Drug Calculation Errors in Anesthesiology Residents and Faculty: An Analysis of Contributing Factors

Shira Black, DO,* Jerrold Lerman, MD, FRCP, FANZCA,**†‡ Shawn E. Banks, MD,§ Dena Noghrehkar, MD,‡ Luciana Curia, MD,* Christine L. Mai, MD, MS-HPEd,∥ Deborah Schwengel, MD, MEd,¶ Corey K. Nelson, MD,# James M. T. Foster, MD,** Stephen Breneman, MD, PhD,* and Kris L. Arheart, PhD††

BACKGROUND: Limited data exist regarding computational drug error rates in anesthesia residents and faculty. We investigated the frequency and magnitude of computational errors in a sample of anesthesia residents and faculty.

METHODS: With institutional review board approval from 7 academic institutions in the United States, a 15-question computational test was distributed during rounds. Error rates and the magnitude of the errors were analyzed according to resident versus faculty, years of practice (or residency training), duration of sleep, type of question, and institution.

RESULTS: A total of 371 completed the test: 209 residents and 162 faculty. Both groups committed 2 errors (median value) per test, for a mean error rate of 17.0%. Twenty percent of residents and 25% of faculty scored 100% correct answers. The error rate for postgraduate year 2 residents was less than for postgraduate year 1 (P = .012). The error rate for faculty increased with years of experience, with a weak correlation (R = 0.22; P = .007). The rate errors were independent of the number of hours of sleep. The error rate for percentage-type questions was greater than for rate, dose, and ratio questions (P = .001). The error rates varied with the number of operations needed to calculate the answer (P < .001). The frequency of large errors (100-fold greater or less than the correct answer) by residents was twice that of faculty. Error rates varied among institutions ranged from 12% to 22% (P = .021).

CONCLUSIONS: Anesthesiology residents and faculty erred frequently on a computational test, with junior residents and faculty with more experience committing errors more frequently. Residents committed more serious errors twice as frequently as faculty. (Anesth Analg 2019;128:1292–9)

KEY POINTS

• Question: What is the frequency and magnitude of drug calculation errors by residents and faculty in US anesthesiology programs?
• Findings: Using the results of a 15-question drug calculation test, we determined that anesthesiology residents and faculty frequently commit computational drug dosing errors; and inexperienced residents and aging faculty commit more errors, although residents commit errors of more extreme magnitude.
• Meaning: We posit that the frequency of computational drug errors, particularly of a large (or small) order of magnitude by residents in this study, warrants remedial training for residents in drug dilutions and dosing, together with clinical safeguards to preclude harm from these errors.

Early critical event reporting systems identified medication errors as one of the most common causes of perioperative morbidity and mortality.1,2 Preventable medication errors that resulted in harm to patients ranged from 0.3%–4.7% of hospital admissions in Australia to 2%–15% in the United Kingdom.3,4 Direct observation indicated that the frequency of medication errors in hospital wards was 19%, of which 17% were dosing-related errors.5 In 2003, the American Society of Anesthesiologists (ASA) closed claims study confirmed that drug administration

The authors declare no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal’s website (www.anesthesia-analgesia.org).

Listen to this Article of the Month podcast and more from OpenAnesthesia.org by visiting http://journals.lww.com/anesthesia-analgesia/pages/default.aspx.

This study was presented in part at the American Society of Anesthesiologists meeting, San Francisco, CA, October 2013. Reprints will not be available from the authors.

Address correspondence to Jerrold Lerman, MD, FRCP, FANZCA, Department of Anesthesiology, Great Lakes Anesthesiology, 1001 Main St, Suite K-3502, Buffalo, NY 14203. Address e-mail to jerrold.berman@gmail.com.

