Impact of an Online Question Bank on Resident In-Training Exam Performance

Nicole P. Scott, Terrell W. Martin, Alison M. Schmidt and Anthony L. Shanks

Indiana University School of Medicine, Indianapolis, IN, USA.

Journal of Medical Education and Curricular Development
Volume 10: 1–5

DOI: 10.1177/23821205231206221

ABSTRACT

OBJECTIVE: In-training exams (ITEs) are administered annually to Obstetrics and Gynecology (OBGYN) residents and have been demonstrated to correlate with success on licensing examinations. Our study objective was to determine the impact of a question bank and mock exam on the performance of Council on Resident Education in Obstetrics and Gynecology (CREOG) ITEs. Secondarily, we investigated the correlation between the extent of question bank usage and performance on the exam.

METHODS: Pre–post intervention study of resident performance on CREOG ITE before and after implementation of the question bank and mock ITE at Indiana University in 2018. Performance was measured as year-to-year improvement in percent correct on ITE exams. Scores were excluded if a resident did not have a prequestion bank score report or if they did not sit for all eligible ITE exams.

RESULTS: There were 51 OBGYN residents at Indiana University during the study period, with 38 available for analysis (75%). Before implementation, average year-to-year improvement for PGY1-2, PGY2-3 and PGY3-4 classes were 0.60%, 1.0% and −1.6%, respectively. After implementation, all resident classes had significant improvements in ITE scores of 6.6% ($P < .01$), 9.0% ($P < .01$), and 7.2% ($P < .01$), respectively. There was a moderate program-wide correlation between the number of questions completed and the percent improvement on the ITE of $R = 0.36$ ($P = .046$).

CONCLUSIONS: Our study demonstrated that access to a question bank and mock ITE significantly improved CREOG ITE performance of OBGYN residents at Indiana University.

KEYWORDS: Question bank, in-service training exam, retrieval practice, quiz, residency, education

RECEIVED: December 20, 2022. **ACCEPTED:** June 12, 2023.

TYPE: Original Research Article

FUNDING: The authors received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Anthony, L, Shanks, Indiana University School of Medicine, 550 University Blvd, University Hospital Suite 2440, Indianapolis, IN 46202, USA.
Email: ashanks@iupui.edu

Introduction

Obstetrics and gynecology (OBGYN) residents complete an annual in-training examination (ITE) through a nationwide exam issued by the Council on Resident Education in Obstetrics and Gynecology (CREOG). The ITE provides a measure to evaluate cognitive knowledge and assess a program's didactic curriculum. A strong performance on ITEs has been demonstrated to predict a passing score on the American Board of Obstetrics and Gynecology (ABOG) Qualifying Exam which in turn is used by ACGME to help inform accreditation decisions.[1,2] Given this, OBGYN residency training programs can use these exams to identify gaps in resident education to ensure resident success on qualifying board exams.

Currently, there is no standardized OBGYN residency education curriculum and no consensus in the literature on tools to improve performance on ITE.[3] The CREOG Educational Objectives provide a framework for resident instruction, but didactic curriculum vary between residency programs.[4] Retrieval methods, including the use of a question bank, have been demonstrated to be effective strategies for learning but the optimal implementation during residency is unknown.[5–8] This study aimed to improve resident performance on the CREOG ITE through the implementation of a question bank and mock ITE.

We hypothesize that universal access to a question bank and mock ITE will improve resident performance on the CREOG ITE. Furthermore, we hypothesize that the extent of question bank usage as measured by number of questions completed will positively correlate with the magnitude of ITE improvement.

Methods

We performed a pre–post intervention study of the CREOG ITE score reports from 2017 to 2019 for all OBGYN residents at Indiana University. The Indiana University Institutional Review Board (IRB) reviewed the protocol and determined the study to be IRB-exempt with waived informed consent as the study fell into research conducted in common educational settings involving normal educational practices (IRB# 1903206729). CREOG ITE examinations are given during January of each year. Beginning in July 2018, residents were given unlimited access to the general TrueLearn question bank. Didactic content and reviews were unchanged during the study and dedicated time was not provided for use of the question bank. Residents had 3 h of dedicated didactic time per week with traditional lectures provided by faculty and fellows. The schedule adheres to topics provided in the CREOG Educational Objectives.

