

Subtest Scores From the In-Training Examination: An Evaluation Tool for an Obstetric-Anesthesia Rotation

ROBERT GAISER, MD

Abstract

Objective To evaluate resident performance in the obstetric-anesthesia rotation using resident portfolios and their In-Training Examination scores, which are provided by the American Board of Anesthesiology/American Society of Anesthesiologists.

Methods We reviewed academic portfolios for second- and third-year anesthesiology residents at a single institution from 2006–2008 to examine United States Medical Licensing Exam Step 1 and 2 scores, grade for obstetrics-gynecology in medical school, and performance on the In-Training Examination. Faculty evaluation of medical knowledge and correlations for the various scores were obtained.

Results We examined scores for 43 residents. The subtest score for obstetric anesthesia increased after completing

a rotation in obstetric anesthesia, 26.1 ± 10.3 versus 36.3 ± 10.6 ($P = .02$). The subtest score correlated with United States Medical Licensing Exam Step 2, $r = 0.46$ ($P = .027$) but not with United States Medical Licensing Exam Step 1 or with the grade obtained in medical school. There was no correlation between faculty evaluations of medical knowledge and resident subtest scores in obstetric anesthesia.

Conclusions Subtest scores in obstetric anesthesia are valid and provide a tool for the assessment of the educational program of a rotation. Knowledge as assessed by a faculty member is different from the knowledge assessed on a written examination. Both methods can help provide a more complete assessment of the resident and the rotation.

Introduction

The American Board of Anesthesiology/American Society of Anesthesiologists In-Training Examination (ITE) is administered annually to assess residents' knowledge in anesthesiology.¹ Another purpose of the examination is to provide the program and the resident feedback to improve the educational process.

The ITE is a useful tool, possessing content validity and reliability. Content validity concerns whether the test generates responses that reflect the knowledge residents are expected to possess.² The ITE incorporates a number of questions from each area of anesthesiology. Before scoring the examination, the items are analyzed to ensure that they are correctly keyed and free of defects. Items are rechecked if fewer than 30% of examinees responded correctly or if it appears that there is more than 1 answer to a question.

Reliability refers to the reproducibility of the scores. Given the effort required to generate and administer the ITE, it is not possible to perform test-retest reliability or parallel forms reliability. For the ITE, internal consistency is checked. This method involves analyzing the test to determine if 2 halves containing equal numbers of items correlate.³ The overall reliability represents an average of the reliability coefficients obtained from all split halves (coefficient α). For the ITE, there is a strong coefficient α (>0.8).⁴ In addition to its reliability and content validity, the ITE subtest score in obstetric anesthesia also has construct validity. Construct validity refers to the extent to which the test measures a particular construct. Further construct validity is reflected by the correlation with United States Medical Licensing Exam (USMLE) Step 2 but not with USMLE Step 1.⁵ USMLE Step 1 reflects the knowledge gained during the first 2 years of medical school, whereas USMLE Step 2 reflects the core clinical clerkships. The USMLE Step 2 is viewed as a clinical test and the ITE subtest score is a similar measure of clinical knowledge.⁶ The subtest score may reflect knowledge that is clinically relevant.

Domain knowledge in anesthesiology is subdivided into 14 distinct areas: anatomy, anesthesiology processes, cardiovascular, hematology, neurologic, obstetrics, pain medicine, statistics, pharmacology, physics equipment,

Robert Gaiser, MD, is Professor of Anesthesiology and Critical Care at Hospital of the University of Pennsylvania.

Corresponding author: Robert Gaiser, MD, Hospital of the University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104, 215.662.3773, gaiserr@uphs.upenn.edu

Received October 10, 2009; revision received January 18, 2010; accepted April 22, 2010.

DOI: 10.4300/JGME-D-09-00066.1

pediatrics, physiology, regional anesthesia, and respiratory function. In 2006, the ITE provided subtest scores for each area, one of which was for obstetric anesthesia. The ITE scores are valid, and they should help to provide an assessment of the obstetric-anesthesia education program at the Hospital of the University of Pennsylvania. This study compared scores before and after the obstetric rotation and determined factors that influence these subscores of the ITE.

Methods

This is a retrospective case-control study using a convenience sample. The study was approved by the Institutional Review Board. All residents provided oral consent to have their academic files reviewed. Residents in clinical anesthesiology years II (postgraduate year 3) and III (postgraduate year 4) for the years 2006–2007 and 2007–2008 had their academic file reviewed. Residents complete the mandatory 1-month rotation in obstetric anesthesia during the clinical anesthesiology II year, and residents in clinical anesthesiology I year were excluded because they would not have completed the obstetric-anesthesia rotation prior to taking the ITE. The rotation consists of clinical cases, required readings, and lectures. The curriculum is designed to cover the physiologic changes of pregnancy and care of the high-risk parturient. Five faculty members, who completed fellowship training in obstetric anesthesia provided the didactics, supervised the 4 residents rotating that month, and completed the evaluations. Evaluations were completed at the end of the rotation and categories for evaluation were based on the Accreditation Council for Graduate Medical Education core competencies. The categories are scored on a 5-point Likert scale, ranging from 1 (requires remediation) to 5 (outstanding). Anonymous evaluations were solicited and completed by computer. The faculty members received a computer session explaining how to complete the form but did not receive training in the evaluation process.

