Significantly Improved American Board of Surgery In-Training Examination Scores Associated With Weekly Assigned Reading and Preparatory Examinations

Christian de Virgilio, MD; Bruce E. Stabile, MD; Roger J. Lewis, MD, PhD; Colleen Brayack, MS

Hypothesis: Weekly reading assignments combined with weekly examinations can significantly improve American Board of Surgery In-Training Examination (ABSITE) scores among a group of residents already scoring above the national average.

Design: Prospective educational study of surgical residents.

Intervention: Beginning in July 2001, residents were given weekly reading assignments, followed by a multiple-choice examination, prepared and scored by the program director.

Main Outcome Measure: Mean change in ABSITE scores from 2001 to 2002.

Results: The mean total test ABSITE scores significantly improved from 58.7% in 2001 to 75.2% in 2002 (P = .008). The improvement was significantly greater in the basic science portion of the ABSITE (57.6% in 2001 vs 72.5% in 2002, P = .04) than in the clinical management section (57.6% in 2001 vs 68.9% in 2002, P = .11). There were no differences in mean ABSITE scores in other years (1999 to 2000 or 2000 to 2001). There was no correlation between the change in ABSITE scores from 2001 to 2002, level of residency training, United States Medical Licensing Examination scores, or performance on weekly preparatory examinations.

Conclusion: Weekly reading assignments combined with weekly preparatory examinations significantly improved mean overall ABSITE scores among a group of residents who were already scoring above the national average.

Arch Surg. 2003;138:1195-1197

Several previous studies have assessed factors that affect American Board of Surgery In-Training Examination (ABSITE) scores. Resident survey findings suggest that performance is affected by conference attendance, amount of sleep, anxiety level, amount of study, and previous examination results. Good performance on the ABSITE is important as there is evidence that a surgical resident's performance correlates with success on subsequent board examinations.

In recent years, resident performance on the ABSITE at Harbor-UCLA Medical Center has averaged just above the national average. Despite reasonably good test results, we embarked on an educational endeavor to improve ABSITE scores among our categorical residents. We hypothesized that weekly reading assignments combined with weekly examinations would significantly improve mean ABSITE scores (year 2002) compared with scores from previous years (2001, 2000, and 1999) in a group that was already performing at or above the national average. We further hypothesized that the actual weekly test scores would correlate with ABSITE performance.

METHODS

On a weekly schedule, beginning in July 2001 through January 2002, residents were assigned to read a chapter in a major surgery textbook, with the eventual goal of covering all areas of surgery within the period. A 15-question, multiple-choice examination (MCE) was prepared by the program director (PD) each week, in a format similar to that of the ABSITE, which covered the material in the assigned chapter. Residents were given 15 minutes in which to complete the closed-book MCE, under the proctoring of the PD. On completion of the MCE, the tests were exchanged with and corrected by an adjacent resident. The PD announced the correct answers and the number of questions answered correctly was written at the top of the test.
training.

Univariate descriptive statistics were expected to improve for an individual from year to year as he or she progresses in residency. The mean test scores were used as a measure of proficiency on the weekly examinations. The overall mean ABSITE score was used as a final compendium of miscellaneous topics.

Twenty-eight MCEs were administered, each with 15 questions. Subjects covered included small bowel, retroperitoneum and spleen, liver, biliary and pancreas, head and neck, trauma, shock, peripheral arterial disease, skin and soft tissue, surgical infection, surgical complications, cardiothoracic and pediatric surgery, critical care, thyroid and parathyroid disease, esophagus, adrenal and pituitary, neurosurgery and urology, breast, fluid and electrolytes, orthopedics and plastic surgery, hernia, stomach, and lymphatic and venous topics. Three additional 30-question MCEs were given in the last weeks before the ABSITE, which served as a final compendium of miscellaneous topics.

Data from individual residents’ scores on weekly MCEs, number of MCEs taken, United States Medical Licensing Examination 1 and 2 scores, year of training, and ABSITE scores for 2002 and for 2001, 2000, and 1999 (if applicable) were recorded on an Excel spreadsheet (Microsoft Corp, Redmond, Wash). The mean scores on the MCE and the number of MCEs were recorded. The data were translated to native SAS format using DBMS/COPY (Conceptual Software, Inc, Houston, Tex). All data analyses were performed using SAS version 8.1 (SAS Institute, Cary, NC).