The TrueLearn question bank is a commercial product with over 1,200 unique questions available for self-study and access both online and through a mobile device.[9] Administrators were able to track the use of the question bank, but use was not mandatory. The total number of practice questions performed up to the ITE test date was collected. As a part of the implementation, a simulated ITE through the question bank provider was administered two months prior to the CREOG ITE. This simulated ITE utilized 150 novel questions and was provided in a timed manner during protected didactic time. The questions were selected by TrueLearn and covered a blueprint similar to the CREOG ITE. Performance on the simulated ITE was not measured, but all residents completed the mock ITE.

CREOG ITE performance was evaluated as individual improvements, assessed by the percent of correct questions answered in the 2018 ITE compared to the 2019 ITE. This year-to-year change was calculated for each individual resident and mean changes were reported for each post-graduate year (PGY), termed postintervention. The preintervention cohort consisted of the average improvement from one PGY year to the next (PGY 1 to 2, PGY 2 to 3, and PGY 3 to 4) from 2017 to 2018. The postintervention cohort consisted of the average improvement from one PGY year to the next from 2018 to 2019.

Statistical Analysis

A dependent t-test was conducted to compare the average annual improvement in CREOG ITE scores before and after question bank implementation. Correlation between question bank usage (defined by the number of questions performed) and improvement on the 2019 ITE (defined by percent increase) were calculated for each individual using Pearson correlation testing and a linear regression was used to determine their association. We used SPSS Statistics 26[10] to perform all statistical analyses.

This study was submitted to Indiana University IRB and found to be exempt. All scores on the CREOG ITE and the question bank were deidentified.

Results

The Indiana University OBGYN Residency Program has 10 residents per year. Over the course of the study, there were a total of 51 OBGYN residents (one class had a temporary increase to 11). Those who did not have an ITE score report prior to the implementation of the question bank and those who did not sit for all eligible exams were excluded from this study. A total of 13 residents were excluded, leaving 38 OBGYN residents in the analysis. The overall mean ITE for all PGY levels was 61.4% (SD 6.9) prior to the implementation of the TrueLearn question bank and the overall mean ITE was 64.5% (SD 4.9, 3.1% increase, $P = .02$) postimplementation.

Prior to the implementation of the question bank, the mean ITE improvement was 0.6% (SD = 3.1) for the PGY1 to 2 cohort and 1.0% (SD = 3.3) for the PGY 2 to 3 cohort. The residents in the PGY3 to PGY4 cohort demonstrated a mean decrease in ITE performance of 1.6% (SD = 2.6) prior to access to this question bank.

Following the implementation of the question bank, there was a statistically significant improvement in the CREOG ITE for all levels of residency training. The PGY 1 to 2 showed a mean improvement of 6.6% (SD = 3.1), with an absolute increase of 6.0% with the question bank, $P = .018$, 95% CI [2.17–9.72] (Figure 1). Residents in the PGY 2 to 3 cohort showed a postintervention improvement of 9.0% (SD = 4.7), demonstrating an absolute increase of 8.0% with the question bank, $P < .01$, 95% CI [4.0–11.96] (Figure 2). The PGY 3 to 4 cohort demonstrated a mean improvement of 7.2% (SD = 3.6), with an absolute increase of 8.9%, $P < .01$, 95% CI [5.56–12.14] (Figure 3).

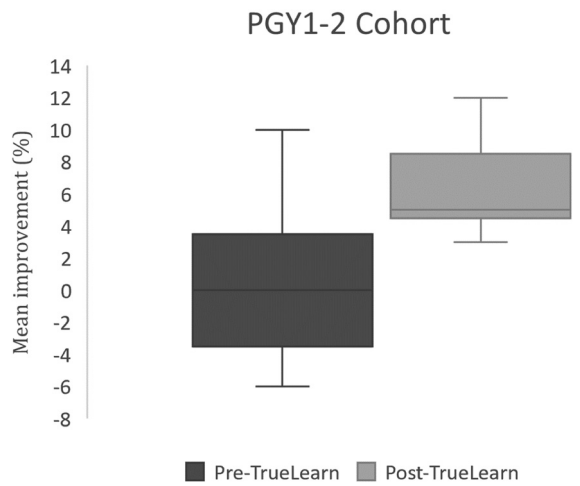

Figure 1. Mean improvement on ITE performance from PGY 1 to 2 before (0.60 [SD 5.0]) and after (6.6 [SD 3.1]) implementing TrueLearn ($P < .01$).