Copyright © 2019 International Anesthesia Research Society
errors were also a major source of iatrogenic harm to patients during anesthesia; \(^{38}\) of drug errors have been attributed to physician error.\(^ {7}\) International data supported these concerns because 89% of anesthesiologists in New Zealand reported committing \(\geq 1\) drug administration error, and 12.5% believed that their errors had actually harmed patients.\(^ {8}\) A subsequent prospective study of 7794 anesthesiologists in New Zealand reported an overall incidence of drug errors of 0.75% (or 1 drug error for every 133 anesthetics), based on self-reporting by anesthesiologists.\(^ {9}\) In 2009, evidence indicated that in the United States, \(\geq 1\) death occurred daily, and 1.3 million patients were injured annually, all attributable to drug errors.\(^ {3}\)

The level of provider experience is an important determinant of the rate of drug errors.\(^ {10}\) Junior residents and physicians committed more drug calculation errors than more experienced physicians.\(^ {11-14}\) The magnitude of the drug errors in one study increased with fewer hours of sleep during the preceding night and with more years of professional experience.\(^ {15}\) Although several strategies, including electronic medical records, decision-supported software, unit dosing by pharmacy, 2-person drug dosing checks, workshops, and programmable smart pumps, have reduced drug errors,\(^ {16-23}\) most anesthesiologists continue to calculate drug doses and administer drugs independently, in many instances without checks and balances (eg, in operating theaters, pediatrics, intensive care, obstetrics, emergency medicine, trauma, and cardiac arrest).\(^ {11,10,24}\) The consequences of which may contribute to serious morbidity and mortality.\(^ {25}\)

The severity of dosing errors in children is often greater than in adults because drug doses in children are usually administered on a weight basis, and the range of weights in children varies by 300-fold, whereas that in adults varies only by 5- to 6-fold.\(^ {8,11,26,27}\) For some medications, the calculations are challenging, and they introduce the risk of errors from a number of different sources,\(^ {25}\) the most common being computation, units, dilution, weight, and transcription. Whether considered individually or in aggregate, these errors may result in 10-fold or greater overdoses in children, which could lead to serious and potentially fatal sequelae.\(^ {11,12,26}\)

The purpose of this study was to investigate the frequency of computational drug errors in a large national sample of anesthesia residents and faculty via a standardized computational test, to identify potential contributing factors, and to determine the magnitude of the dosing errors.

**METHODS**

The directors of 11 anesthesia residency programs were invited along with their anesthesiology residents and faculty to participate in this prospective study to test the computational skills in their departments at grand rounds by the senior author (J.L.). Seven directors agreed to participate.

Institutional review board approval was obtained at each of the 7 academic anesthesiology departments in the United States with a waiver of consent. This study adhered to the applicable Enhancing the Quality and Transparency of Health Research guidelines. This study was not registered.

At a single grand rounds at each site, an information sheet and the drug calculation test were distributed (without previous notification) to all anesthesiology residents and faculty in attendance. The information sheet described the purpose and nature of the study, emphasizing that completion of the test at the rounds was voluntary and anonymous. Written consent was not required by any review board; if the participant completed the test, then consent was implied. The test was proctored by the program director or his/her associate at each participating center to ensure that participants did not collaborate in their answers. Those who declined to complete the test were excused from the rounds. Demographic data were not collected from nonresponders. All sites completed the same drug calculation test.

The test was not time constrained, and participants were allocated as much time as needed to answer the 15 questions. Participants were permitted to use any instrument on their person (calculator, smartphone, and computer) to calculate their answers. Completed tests were collected, graded by 2 independent and blinded investigators, and then collated.

**Test Construction and Psychometric Properties**

The test was constructed in a 3-step manner as outlined below.

A preliminary version of the test instrument, which consisted of 25 questions, was first vetted by 3 academic anesthesiologists (Dr R. Cox, University of Alberta, Edmonton, AB, Canada; Dr T. Erb, University Hospital, Basel, Switzerland; and Dr T. Yemen, University of Virginia, Charlottesville, VA) for accuracy, relevance, and range of difficulty. After editing, a second version of the questions was trialed in a pilot study with 20 anesthesiologists at an international anesthesia meeting, after which the questions were further edited to yield the final version of the test. The final test consisted of 15 questions to address the spectrum of clinical relevance and numerical computations. The test questions were worded to require answers in different units (eg, as a dose, dose per body weight or an infusion rate over time) to reflect the types of units used when administering drugs in anesthesia. To ensure a consistent answer format, most questions stipulated the number of decimals places for the answer (Supplemental Digital Content, Appendix 1, http://links.lww.com/AA/C697). Specifically, the questions tested the examinees’ ability to calculate drug dilutions, drug concentrations, and infusion rates, and thus required knowledge of fundamental mathematical principles that are applicable to adult and pediatric age drug doses. Several drugs that are not commonly used in anesthesia were included to test the participants’ computational skills without relying on any prior knowledge of and familiarity with the usual dosages for these drugs. The purpose of the study was to evaluate the medication-related computational skills of resident and faculty anesthesiologists.