Because not all residents completed all weekly MCEs, primarily because of ongoing clinical responsibilities, the mean test score was used as a measure of proficiency on the weekly examinations. Changes in national percentile rankings of ABSITE scores between years were considered the outcome of interest. The mean percentile score was analyzed rather than the percentage correct or the standard score, as these latter scores would be expected to improve for an individual from year to year as he or she progresses in residency. Univariate descriptive statistics were calculated for all numerical variables. In addition, numerical variables were compared between postgraduate year (PGY) classes using nonparametric testing (Wilcoxon signed rank test and the Spearman rank correlation coefficient), used to detect correlations between PGY level, United States Medical Licensing Examination scores, ABSITE scores, mean weekly test scores, number of quizzes taken, and changes in percentile ranking on the ABSITE. Nonparametric testing was used to avoid sensitivity to extreme values and to allow us to use national percentile rank for comparison. Changes in ABSITE national percentile scores from 2001 to 2002 were analyzed using paired observations. Paired measurements of the same individual residents’ percentile scores were used in comparing 2001 and 2002 scores to avoid bias that may have otherwise occurred due to secular trends in the knowledge bases of classes of residents enrolled in the program. P = .05 was considered statistically significant, and no adjustment was made for multiple comparisons.

Results who did not take the ABSITE before 2002 were excluded from the analysis, because there were no comparison data. Therefore, interns were excluded, as they had not previously taken the ABSITE. Likewise, residents who were undertaking a year of research at an outside institution and did not participate in the MCEs were also excluded. Resident names were not identified. Scores on the MCEs were not included in the overall resident evaluation process.

### RESULTS

The mean ABSITE scores by PGY level for 1999 to 2002 are shown in Table 1. The overall mean ABSITE score was 75.2% for all residents in 2002. This was significantly higher than the 2001, 2000, and 1999 scores (P = .008). There were no differences in the mean ABSITE scores from 1999 to 2000 (P = .9) or from 2000 to 2001 (P = .5). In fact, there was a small but nonsignificant downtrend in the overall mean scores during those years.

The mean percentile score on the basic science section of the examination was 72.5% in 2002 vs 57.0% in 2001 (P = .04). The mean percentile score on the clinical management section was 68.9% in 2002 vs 57.6% in 2001 (P = .11).

The largest increase in mean scores in 2002 occurred among the PGY-4 residents, whose mean percentile score increased from 61.0% in 2001 to 89.5% in 2002. The smallest increase in mean scores was noted among the PGY-5 residents, whose mean score increased only from 50.6% to 54.0%. Ten (53%) of 19 residents scored at or above the 80th percentile on ABSITE 2002, compared with only 2 (11%) of 19 residents on ABSITE 2001 (P = .005).

**Table 1. American Board of Surgery In-Training Examination (ABSITE) Scores (1999-2002) by Postgraduate Year (PGY) Level**

<table>
<thead>
<tr>
<th>PGY Level as of 2002</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>71.3</td>
<td>82.5</td>
</tr>
<tr>
<td>3</td>
<td>NA</td>
<td>48.2</td>
<td>55.7</td>
<td>78.7</td>
</tr>
<tr>
<td>4</td>
<td>71.0</td>
<td>61.3</td>
<td>61.0</td>
<td>89.3</td>
</tr>
<tr>
<td>5</td>
<td>58.4</td>
<td>71.2</td>
<td>50.6</td>
<td>54.0</td>
</tr>
<tr>
<td>Overall</td>
<td>63.9</td>
<td>60.1</td>
<td>58.7</td>
<td>75.2†</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable, as these residents were not yet in training.

*Data are given as mean percentile scores.
†P = .008.
‡P = .04.
§P = .008.

**Table 2. Comparison of 2001 and 2002 American Board of Surgery In-Training Examination Scores**

<table>
<thead>
<tr>
<th>Year</th>
<th>Basic Science</th>
<th>Clinical Management</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>57.6</td>
<td>57.6</td>
<td>58.7</td>
</tr>
<tr>
<td>2002</td>
<td>72.5†</td>
<td>68.9‡</td>
<td>75.2§</td>
</tr>
</tbody>
</table>

*Data are given as mean percentile scores.
†P = .04.
‡P = .11.
§P = .008.

At the time of the study, there were 25 categorical residents in our program (21 categorical residents in the clinical years and 4 categorical residents performing research). All 4 interns were excluded from the primary analysis, as they had no prior ABSITE scores for comparison. Two research residents were excluded because they were performing research at outside institutions and did not participate in the MCEs. The remaining 19 residents were included in the study.

**COMPARISON OF ABSITE 2002 WITH PREVIOUS YEARS**

The mean ABSITE scores by PGY level for 1999 to 2002 are shown in Table 1. The overall mean ABSITE score was 75.2% for all residents in 2002. This was significantly higher than the 2001, 2000, and 1999 scores (P = .008). There were no differences in the mean ABSITE scores from 1999 to 2000 (P = .9) or from 2000 to 2001 (P = .5). In fact, there was a small but nonsignificant downtrend in the overall mean scores during those years.

The mean percentile score on the basic science section of the examination was 72.5% in 2002 vs 57.0% in 2001 (P = .04). The mean percentile score on the clinical management section was 68.9% in 2002 vs 57.6% in 2001 (P = .11).