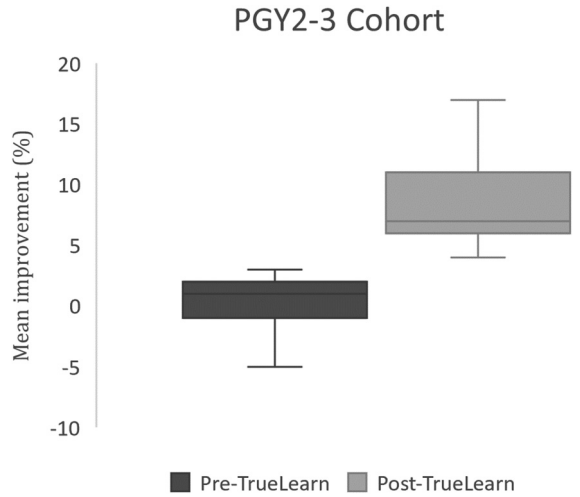

Figure 2. Mean improvement on ITE performance from PGY 2 to 3 before (1.0 [SD 3.3]) and after (9.0 [SD 4.7]) implementing TrueLearn. ($P < .01$).

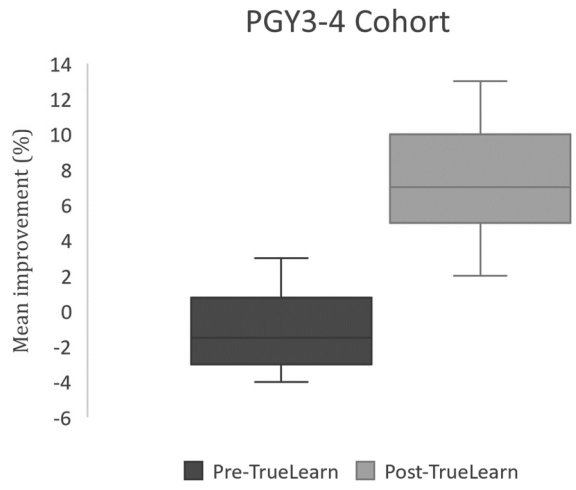

Figure 3. Mean improvement on ITE performance from PGY 3 to 4 before (−1.6 [SD 2.6]) and after (7.2 [SD 3.6]) implementing TrueLearn ($P < .01$).

Table 1. Correlation of TrueLearn Usage and Improvement on CREOG ITE.

	Mean post-TrueLearn improvement	Average number of questions completed	Pearson correlation coefficient	*P*-value
PGY 3-4	7.2	397	0.29	.41
PGY 2-3	9.0	965	0.50	.14
PGY 1-2	6.6	503	−.019	.58
Total	7.6	618	0.361	.046[a]

CREOG: Council on Resident Education in Obstetrics and Gynecology; ITE: in-training exam; PGY: post-graduate year.
[a] Significance $P < .05$.

Pearson correlation coefficients were calculated to assess the relationship between the extent of question bank usage, as measured by the number of practice questions completed, and percentage improvement on the ITE. Overall, there was a positive moderate correlation between the number of questions accessed in the question bank and the percentage point improvement on the CREOG ITE; $r = .361$, $P = .046$, though this was not significant for any individual year (Table 1). A scatter plot demonstrates a linear correlation between the number of questions answered in the question bank and the CREOG ITE result (Figure 4).

Discussion

Our study demonstrates a statistically significant increase in year-to-year improvement on CREOG ITE after the implementation of a question bank and mock ITE. When assessing independent question bank usage, our data demonstrates a positive correlation with ITE performance. The focus of our study was to discern if universal access to an online question bank and mock ITE proved beneficial to the global performance of the CREOG ITE at our institution. These findings support our hypotheses.

The positive effects of practice question use on scores on the United States Medical Licensing Examination (USMLE) Step 1 examination have previously been demonstrated.[11–13] Similarly, this benefit has been shown in performance on the American Board of Surgery In-Training Examination (ABSITE), the CREOG ITE equivalent for general surgery residents.[14,15] Imran et al. have identified that the use of the TrueLearn question bank specifically is associated with a greater percentage correct on ABSITE.[16]

Our findings differ from a separate evaluation of the same question bank.[17] Both studies found a positive correlation between question bank usage and CREOG scores, however, the authors did not find a statistically significant increase in CREOG performance with question bank implementation. Though our study also provided unlimited access to the question bank, participants in our study differed as they had access to a mock ITE. The opportunity to simulate exam conditions may have helped to reduce anxiety or improve time management for some test takers, leading to the difference in scores observed in our study.