A confirmatory factor analysis was used to determine whether the 15 questions represented a single factor. Cronbach alpha was used to determine the internal consistency and reliability of the summary question scale. The confirmatory factor analysis suggested that all 15 items loaded on the same latent scale (“knowledge”) with standardized factor loadings between 0.32 and 0.71 (>0.30 is good), a root mean square error of approximation fit statistic of 0.033 (<0.06 is good), and a comparative fit index fit statistic of 0.93 (>0.95 is good),\(^ {29}\) supporting the interpretation of a total score. All items correlated positively with the total score and with each other, resulting in an acceptable level of internal consistency.
Cronbach $\alpha = .71$. Confirmatory factor analysis was performed using MPlus (Muthen & Muthen, Los Angeles, CA).

**Statistical Methods and Analysis**

In the absence of data on which to base a sample size calculation, a convenience sample of 2%-5% of the total number of anesthesia residents in the United States was expected from the 7 participating institutions. We posited that this sample would be representative of the national population of US anesthesiology residents and faculty and provide a reliable geographic representation to reliably estimate the computational error rates.

Demographic data of the participants, including training status (resident or faculty), training (in resident year), or years of experience (post completion of residency), and the number of hours of sleep during the preceding night were collected. The numbers of residents and faculty with correct answers (%) correct answers) and the number of errors per participant were determined, the latter compared among institutions. Unanswered questions were considered incorrect answers.

Factors that significantly affected the error rates including the institution, the participant’s position (resident or faculty), and the resident’s year of training were compared using generalized linear models. Model mean percents and SEs are presented along with the overall $P$ value for each model. When the overall $P$ value was significant, pairwise comparison $P$ values from a Fisher-protected least significance test were calculated.

A generalized linear mixed model for repeated measures was used to determine the effect of type of drug calculation (percent, rate, dose, and ratio) and the number of calculations involved on the error rate (categories of question type and number of calculations were repeated within participants). When the overall $P$ value was significant, pairwise comparison $P$ values from a Fisher-protected least significance test were calculated.

A generalized linear model was used to perform an analysis of a Poisson distributed variable to compare the number of errors per person between residents and faculty. Means and SEs are presented in the Results text. Also, a generalized linear model for a normally distributed variable was used to perform a repeated-measures analysis to compare the average magnitude of the errors per participant between residents and faculty; results shown as means and SEs. Pearson correlations were used to assess the strength of the relationship of years of experience and hours of sleep with the number of errors and mean magnitude per participant.

Post hoc, the magnitude of the calculation errors was analyzed. If an answer was within 5% of the correct answer, it was defined to be a “correct” answer. Incorrect answers were those that were outside of 5% greater or less than the correct answer. This permitted an analysis of the magnitude of incorrect answers. The magnitude of incorrect answers was grouped empirically into 5 ranges\(^{13,14}\): $\geq$2-fold greater or less than the correct answer; $\geq$2-fold to $<$10-fold greater or less; $\geq$10-fold to $<$100-fold greater or less; $\geq$100-fold to $<$1000-fold greater or less; and $\geq$1000-fold greater or less than the correct answer.

Statistical analysis was undertaken by a statistician who was blind to the institutions. A 2-tailed $P < .05$ was accepted. Data are displayed as means ± SEM unless indicated otherwise. Statistical analyses other than the confirmatory factor analysis were performed using SAS 9.3 (SAS Institute Inc, Cary, NC).

