesth Reanim 2009; 28: 206-10.
9. *American Society of Anesthesiologists Committee*. Practice guidelines for preoperative fasting and the use of pharmacologic agents to reduce the risk of pulmonary aspiration: application to healthy patients undergoing elective procedures: an updated report by the American Society of Anesthesiologists Committee on Standards and Practice Parameters. Anesthesiology 2011; 114: 495-511.
10. *Fujigaki T*, *Fukusaki M*, *Nakamura H*, *Shibata O*, *Sumikawa K*. Quantitative evaluation of gastric contents using ultrasound. J Clin Anesth 1993; 5: 451-5.
11. *Perlas A*, *Chan VW*, *Lupu CM*, *Mitsakakis N*, *Hanbidge A*. Ultrasound assessment of gastric content and volume. Anesthesiology 2009; 111: 82-9.
12. *Perlas A*, *Davis L*, *Khan M*, *Mitsakakis N*, *Chan VW*. Gastric sonography in the fasted surgical patient: a prospective descriptive study. Anesth Analg 2011; 113: 93-7.
13. *Bouvet L*, *Miquel A*, *Chassard D*, *Boselli E*, *Allaouchiche B*, *Benhamou D*. Could a single standardized ultrasonographic measurement of antral area be of interest for assessing gastric contents? A preliminary report. Eur J Anaesthesiol 2009; 26: 1015-9.
14. *Bouvet L*, *Mazoit JX*, *Chassard D*, *Allaouchiche B*, *Boselli E*, *Benhamou D*. Clinical assessment of the ultrasonographic measurement of antral area for estimating preoperative gastric content and volume. Anesthesiology 2011; 114: 1086-92.
15. *Wong CA*, *Loffredi M*, *Ganchiff JN*, *Zhao J*, *Wang Z*, *Avram MJ*. Gastric emptying of water in term pregnancy. Anesthesiology 2002; 96: 1395-400.
16. *Wong CA*, *McCarthy RJ*, *Fitzgerald PC*, *Raikoff K*, *Avram MJ*. Gastric emptying of water in obese pregnant women at term. Anesth Analg 2007; 105: 751-5.
17. *Kubli M*, *Scrutton MJ*, *Seed PT*, *O'Sullivan G*. An evaluation of isotonic "sport drinks" during labor. Anesth Analg 2002; 94: 404-8.
18. *Vandenbroucke JP*, *von Elm E*, *Altman DG*, *et al*. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology 2007; 18: 805-35.
19. *Cubillos J*, *Tse C*, *Chan VW*, *Perlas A*. Bedside ultrasound assessment of gastric content: an observational study. Can J Anesth 2012; 59: 416-23.

20. *Bould MD*, *Crabtree NA*, *Naik VN*. Assessment of procedural skills in anaesthesia. Br J Anaesth 2009; 103: 472-83.
21. *Barrington MJ*, *Wong DM*, *Slater B*, *Ivanusic JJ*, *Ovens M*. Ultrasound-guided regional anesthesia: how much practice do novices require before achieving competency in ultrasound needle visualization using a cadaver model. Reg Anesth Pain Med 2012; 37: 334-9.
22. *Naik VN*, *Devito I*, *Halpern SH*. Cusum analysis is a useful tool to assess resident proficiency at insertion of labour epidurals. Can J Anesth 2003; 50: 694-8.
23. *De Oliveira Filho GR*, *Helayel PE*, *Da Conceicao DB*, *Garzel IS*, *Pavei P*, *Ceccon MS*. Learning curves and mathematical models for interventional ultrasound basic skills. Anesth Analg 2008; 106: 568-73.
24. *Margarido CB*, *Arzola C*, *Balki M*, *Carvalho JC*. Anesthesiologists' learning curves for ultrasound assessment of the lumbar spine. Can J Anesth 2010; 57: 120-6.
25. *Kestin IG*. A statistical approach to measuring the competence of anaesthetic trainees at practical procedures. Br J Anaesth 1995; 75: 805-9.
26. *Halpern SH*, *Banerjee A*, *Stocche R*, *Glanc P*. The use of ultrasound for lumbar spinous process identification: A pilot study. Can J Anesth 2010; 57: 817-22.
27. *Bush RR*, *Mosteller F*. A mathematical model for simple learning. Psychol Rev 1951; 58: 313-23.
28. *Norris A*, *McCahon R*. Cumulative sum (CUSUM) assessment and medical education: a square peg in a round hole. Anaesthesia 2011; 66: 250-4.
29. *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.
30. *Sivaprakasam J*, *Purva M*. CUSUM analysis to assess competence: what failure rate is acceptable? Clin Teach 2010; 7: 257-61.
31. *Komatsu R*, *Kasuya Y*, *Yogo H*, *et al*. Learning curves for bag-and-mask ventilation and orotracheal intubation: an application of the cumulative sum method. Anesthesiology 2010; 112: 1525-31.
32. *Niazi AU*, *Haldipur N*, *Prasad AG*, *Chan VW*. Ultrasound-guided regional anesthesia performance in the early learning period: effect of simulation training. Reg Anesth Pain Med 2012; 37: 51-4.
33. *Altman DG*, *Royston JP*. The hidden effect of time. Stat Med 1988; 7: 629-37.
34. *Boulet JR*, *Murray D*. Review article: Assessment in anesthesiology education. Can J Anesth 2012; 59: 182-92.
35. *Birnbach DJ*, *Santos AC*, *Bourlier RA*, *et al*. The effectiveness of video technology as an adjunct to teach and evaluate epidural anaesthesia performance skills. Anesthesiology 2002; 96: 5-9.
36. *Gaba DM*. Two examples of how to evaluate the impact of new approaches to teaching. Anesthesiology 2002; 96: 1-2.
37. *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.
38. *Young A*, *Miller JP*, *Azarow K*. Establishing learning curves for surgical residents using cumulative summation (CUSUM) analysis. Curr Surg 2005; 62: 330-4.
39. *Johnson DW*, *Oren-Grinberg A*. Perioperative point-of-care ultrasonography: the past and the future are in anesthesiologists' hands. Anesthesiology 2011; 115: 460-2.
40. *Naslund E*, *Bogefors J*, *Gryback P*, *Jacobsson H*, *Hellstrom PM*. Gastric emptying: comparison of scintigraphic, polyethylene glycol dilution, and paracetamol tracer techniques. Scand J Gastroenterol 2000; 35: 375-9.
41. *Merchant R*, *Chartrand D*, *Dain S*, *et al*. Guidelines to the practice of anesthesia revised edition 2013. Can J Anesth 2013; 60: 60-84.