Additionally, we chose to report mean improvement from one PGY year to the next, with the control group representing the same PGY transition year, as resident knowledge is naturally expected to change throughout this time. Using the average improvement in percent correct over a year rather than absolute ITE scores allows for individual feedback for residents and mitigates potential outliers for individual scores that might skew the mean. Scores on the CREOG are relatively stable over time. In 2019, the average CREOG percent correct was 55% for PGY1, 60% for PGY2, 64% for PGY3, and 66% for PGY4. This represents a national average increase from PGY 1 to 2 of 5%, PGY 2 to 3 of 4% and PGY 3 to 4 of 2%. Our results in the post-TrueLearn implementation group of 6.6%, 9.0%, and 7.2% exceeded these thresholds.

Our study has several limitations. First, our sample size was small with no a priori size calculation and limited to a single institution. Although the resident class at Indiana University is large at 10 persons per year, it cannot be concluded whether the question bank and mock ITE filled a gap in resident education specific to the Indiana University residency program, or if this service would be beneficial across all OBGYN programs. Our R square of 0.13 suggests that only 13% of the variability in percent improvement on the CREOG ITE was explained by TrueLearn usage therefore other factors could explain the improvement aside from the question bank access. Additionally, it is unknown from this study whether any peer-reviewed question bank would yield similar improvements, or if this benefit is limited to TrueLearn only.

One potential area of concern for many residency programs could be the cost of providing universal access to a question bank for its residents. A one-year subscription to the OBGYN ITE TrueLearn question bank is $320 per resident.[9]

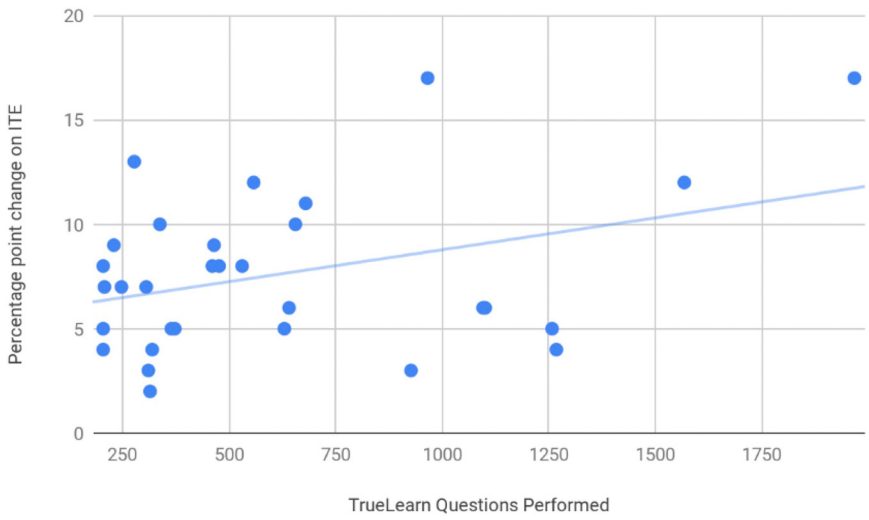

Figure 4. Correlation between individual TrueLearn usage and their percentage point improvement on CREOG ITE; $r = .361$, $P = .046$.

This may be prohibitive for many programs, particularly those with larger resident cohorts. A potentially less cost-prohibitive option may be to use shorter subscription plans, as TrueLearn offers one-, three- and six-month options. Other question banks exist with different payment structures, such as Board Vitals or Rosh Review, with the latter offering a mock ITE in addition to question bank access.

Historically, standardized multiple choice exams, particularly the USMLE Step 1, have sometimes been used broadly as a screening metric by which OBGYN residencies offer interviews.[18] One reason for this is that those who score highly on STEP 1 tend to perform well on the CREOG ITE and ABOG qualifying exams.[19,20] However, with the implementation of Pass/Fail grading on the USMLE Step 1 exam, residency programs will no longer have this metric during the recruitment process. This study suggests that with additional resources, including a question bank and simulated ITE, resident performance can be positively affected by retrieval practice. Future studies will continue to refine the optimal implementation of these resources.

Conclusions

Our study demonstrated that access to a question bank and mock ITE significantly improved the CREOG ITE performance of OBGYN residents at Indiana University.

Acknowledgments

The authors want to acknowledge the work of our residents and Program Manager, Ali Cotton, on this project.