**RESULTS**

Seven institutions participated in this study: Massachusetts General Hospital, Boston, MA; State University of New York at Buffalo, Buffalo, NY; Johns University Hopkins, Baltimore, MD; University of Miami, Miami, FL; University of Rochester, Rochester, NY; University of New York Upstate, Syracuse, NY; and University of California at Irvine, Irvine, CA. A total of 371 anesthesia residents in the first through fourth postgraduate years, and anesthesia faculty consented to participate. The number of residents and faculty among the 7 institutions were evenly distributed among the institutions (Table 1). However, only 5 first-year residents completed the test: 1 from the 1 institution, and

| Hospital Designation | Faculty Number | Experience (y) | Resident Numbers | Postgraduate Year 2 | Postgraduate Year 3 | Postgraduate Year 4 | Overall Error Rates (%) |
|----------------------|----------------|----------------|------------------|---------------------|---------------------|-----------------------|
| A                    | 27             | 9.6 ± 1.1      | 33               | 14                  | 9                   | 10                    | 12.2 ± 2.1             |
| B                    | 12             | 7.1 ± 1.4      | 30               | 10                  | 10                  | 10                    | 21.0$^b$ ± 2.5         |
| C                    | 17             | 7.3 ± 1.7      | 12               | 7                   | 5                   | 5                     | 16.3 ± 3.0             |
| D                    | 41             | 6.4 ± 1.4      | 49               | 22                  | 16                  | 11                    | 14.4 ± 1.7             |
| E                    | 32             | 8.4 ± 1.1      | 42               | 11                  | 16                  | 14                    | 19.6$^c$ ± 1.9         |
| F                    | 20             | 6.4 ± 1.3      | 27               | 9                   | 6                   | 8                     | 16.9 ± 2.3             |
| G                    | 13             | 8.4 ± 1.7      | 16               | 8                   | 5                   | 3                     | 22.1$^d$ ± 3.0         |
| H                    | 17             | 7.3 ± 1.7      | 12               | 7                   | 5                   | 5                     | 16.3 ± 3.0             |
| Totals               | 162            | 209            | 81               | 62                  | 61                  | 371                   | 17.0 ± 0.8             |

The data from 5 interns (postgraduate year-1) who completed the test were omitted from the table (see text).

Experience years and error rates are presented as mean ± SE.

Generalized linear model analyses of the error rates were performed. Model means and SEs are reported, and $P$ values for Fisher-protected least significant difference tests are reported when the main effect is significant.

$^aP = .007$ compared with hospital A.

$^bP = .029$ compared with hospital D.

$^cP = .008$ compared with hospital A.

$^dP = .038$ compared with hospital D.

$^eP = .026$ compared with hospital D.

$^fP = .012$ compared with postgraduate year 1.
4 from another. Given the small number of first-year residents who participated, their responses were not included in the analysis. A total of 209 residents (postgraduate year 2 to postgraduate year 4) and 162 faculty completed the test and were included in the analysis; 1.2% of the questions on all of the tests combined were unanswered.

Only 20% of residents and 25% of faculty scored 100% on the test. A detailed analysis of the error rates by residents and faculty according to their institution, level of training, and years of experience is presented in Table 1. The mean error rate (%) was calculated as the number of incorrect answers \( \times 15 \times 1 \times 100 \). For the entire sample, the mean correct answer rate was 83% (95% CI, 81.4–84.8), yielding a mean error rate of 17% (95% CI, 15.2–18.6). Figure 1A displays percentage of residents and faculty who answered each of the questions correctly. The error rates for the residents (17.5% [1.1]) and faculty (16.3% [1.2]) were similar (P = .51) (Table 1). The total number of incorrect answers by residents (2.9 ± 0.1 errors/resident) and faculty (3.0 ± 0.2 errors/faculty) did not differ significantly (P = .504) (Figure 1B). However, the error rates among the residents differed according to their level of training (P = .031); the error rate for postgraduate year 1 residents was significantly greater than that for postgraduate year 2 (P = .012), but not for postgraduate year 3 (P = .076).

The error rates for postgraduate year 2 and postgraduate year 3 residents did not differ (P = .48) (Table 1). The overall error rates differed significantly among the 7 institutions (P = .021) (Table 1): the error rate for institution A was significantly less than that for institutions B (P = .007), E (P = .008), and G (P = .007). Similarly, the rate for institution D was significantly less than that for institutions B (P = .029), E (P = .038), and G (P = .026).