The data extracted from the residents' files included USMLE Step 1 score, USMLE Step 2 score, ITE for clinical anesthesiology year I, and the medical school grade obtained in obstetrics-gynecology. Data extracted from the education portfolio included the faculty evaluations of medical knowledge in obstetrics regarding residents' "ability to propose a plan," knowledge of "medical literature," and "clinical judgment." These 3 scores were averaged, and the mean score was used as the faculty's rating for medical knowledge. Finally, the residents' ITE-scaled subtest score for obstetric anesthesia was recorded.

Subtest scores were compared for residents who completed the obstetric rotation versus those who did not, using Student *t* test. Linear regression was used for USMLE Step 1, USMLE Step 2, ITE score at end of the clinical anesthesiology I year, and medical school grade in obstetrics-gynecology. The Pearson correlation coefficient was determined for faculty evaluation of medical

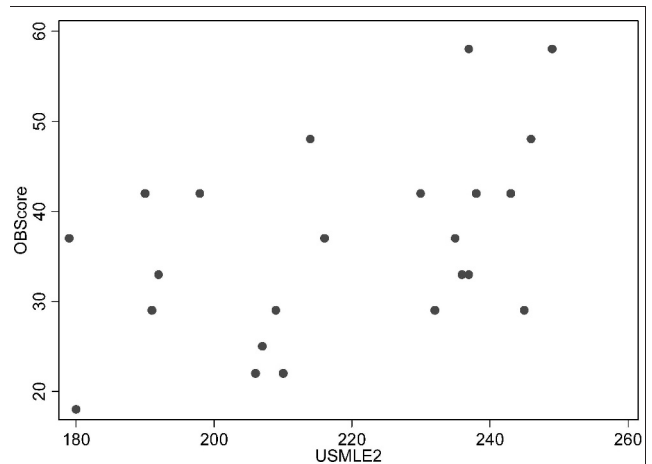


FIGURE 1

THE CORRELATION BETWEEN THE SUBTEST SCORE FOR OBSTETRIC ANESTHESIA FOLLOWING A MANDATORY 1-MONTH ROTATION IN OBSTETRIC ANESTHESIA WITH PERFORMANCE ON THE USMLE STEP 2 WAS 0.46

Abbreviations: OBSScore, score on subtest section; USMLE2, score on United States Medical Licensing Examination Step 2.

knowledge and subtest score. For data analysis, Intercooled Stata 9 (StataCorp LP, College Station, TX) was used.

Results

A total of 43 residents were included in this study. Of these residents, 21 residents were in the clinical anesthesiology II year (prerotation) and 22 residents were in the clinical anesthesiology III year (postrotation). Of the clinical anesthesiology II residents, 3 residents were included in the clinical anesthesiology III analysis as they had completed the obstetric rotation prior to the ITE (eg, during the clinical anesthesiology I year).

Residents demonstrated a statistically significant increase in knowledge after completing the obstetric-anesthesia rotation, with the mean subtest score increasing from 26.1 ± 10.3 prerotation to 36.3 ± 10.6 postrotation ($P = .02$). The correlations between subtest scores after completion of the rotation and USMLE Step 1 ($r = 0.38$) and between the medical school obstetric rotation grade ($r = -0.13$) were weak and not statistically significant. There was a moderate correlation with USMLE Step 2 ($r = 0.46$, $P = .027$) and with performance on the ITE during the clinical anesthesiology I year ($r = 0.43$, $P = .048$). FIGURE 1 plots the data for subtest scores and USMLE Step 2, and FIGURE 2 plots the data for subtest scores and ITE score for the clinical anesthesiology I year. Using linear regression, the predicted score on the obstetric section of the ITE following completion of the rotation can be predicted by the following formula:

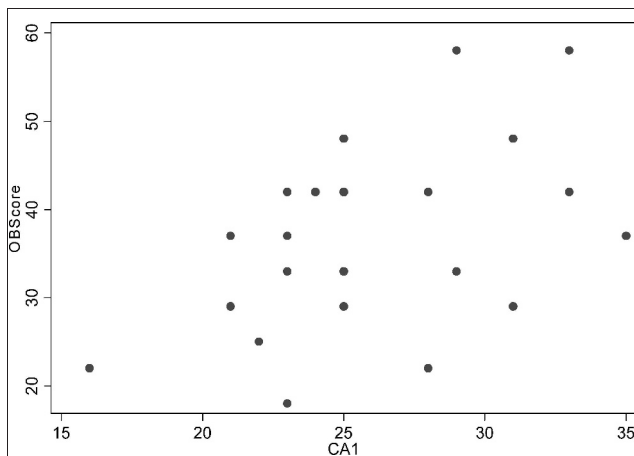


FIGURE 2

THE CORRELATION BETWEEN THE SUBTEST SCORE FOR OBSTETRIC ANESTHESIA FOLLOWING A MANDATORY 1-MONTH ROTATION IN OBSTETRIC ANESTHESIA WITH PERFORMANCE ON THE IN-TRAINING EXAMINATION DURING THE FIRST YEAR OF RESIDENCY WAS 0.48

