The Journal of Education in Perioperative Medicine

Original Research

Correlation Between Question Bank Usage and In-Training Exam Scores for Anesthesiology Residents

Daniel P. Walsh, MD
Nanaama O'Hene, MD
Vanessa T. Wong, BS
Sara E. Neves, MD

Introduction

In order to become a board-certified anesthesiologist in the United States, anesthesiologists must take and pass 3 staged examinations administered by the American Board of Anesthesiology (ABA).[1] These exams are the BASIC Exam, which is a written exam taken during the first year of residency; the ADVANCED Exam, which is a written exam taken after graduation from residency; and the APPLIED Exam, which is an oral exam and Objective Structured Clinical Exam taken after passing the ADVANCED Exam. The In-Training Examination (ITE) for anesthesiology residents, which many programs require residents to take, has been given annually by the ABA since it was developed in 1975.[2] Trainees and residency programs receive feedback on performance on the ITE that can assist in identifying potential gaps in knowledge that can inform individualized formative educational plans. The exam can be used longitudinally as an important indicator of trainees' academic progress through training. Additionally, the anesthesiology ITE is a useful predictor of success on the ABA written certification exams.[3–5] Whereas initially intended as a low-stakes formative assessment of trainees' knowledge, the ITE has increasingly been used by trainees and training programs to predict outcomes and prepare trainees for the high-stakes summative assessments of the ABA certification exams.[4] Moreover, trainees who want to complete a fellowship (or fellowships) are often asked to report their ITE scores as part of the application process, so performance on the ITE can potentially influence their qualification for fellowships even though the initial intentions of the exam were for it to be a low-stakes formative assessment.

Significant effort is spent by both trainees and educators in preparing for success on the ITE. There is some data among surgery and obstetrics/gynecology residents that increasing usage of question banks (Qbank) is associated with improved ITE scores with some data showing completion of a question bank set increasing ITE score percentile by as much as 20 percentiles.[6–8] This is likely due in part to the benefits of retrieval practice by which actively retrieving information from memory while answering the Qbank questions strengthens learning more than passive review as supported by a previous study on the effects of retrieval practice on exam scores.[9]

At our institution (Beth Israel Deaconess Medical Center, a tertiary care academic medical center in Boston, MA), anesthesiology residents are given resources to study for the ITE as part of their regular training. One resource is TrueLearn (TrueLearn, LLC, Charlotte, NC), a Qbank specifically designed to prepare trainees for medical examinations in different specialties, including anesthesiology. We sought to investigate how usage of the Qbank among anesthesiology residents correlates to performance on the anesthesiology ITE, specifically regarding whether there are inflection points of usage that correlate with low or high performance. This information could be used to develop expectations for trainees and educators on the effectiveness of efforts put into Qbanks. It may also help educators develop a more effective individualized study plan for trainees who are struggling academically.

Materials and Methods

This retrospective cohort study received institutional review board approval for exempt status with a waiver of documentation of informed consent by the Committee on Clinical Investigations at Beth Israel Deaconess Medical Center.

Data Collection

We gathered the number of Qbank (TrueLearn) questions completed by residents and their ITE percentile scores over the course of 3 academic years (2021–2022, 2022–2023, and 2023–2024). The number of questions completed is within each academic training year and is not cumulative over the training years. Of note, residents are allowed to complete QBank questions more than once. In addition, we gathered the ITE percentile scores of residents for these years from the residency program administration in our

continued on next page

continued from previous page

department, which receives these scores each year from the ABA after the residents complete the exam. Data (number of Qbank questions completed and ITE percentile score) from 131 residents were reviewed for a total of 155 records. Due to anonymization of the data, we were unable to link records to individual residents and, therefore, could not perform complete longitudinal analysis of performance across the years.

Statistical Analysis

Statistical analyses were performed with Stata/Special Edition 18 (StataCorp LP, College Station, TX). For the purpose of analysis, we grouped the number of Qbank questions completed by each resident each year into the following categories:

1. 0–100
2. 101–500
3. 501–1000
4. 1001–2000
5. >2000

The intention behind this categorization method was to assess broadly if there were differences on ITE performance among residents who answered very few questions (0–100), some questions (101–500), many questions (501–1000), a significant number of questions (1001–2000), or a very high number of questions (>2000). We performed a regression analysis using nonparametric series estimation with a B-spline basis. ITE percentile score was the dependent (continuous) variable, whereas number of Qbank questions completed (categorized as described above) was the independent (categorical) variable with using 0–100, 101–500, 501–1000, and 1001–2000 as base levels to make pairwise comparisons among the 5 categories. Because of the multiple (4) regressions performed, we applied the Bonferroni correction and considered $p < .0125$ as significant.