The error rates differed significantly among the types of questions (P = .001) (Table 2). Percentage questions yielded the greatest error rates compared with rate (P < .001), dose (P = .001), and ratio (P = .015) questions. Ratio questions yielded significantly greater error rates than for rate questions (P < .001), and dose questions yielded significantly greater error rates than for rate questions (P = .005). The error rates caused by the number of operations to obtain an answer also differed significantly overall (P < .001) (Table 2). The error rates from greatest to least were for 5 > 4 > 2 > 3 > 1 mathematical operations (pairwise at P < .001, except for 1 to 3 operations [P = .796] and 4 to 5 [P = .70]). Six questions included drugs with which most anesthesiologists are unfamiliar: Nos. 3, 4, 7, 8, 13, and 15; the remaining 9 questions included familiar drugs. The list of unfamiliar drugs included drugs to treat pulmonary hypertension and heart failure, as well as for epidural infusions in a child and dantrolene for malignant hyperthermia. The average number of errors for these 6 questions (from all participants combined) (82 ± 15) was similar to that for the remaining 9 (54 ± 10) (P = .12). The questions associated with the greatest frequency of errors were No. 13 (n = 116), 4 (n = 112), and 6 (n = 106).

Total years of experience for residents was 1.9 ± 0.1 years, and for faculty, it was 15.4 ± 0.7; when combined, the years of experience was 7.6 ± 0.5 years. Pearson correlation between experience and error rate for faculty was modest but significant (R = 0.022; P = .007), whereas that for residents was not significant (R = −.011; P = .12). When combined, the years of experience did not correlate with the error rate (R = 0.04; P = .39). The (mean ± SE) number of hours of sleep for residents was 6.2 ± 0.1, and for faculty, it was 6.3 ± 0.1; when combined, it was 6.3 ± 0.1 hours. There were no significant

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Error Rate (Mean % ± SEM)</th>
<th>No. Operations</th>
<th>Error Rate (Mean % ± SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose</td>
<td>19.2 ± 1.1*a</td>
<td>1</td>
<td>7.5 ± 1.4</td>
</tr>
<tr>
<td>Percentage</td>
<td>26.4 ± 2.3</td>
<td>2</td>
<td>15.9 ± 0.9*</td>
</tr>
<tr>
<td>Rate</td>
<td>12.9 ± 0.9*</td>
<td>3</td>
<td>8.1 ± 1.4*</td>
</tr>
<tr>
<td>Ratio</td>
<td>21.0 ± 2.1</td>
<td>4</td>
<td>25.9 ± 1.8*</td>
</tr>
</tbody>
</table>

*P < .001 compared with percentage.

Table 2. Error Rates for the Type of Question and Number of Operations

A repeated-measures analysis was performed because participants are measured several times for question type and number of operations. Model means and SEs are reported. P values for Fisher-protected least significant difference tests are reported when the main effect is significant.

4P < .001 compared with 1 operation to answer the question.

5P < .001 compared with percentage.

6P < .005 compared with ratio questions.

7P < .001 compared with 2 operations to answer the question.

8P < .001 compared with 3 operations to answer the question.
Drug Calculation Errors in Anesthesia

relationships between the error rate and the duration of sleep the previous night (all R < 0.06; P > .42).

Post hoc, the authors posited that drug calculation errors that were up to twice the correct dose might result in minor adverse events but were less likely to result in serious adverse events than those that were >2-fold greater or less than the correct dose based on the therapeutic index for most anesthetic drugs in humans. As a result, we compiled the frequency of errors that was >2-fold greater or less than the correct dose (Figure 2; Supplemental Digital Content, Appendix 1, http://links.lww.com/AA/C697). The magnitude of the errors in aggregate for the residents (2.6 ± 0.1) and faculty (2.5 ± 0.1) was similar (P = .42) and independent of the years of experience (R = 0.07, P = .24) and the number of hours of sleep the night before (R = −0.6; P = .29). The answers to 2 questions, No. 4 (n = 88) and No. 13 (n = 91), comprised 26% of the errors of large magnitude. Overall, 57% of the errors of large magnitude exceeded the correct answer and 43% were less.