ORCID iD

Terrell W. Martin https://orcid.org/0000-0003-1955-8007

REFERENCES

1. McCrary HC, Colbert-Getz JM, Poss WB, Smith BK. A systematic review of the relationship between in-training examination scores and specialty board examination scores. *J Grad Med Educ*. 2021;13(1):43-57. doi:10.4300/JGME-D-20-00111.1
2. Lingenfelter BM, Jiang X, Schnatz PF, O'Sullivan DM, Minassian SS, Forstein DA. CREOG in-training examination results: contemporary use to predict ABOG written examination outcomes. *J Grad Med Educ*. 2016;8(3):353-357. doi:10.4300/JGME-D-15-00408.1
3. Mutter O, Hylton J, Jeronis S, Jaspan D, Rose M. A standardized obstetrics and gynecology residency education curriculum: multi-institutional needs assessment survey. *Reproductive Medicine*. 2020;1(3):195-204. doi:10.3390/reprodmed1030015
4. CREOG Educational Objectives. Accessed September 21, 2021. https://www.acog.org/en/education-and-events/creog/curriculum-resources/creog-educational-objectives.
5. Larsen DP, Butler AC, Roediger HL III. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med Educ*. 2009;43(12):1174-1181. doi:10.1111/j.1365-2923.2009.03518.x
6. Dunlosky J, Rawson KA, Marsh EJ, Nathan MJ, Willingham DT. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol Sci Public Interest*. 2013;14(1):4-58. doi:10.1177/1529100612453266
7. Larsen DP, Butler AC, Roediger HL III. Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Med Educ*. 2013;47(7):674-682. doi:10.1111/medu.12141
8. Karpicke JD, Blunt JR. Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*. 2011;331(6018):772-775. doi:10.1126/science.1199327
9. Team TT. CREOG in-training exam practice questions | OBGYN Exam Prep. TrueLearn SmartBanks for Medical Exam Preparation. Accessed May 23, 2021. https://truelearn.com/obgyn/obgyn-creog-training/
10. IBM Corp. IBM SPSS Statistics for Windows, Version 26.0. IBM Corp; 2021.
11. Baños JH, Pepin ME, Van Wagoner N. Class-wide access to a commercial step 1 question bank during preclinical organ-based modules: a pilot project. *Acad Med*. 2018;93(3):486-490. doi:10.1097/ACM.0000000000001861
12. Bonasso P, Lucke-Wold B, Reed Z, Bozek J, Cottrell S. Investigating the impact of preparation strategies on USMLE step 1 performance. *MedEdPublish*. 2015;4(1). doi:10.15694/mep.2015.004.0005
13. Jackson F, Duane E, Harmon R, Kollar RA, Rainville NM, Smith RM. Resources that improve medical board licensing examination performance. *Cureus*. 2019;11(10):e5927. doi:10.7759/cureus.5927
14. Chang D, Kenel-Pierre S, Basa J, et al. Study habits centered on completing review questions result in quantitatively higher American Board of Surgery In-Training exam scores. *J Surg Educ*. 2014;71(6):e127-e131. doi:10.1016/j.jsurg.2014.07.011
15. Flentje AO, Caturegli I, Kavic SM. Practice makes perfect: introducing a question bank for ABSITE preparation improves program performance. *J Surg Educ*. 2020;77(1):54-60. doi:10.1016/j.jsurg.2019.09.005
16. Imran JB, Madni TD, Taveras LR, et al. Assessment of general surgery resident study habits and use of the TrueLearn question bank for American Board of

Surgery In-Training exam preparation. *Am J Surg*. 2019;218(3):653-657. doi:10.1016/j.amjsurg.2019.02.031
17. Green I, Weaver A, Kircher S, et al. Impact of question bank use for in-training examination preparation by OBGYN residents - a multicenter study. *J Surg Educ*. 2022;79(3):775-782. Published online January 24, 2022. doi:10.1016/j.jsurg.2021.12.014
18. National Resident Matching Program, data release and research committee: results of the 2020 NRMP Program Director Survey. National Resident Matching Program, Washington, DC. 2020.
19. Armstrong A, Alvero R, Nielsen P, et al. Do U.S. medical licensure examination step 1 scores correlate with council on resident education in obstetrics and gynecology in-training examination scores and American board of obstetrics and gynecology written examination performance? *Mil Med*. 2007;172(6):640-643. doi:10.7205/milmed.172.6.640
20. Spellacy WN, Downes KL. United States medical licensing examination scores as a predictor of performance on the annual council of resident education in obstetrics and gynecology examinations. *J Reprod Med*. 2014;59(1-2):17-19.