As a secondary analysis, we performed the same nonparametric series estimation for the individual classes (clinical anesthesia resident years 1 [CA-1], 2 [CA-2], and 3 [CA-3]). In addition, we also grouped the ITE percentile score of each resident each year into the following categories:

1. 0–10
2. 11–25
3. 26–50
4. 51–75
5. 76–90
6. >90

Then, we performed the same nonparametric series estimation with number of Qbank questions completed as the dependent (continuous) variable and ITE score (categorized as described above) as the independent (categorical) variable with using 0–10, 11–25, 26–50, 51–75, and 76–90 as base levels to make pairwise comparisons among the 6 categories. For all secondary analyses, we considered $p < .05$ as significant.

Results

ITE Percentile Scores by Number of Qbank Questions Completed for All Residents

ITE percentile scores of residents for each category of number of Qbank questions completed are summarized in Figure 1 as box-and-whisker plots with average ITE percentile score for each category marked with an "X." Overall, ITE percentile scores of residents who completed more than 500 Qbank questions were significantly higher than ITE percentile scores of residents who completed fewer than 500 Qbank questions (Table 1).

ITE Percentile Scores by Number of Qbank Questions Completed for Each Class

For CA-1s, ITE percentile scores of residents who completed more than 500 Qbank questions were significantly higher than ITE percentile scores of residents who completed fewer than 500 Qbank questions (Table 1). They were also significantly higher for residents who completed 101–500 Qbank questions than for residents who completed 0–100 Qbank questions.

For CA-2s, ITE percentile scores of residents who completed 501–2000 Qbank questions were significantly higher than ITE percentile scores of residents who completed fewer than 500 Qbank questions (Table 1).

For CA-3s, ITE percentile scores of residents who completed more than than 2000 Qbank questions were significantly higher than ITE percentile scores of residents in all other categories (Table 1).

Number of Qbank Questions Completed by ITE Percentile Scores

For All Residents

Number of Qbank questions completed by residents for each category of ITE percentile scores they achieved are summarized in Figure 2 as box-and-whisker plots with average number of Qbank questions completed for each category marked with an "X." The number of Qbank questions completed by residents who scored from the 11th to 99th percentile was significantly higher than the number of Qbank questions completed by residents who scored below the 10th percentile (Table 2). They were also significantly higher for residents who scored greater than the 90th percentile than those who scored between the 11th and 75th percentiles (Table 2).

For Each Class

For CA-1s, the number of Qbank questions completed by residents who scored from the 11th to 99th percentile was significantly higher than residents who scored in the 10th percentile or lower (Table 2).

For CA-2s, the number of Qbank questions completed by residents who scored greater than the 75th percentile was significantly higher than those who scored in the 75th percentile or lower, but they were not significantly different between those who scored between the 76th and 90th percentile and those who scored greater than 90th percentile (Table 2).

For CA-3s, the number of Qbank questions completed by residents who scored greater than the 90th percentile was significantly higher than those who scored in the 10th percentile or less and those who scored between the 26th and 50th percentiles and between the 51st and 75th percentiles (Table 2).

Discussion

The CA-1 residents who completed less than 100 and between 101 and 500 Qbank questions had ITE scores that were

continued on next page

continued from previous page

significantly lower than those in each of the other categories that completed more questions. Data from the ABA suggests that cohorts of trainees with ITE scores below approximately the 30th percentile begin having increasing rates of board exam failure.[5] Our data show that CA-1 residents who completed fewer than 100 Qbank questions scored in ITE percentile ranges that, based on ABA-published data, have been associated with BASIC Exam failure rates of up to 40%.[5] CA-1 residents who completed between 101 and 500 questions achieved average ITE percentile scores that, based on ABA data, have been associated with BASIC Exam failure rates of up to 19%.[5]

There are notable differences between the junior (CA-1) residents and senior (CA-2/3) residents in the effects of Qbank usage. For senior residents, the average ITE percentile score for those who completed fewer than 100 Qbank questions remained low, but it did increase for each year in training, likely reflecting learning that was unrelated to Qbank usage. However, the average remained below the 30th percentile, which may put them at increased risk of ADVANCED Exam failure, so there may be incentive for this cohort to improve their ITE performance still further.[5] As average ITE scores increased with training level, even the highest ITE scores among senior residents who completed fewer than 100 Qbank questions remained modest (above the 50th percentile) and were not in percentile ranges previously associated with increased risk of board certification examination failure. Some of the highest performers in the senior classes were able to perform well while doing fewer questions. This suggests that clinical training exposure over the years along with other methods of exam preparation besides Qbank questions may be enough for some residents to pass the board exams without doing targeted exam studying though many residents may benefit from targeted exam studying using Qbank questions. Alternatively, it may be possible that residents who perform modestly on the ITE while doing minimal Qbank questions may have been using the Qbank more efficiently, such as by using a more effective test mode in TrueLearn (timed mode, untimed mode, or tutor mode)[10] while doing the questions or by more explicitly assessing their knowledge while doing the questions and targeting their studying to specific areas for improvement. It is also possible that residents in this cohort may have been spending more time and effort using alternative methods of studying, such as reading, listening to podcasts or lectures, and/or reviewing notes. It could also be that these were residents who had high Qbank usage as CA-1 residents and no longer completed as many questions as senior residents. Unfortunately, our data did not allow for tracking which test mode was used by residents or tracking individual resident usage of the Qbank over the years. Whereas it is not clear if the use of the Qbank is specifically driving these effects or if increased Qbank usage indicates generally an increased engagement with the studying or exam preparation process overall, it may be reasonable for residency program leadership to use low Qbank usage as a prompt to check in with residents about their overall study strategies and study time allocation, especially with CA-1 residents who have done fewer than 100 Qbank questions.