We also determined that 6% of the participants committed error rates between 10- and 100-fold greater (or less) than the correct answer. For errors (total = 99 errors) that exceeded 100-fold greater (or less) than the correct answer, residents (n = 51) committed 68% of the errors, twice that committed by the faculty (n = 29). Residents erred with >1 incorrect answer of large magnitude (>100-fold greater or less than the correct answer), with a frequency of 13 out of 51 (or 25%) compared with faculty, at 3 out of 29 (or 10%) (P = .10).

**DISCUSSION**

We determined the frequency and magnitude of computational errors in 371 residents and faculty from 7 anesthesiology residency programs in the United States by analyzing their responses to a written test of drug doses and infusions. We undertook this investigation because drug errors occur frequently in anesthesiology (once in every 133–274 anesthetics), are more common in emergency situations, and may lead to serious sequelae.27,30 Multiple strategies have been proposed to mitigate drug errors, although they have not been uniformly adopted in anesthesia practice20,21 and in some cases such as smart pumps, have been overridden by clinicians.23 And yet, one strategy, the standardization, technology, pharmacy, and culture arising from a consensus conference of the Anesthesia Patient Safety Foundation, has been successfully implemented to reduce drug errors.31 In this single, snapshot test of computational skills in a large cohort of residents and faculty in which the time to complete the study was not constrained and technology devices (eg, phone and calculator) were permitted, a minority of residents (20%) and faculty (25%) answered all of the questions correctly. Residents and faculty both erred similarly in 17% of questions, although errors of a large magnitude in residents, 100-fold greater or less than the correct answer, were more frequent than in faculty. The results of this study identify an underappreciated deficiency in the computational skills of both anesthesia residents and faculty.

Computational proficiency is not a prerequisite for medical school, although some entrance examinations in Europe now include a numerical reasoning component.32 When the computational skills of 168 medical students were tested with 3 drug dosing questions, only 10% answered all 3 questions correctly, and 25% answered all 3 incorrectly.33 Several studies reported that the error rate by junior postgraduates was greater than by senior year residents,12,24 which is consistent with our results (Table 1). Furthermore, a small proportion of pediatric residents (10%–30%) committed 10-fold errors, and 5% committed 1000-fold errors, also consistent with our results (Figure 2).11,12,26 What distinguishes this study from previous studies is that we tested computational proficiency in a large cohort of residents and faculty (10-fold greater than in a previous study),26 in a multi-institutional, pan-national design,11,12,33,35 and with a test that included 6 questions focused on drugs with which participants would be less familiar (Supplemental Digital Content, Appendix 1, http://links.lww.com/AA/C697).11,12,25,26 These latter questions minimized the impact of familiarity on the answers (also known as crystallized intelligence) compared with the remaining questions.11,12,33,35 Notably, neither the error rates nor the frequency of large (or small) errors for these 6 questions with unfamiliar drugs differed significantly from the remaining 9 questions with more familiar drugs. We believe that this approach increased the external validity of our results and allows us to recommend that educational strategies should be adopted.

Developing a strategy to improve the computational skills of anesthesia residents requires a multipronged approach. The first strategy would establish a baseline by testing the computational skills of each resident at admission into the program,12,19,32,33 The second would require

![Figure 2. The magnitude of the incorrect answers for each of the 15 questions. Errors that were <2-fold greater or less than the correct answer were omitted from the graph (see the Results section for explanation). Each histogram displays the percentage of the incorrect answers that were >2-fold greater or less than the correct answer, grouped into 4 error size categories as shown in the legend. The size categories were based, in part, on the work of Glover and Sussman.12](http://links.lww.com/AA/C697)
that each resident (with emphasis on those who performed poorly on the admission test) participate in serial workshops to address knowledge gaps and improve their computational skills. The third would involve follow-up assessments periodically throughout the residency program to ensure maintenance of computational proficiency. At the end of the training program, a final test would either confirm whether an acceptable level of computational skills was maintained. If not, then remedial computational training would be provided. We posit that this process for improving computational skills when combined with systemwide strategies that included standardization, technology, pharmacy, and culture, electronic medical record dosing alerts, and/or electronic apps will attenuate the frequency and magnitude of drug errors by anesthesia residents.