Abbreviations: CA1, score on In-training Examination during first year of anesthesia training; OBScore, score on subtest section.

$$\begin{aligned} &\text{obstetric-anesthesia score on ITE} \\ &= 0.15(\text{USMLE Step 2 score}) \\ &+ 0.53(\text{clinical anesthesiology-I ITE score}) - 10.9 \end{aligned}$$

The mean faculty evaluation for residents was 3.8 ± 0.2 . There was no correlation between faculty evaluations and resident subtest score ($r = 0.04$).

Discussion

We used subtest scores from the ITE to assess the educational program in obstetric anesthesia (a required 1-month rotation), in keeping with the expectations of the Outcome Project to use data on resident performance to improve the educational program. The Accreditation Council for Graduate Medical Education requirements stipulate that competency in medical knowledge requires residents to “demonstrate knowledge of established and evolving biomedical, clinical, epidemiological and social-behavioral sciences, as well as the application of this knowledge to patient care.”²⁷

Before a score is used, it should be determined to be reliable and valid. The subtest score in obstetric anesthesia has both content and construct validity. Construct validity refers to the extent to which the test measures a particular construct. The ITE subtest score did not correlate with the grade obtained in the obstetrics-gynecology course in medical school. Reasons may include variation in the obstetrics-gynecology curriculum and in grading systems among medical schools. In addition, the subtest score likely

represents new knowledge learned by the resident during the obstetric-anesthesia rotation. This hypothesis is supported by the increase in scores after completion of the rotation.

It is important to note that the ITE does not have criterion validity. Criterion validity refers to the comparison of the score to faculty ratings. Residents receive numerical ratings by faculty on medical knowledge, and our study found no correlation between faculty ratings of medical knowledge and subtest scores. The lack of correlation may be due to differences in the knowledge being assessed, with faculty evaluations potentially reflecting knowledge different from that measured by the ITE. The faculty has the opportunity to assess the application of knowledge in clinical decisions, allowing an assessment of judgment and adaptability, qualities that are difficult to assess on most written examinations. Rather than assessing whether the resident has the knowledge, the faculty may be assessing the application of the knowledge.

The ITE scores can be used in identifying and providing added education for residents who have trouble with the obstetric portion of the ITE, specifically those who also performed poorly on the USMLE Step 2. Special assignments and focused education efforts can enhance these individuals’ learning environment. Further analysis is needed to assess the efficacy of this type of intervention. Residents and faculty members have embraced the analysis, and the effect of changes in the curriculum will be evaluated in a future study.

Our study has several limitations. The results are from a single institution and may not be generalizable across settings. At the same time, the ITE is a national test, and the concept of evaluating the rotation is applicable to other institutions and broadly useful to program directors. Another limitation to the study is that the ITE is not a high-stakes test, and some residents may feel that there are no significant consequences for poor performance on the test. These results may change if promotion to the next year becomes dependent upon performance on the ITE. Another limitation is that our study could not eliminate the presence of possible confounding variables not related to the obstetric-anesthesia rotation, such as residents’ added self-directed learning or clinical experience, which may independently affect ITE scores.

Conclusion

By providing subtest scores, the American Board of Anesthesiology provided an opportunity for the program director and rotation director to evaluate the educational component of the rotation. Reporting individual section scores provides an opportunity for the faculty responsible for the obstetric-anesthesia rotation to evaluate the effectiveness of the educational program and to assess the effect of curricular changes.

References

- Hall JR, Cotsonis GA. Analysis of residents’ performances on the In-Training Examination of the American Board of Anesthesiology-American Society of Anesthesiologists. *Acad Med*. 1990;65(7):475-477.

- 2 Aiken LR, Groth-Marnat G. *Reliability and Validity in Psychological Testing and Assessment*. 12th ed. Boston, MA: Pearson; 2006:97–98.
- 3 Aiken LR, Groth-Marnat G. *Reliability and Validity in Psychological Testing and Assessment*. 12th ed. Boston, MA: Pearson; 2006:90–91.
- 4 McClintock JC, Gravlee GP. Predicting success on the certification examinations of the American Board of Anesthesiology. *Anesthesiology*. 2010;112(1):212–219.
- 5 Swanson DB, Case SM, Melnick DE, Volle RL. Impact of the USMLE step 1 on teaching and learning of the basic biomedical sciences: United States Medical Licensing Examination. *Acad Med*. 1992;67(9):553–556.
- 6 Cuddy MM, Dillon GF, Clauser BE, et al. Assessing the validity of the USMLE step 2 clinical knowledge examination through an evaluation of its clinical relevance. *Acad Med*. 2004;79(suppl 10):S43–S45.
- 7 Accreditation Council for Graduate Medical Education. Outcome Project. Available at: <http://www.acgme.org/outcome/comp/compMin.asp>. Accessed December 29, 2008.