The largest increase in mean scores in 2002 occurred among the PGY-4 residents, whose mean percentile score increased from 61.0% in 2001 to 89.3% in 2002. The smallest increase in mean scores was noted among the PGY-5 residents, whose mean score increased only from 50.6% to 54.0%. Ten (53%) of 19 residents scored at or above the 80th percentile on ABSITE 2002, compared with only 2 (11%) of 19 residents on ABSITE 2001 (P = .005).
CORRELATION WITH CHANGE IN MEAN ABSITE SCORES FROM 2001 TO 2002

There was no correlation between the change in mean ABSITE scores from 2001 to 2002 and scores on the United States Medical Licensing Examination 1 (0.17, \( P = .49 \)) or 2 (\(-0.06, P = .89\)), PGY level (\(-0.17, P = .49\)), or 1999 (0.28, \( P = .4 \)) or 2000 (\(-0.18, P = .54\)) ABSITE scores. There was also no correlation with the mean overall score on the MCEs or with the number of MCEs taken. Of note, there was also no correlation between PGY level and the mean score on the weekly MCEs.

The present study demonstrates that the institution of weekly reading assignments combined with weekly review examinations significantly improved mean ABSITE scores in a group of surgical residents who were already performing above the national average. The mean ABSITE percentile score increased from 58.7% to 75.2%. Scores on the basic science portion of the examination improved more than the clinical management scores. There was a remarkable jump in the number of residents scoring above the 80th percentile. In 2002, more than half of the residents (10/19) scored above the 80th percentile nationally, compared with only 2 (11%) of the residents in the prior year. When analyzing the trend in ABSITE scores from previous years, there were no significant changes in mean ABSITE scores from 1999 to 2000 or from 2000 to 2001.

The present study differs from previously published studies in several respects. First, we included only residents who had taken at least 1 prior ABSITE. This allowed for a valid comparison. Second, in addition to assigned reading, a weekly examination was prepared and administered by the PD, attendance records were taken, and weekly scores were tallied. Therefore, a specific new quantifiable intervention was undertaken to determine if ABSITE scores would improve.

The reason for the significant rise in ABSITE scores in 2002 may be multifactorial. Certainly, one may argue that the existence of a specific reading assignment, coupled with a weekly review prepared by the PD, sent a message to the residents as to the significance of ABSITE preparation. The format of the weekly MCEs themselves may have familiarized and prepared the residents for the ABSITE format. Positive reinforcement, in the form of public praise for top performances, may have induced residents to better prepare. In addition, following each examination, the PD provided a review of the questions and emphasized important points. Interestingly, the actual scores on the MCEs did not correlate with ABSITE scores. This finding suggests that the process of creating a formal reading program with weekly examinations and review was more important in improving the scores than the actual content of the MCEs. One potential weakness of our study is that we did not quantify the exact amount each resident read.

Previous studies have analyzed factors associated with improved ABSITE scores. Most of these studies have focused on survey responses before and after taking the ABSITE. Godellas and Huang demonstrated via questionnaires that conference attendance, previous performance, probationary status, amount of sleep, and amount of study were significant in explaining most of the variance in ABSITE scores. Hirvela and Becker found that programmed reading significantly improved ABSITE scores. Their study differs from the present one in several important respects. The present study compared the year-to-year change in individual residents’ ABSITE performances using national percentile scores, which were analyzed using paired measurements and nonparametric comparisons. In the study by Hirvela and Becker, improvement in ABSITE scores for interns in 2 separate periods was assessed with parametric statistical analysis using standard scores rather than national percentiles. Because our study used paired measurements, interns were excluded, as they had not previously taken the ABSITE. In addition, only residents who took part in the MCEs were included. Therefore, chief residents who graduated in 2001 were excluded, because they did not participate in the MCEs, which were instituted after they graduated. There was no significant change in ABSITE scores in our program from 1999 to 2000 or from 2000 to 2001, before the reading program was instituted. In the present study, the PD created weekly examinations that were mandatory for all residents, and scores were tallied. More important, in the present study, residents were already performing at or above the national average (mean, 58%) before institution of the reading program and MCEs, demonstrating that programmed reading can further improve scores. The high scores of our residents did not correlate with previous United States Medical Licensing Examination scores. In addition, the present study demonstrated that ABSITE scores could be significantly improved during a short period, as the programmed reading and MCEs were instituted only 6 months before the scheduled ABSITE.

Accepted for publication March 22, 2003.

Corresponding author: Christian de Virgilio, MD, Department of Surgery, Harbor–UCLA Medical Center, 1000 W Carson St, Mail Box 25, Torrance, CA 90509 (e-mail: cdevirgilio@rei.edu).

REFERENCES