Addressing the error rates in faculty requires an approach that is distinct from that for the residents. In this study, even experienced anesthesiologists committed computational errors at a rate ≥15%, with 10% committing >1 error of a large order of magnitude. Available strategies to attenuate faculty computational errors include adopting smart pumps, relying on pharmacy to prepare correct doses, including electronic medical record alerts and the standardization, technology, pharmacy, and culture strategy. The results of the few studies of the effectiveness of smart pumps to prevent dosing errors are mixed. In some, the reduction in errors was limited because clinicians overrode the soft limits on the pumps. Cost-cutting strategies in pharmacies have increasingly limited their ability to prepare drug dilutions, infusions, and unit dosing. Nonetheless, preparing drug infusions in a central pharmacy or from commercially prepared stock solutions reduces drug concentration errors compared with those prepared by nurses on the ward.

Furthermore, automated preparation techniques in the pharmacy reduce errors compared with manually prepared solutions and when ward nurses prepared drug infusions using prefilled drug syringes, resulted in fewer drug errors compared with infusions prepared from standard drug vials. A novel notion is to maintain computational skills testing in maintenance of competence in anesthesia educational programs as well to institute remedial training for those who fail or commit frequent errors or errors of large magnitude. In aggregate, these strategies may reduce drug dosing errors and the magnitude of the errors by faculty anesthesiologists, although studies are required to establish their effectiveness.

The frequency of incorrect answers by faculty increased with years of experience, although experience accounted for only 5% of the variability in the error rate. We used the respondents’ “years of experience” as a surrogate index for aging in the after-analysis. Aging is associated with a gradual deterioration in memory and executive cognitive functions; however, its effect on computational skills has not been fully elucidated. The test used in this study assessed 2 different skills: computation or arithmetic skills, and problem solving. The former skills are well preserved with age and may actually improve as a result of the years of experience during which these skills become firmly established. However, the latter, problem solving, may deteriorate. Problem solving depends on 2 executive functions: the inhibitory process, and fluid intelligence. The inhibitory process, which becomes progressively impaired with age (beyond 50 or 60 years of age), is the ability to separate extraneous information from that required to solve the problem. The second, fluid intelligence, which may begin to deteriorate as early as 50 years of age, allows innovative thinking to solve unfamiliar problems, without relying on previous experience or knowledge. In our test, 40% of the questions involved drugs with which the participants were unfamiliar, thus requiring a greater element of problem solving. Thus, the small decrease in the number of correct answers with years of experience may reflect the offsetting effects of preserved computational skills and waning inhibitory and fluid intelligence processes.

Computational proficiency has been investigated to only a limited extent in anesthesia. In a study of 141 physicians and surgeons, drug calculation errors occurred in the majority of participants who completed the 12-question test, although anesthesiologists committed fewer errors than both physicians and surgeons. The source of the errors was multifactorial: misreading the question, using incorrect operations (eg, dividing instead of multiplying), and/or procedures (eg, incorrectly placing the decimal point or reporting the answer in incorrect units). Although many errors may go unnoticed and have minimal impact on patient outcomes, others may lead to serious sequelae. Anesthesiologists administer a small number of drugs on a daily basis, possibly limiting the frequency of errors they commit, although the frequency and magnitude of errors with familiar and unfamiliar drugs in this study were similar. Nonetheless, several approaches have demonstrated that computational skills may be improved long term (up to 4 years after the initial workshop training), including drug calculation lectures and participation in workshops and simulation in both medical and pharmacy schools. Further studies are needed to identify the source(s) of and to develop effective strategies to reduce the most frequent and largest drug errors in anesthesia.

Drug concentrations are often labeled differently on the packaging, resulting in errors when the provider has to convert the units to a measurable unit. Studies demonstrated that dosing errors were more common when concentrations were expressed as ratios or percentages. Our results are consistent with those findings (Table 2).

Computational errors of the order of magnitude reported here are exceedingly dangerous in anesthesia. Ten-fold or greater errors have been reported in several studies, with 1 student (5%) reportedly committing a 1000-fold calculation error. In this study, 6% of the participants committed drug errors 10- to 100-fold greater (or less) than the correct answer, and 6.7% of the residents committed >1 error with a magnitude in excess of 100-fold greater or less than the correct answer, which are consistent with previous studies. The clinical impact of errors depends on the therapeutic index of the drugs involved in the dosing errors. Dosing errors that are 100-fold greater (or less) than the correct dose almost certainly will result in substantial consequences that may range from prolonged recovery or intensive care admission to the need for cardiopulmonary resuscitation; drugs with narrow therapeutic indices are twice as likely to yield
Drug Calculation Errors in Anesthesia

There are several limitations in this study. First, we studied only a sample of residents compared with the total number in residency programs in anesthesiology in the United States (209 of the 5578 residency positions or 3.7%) and a small fraction of faculty. Whether this sampling accurately reflects the computational proficiency of the entire population of residents and faculty remains unclear. Moreover, all participants work in the United States, possibly limiting the external validity of our data. Second, we did not randomly select the anesthesia programs from across the country to participate in this study, but rather used a convenience sample of those institutions that had program directors who were interested in this study. Thus, sampling bias of the residency programs may have skewed our results. Because individuals could opt out of participating in the test, there is a risk that those who chose not to participate could have included a disproportionately large number of residents and faculty with weaker math skills, known as response bias. Third, although we vetted the questions through experts and a sample of clinical before finalizing the syntax and number of questions, we standardized neither the number of questions with similar concentration units (eg, percentages) nor the number of computational operations per question that may have resulted in a disproportionate number of questions of a specific type. Finally, by conducting the study in a quiet “laboratory” setting (outside of the operating room), the error rates reported here may actually underestimate the rates that would have been obtained had the tests been conducted in a stressful operating room environment with multiple sensory distractions.43,56

In conclusion, we determined that both residents and faculty in anesthesiology at 7 academic institutions in the United States committed a substantial number of drug calculation errors during computational skill testing. Error rates occurred more frequently in less experienced residents and in more experienced faculty. Very large or very small dosing errors occurred infrequently, although they were, more commonly committed by residents than faculty. These findings represent a serious potential risk for harm to patients.

DISCLOSURES

Name: Shira Black, DO.
Contribution: This author helped collect data, submit to the institutional review board at University of Rochester, write the first draft, edit subsequent drafts, and approve the final copy.

Name: Jerrold Lerman, MD, FRCP, FANZCA.
Contribution: This author helped develop the test, submit to the institutional review board at University of Buffalo and University of Rochester, collect data, write and edit all drafts, prepare figures, and approve all copies.

Name: Shawn E. Banks, MD.
Contribution: This author helped submit to the institutional review board at University of Miami, collect data, edit drafts, and approve all copies.

Name: Dena Noghrehkar, MD.
Contribution: This author helped submit to the institutional review board at University of Buffalo, collect data, edit drafts, and approve the final copy.

Name: Luciana Curia, MD.
Contribution: This author helped submit to the institutional review board at University of Rochester, collect data, edit drafts, and approve the final copy.

Name: Christine L. Mai, MD, MS-HPEd.
Contribution: This author helped submit to the institutional review board at University of Rochester, collect data, edit drafts, and approve the final copy.

Name: Deborah Schwengel, MD, MEd.
Contribution: This author helped submit to the institutional review board at Johns Hopkins University, collect data, edit drafts, and approve the final copy.

Name: Corey K. Nelson, MD.
Contribution: This author helped submit to the institutional review board at UC Irvine, collect data, edit drafts, and approve the final copy.

Name: James M. T. Foster, MD.
Contribution: This author helped submit to the institutional review board at SUNY Upstate, collect data, edit drafts, and approve the final copy.

Name: Stephen Breneman, MD, PhD.
Contribution: This author helped collect data, write the first draft, edit subsequent drafts, and approve the final copy.

Name: Kris L. Arheart, PhD.
Contribution: This author helped prepare the manuscript, edit drafts, analyze the data statistically, and approve the final copy.

This manuscript was handled by: Edward C. Nenengut, MD.

REFERENCES


30. Wright SR, Bradley PM. Has the UK clinical aptitude test improved performance and regulatory compliance as measured by clinical pertinence indicators before and after implementation of Anesthesia Information Management System (AIMS). J Med Syst. 2014;38:5.