Across all training years, the average number of questions done by those who scored above 75th percentile on the ITE was well above the number of questions in the complete Qbank. The lowest number of Qbank questions completed by a resident scoring above 90th percentile on the ITE was 725 questions, which is approximately 70% of the complete Qbank. This high volume of completed Qbank questions may have been influenced by baseline knowledge and/or question-answering efficiency. Residents with higher baseline knowledge and/or efficiency in answering questions may be able to answer questions more quickly, and this limits using the volume of completed Qbank questions as a proxy for study time or exam preparation intensity. Despite this, this positive association between the number of Qbank questions done and exam scores is in line with several learning theories. First, by initially attempting to answer a question, residents must retrieve information from memory (retrieval practice), which strengthens their neural pathways and memory. Moreover, repeating questions in the question bank likely reinforces the concepts in their memory.[11] Second, because the questions are available for residents to practice throughout the year, the intermittent use of questions and retrieval attempts likely strengthens their long-term memory through the spacing effect.[12] Third, the practice Qbank questions, along with their corresponding explanations, break down complex material into smaller, more manageable parts, and this helps reduce cognitive load for residents while they study.[13] Fourth, doing the practice questions also provides feedback to residents on what they actually know (as opposed to what they think they know) and helps identify knowledge gaps that can guide further focused study efforts, both of which support metacognition and self-regulated learning.[14]

Despite the overall positive association between number of Qbank questions completed and ITE score, there is a small cohort of residents across all classes that completed a large number of Qbank questions and still did not perform well on the ITE with some even performing at less than the 25th percentile. This suggests that, whereas our findings showed that residents who performed above average on the ITE completed a high volume of Qbank questions, completing a high volume of Qbank questions does not guarantee success. Therefore, residency programs and residents should not use high Qbank usage as a predictor of success on the ITE. Instead, residency programs may consider targeted check-ins with residents reporting low Qbank usage and discuss overall study engagement, such as use of other study methods or resources, use of the different Qbank test modes, or use of repeated exposure to topics to support retention and/or understanding.

Although we were unable to find prior data on this topic for anesthesiology residents, this data fits with what has been studied for surgery and obstetrics/gynecology residents. One study looked at the performance on test scores for obstetrics/gynecology residency programs before and after they made available an online Qbank.[8] This study showed no difference

continued on next page

continued from previous page

in program scores though individuals who did more questions had higher scores, generally. There was no data on what types of resources were being used to study prior to when the online Qbank was made available. This differs from another study on an obstetrics/gynecology residency program which showed an increase in year-to-year ITE scores for each residency class year after the availability of an online Qbank.[15] Our study is generally consistent with these studies on nonanesthesia graduate medical education programs in that it shows there is value in using a Qbank but that using a Qbank does not guarantee success.

This study is limited by the retrospective nature of the data collection, which restricts our ability to link scores to specific residents and account for repeated measures. This also does not allow us to have year-to-year data linking each resident's question usage during the prior years, which would have allowed us to evaluate whether there are either continued effects from doing many questions early in training or a compounding effect of doing questions year over year. This study is also limited by not having further data regarding other studying habits that residents had beyond the number of Qbank questions completed such as total time spent studying, time spent using non-Qbank resources, which Qbank test mode was used while doing the questions (timed, untimed, or tutor), and average time spent completing each question and reading the explanations as well as any other possible confounders.

In the future, a prospective study that obtains information on Qbank test mode usage, links resident usage and performance throughout their training years, and combines this with survey data on study habits accompanying Qbank usage data may be able to provide more insight into potential effects of the different ways residents used the Qbank questions and of the different (non-Qbank) resources that residents used to study. In addition, further research is needed to assess the generalizability of our results at other centers as well as the retention of knowledge gained through doing Qbank questions. Despite not having the benefits that would likely be gained from having all this further information, this current study does provide some insight into how much an anesthesiology resident should consider using review questions as part of their test preparation strategy.

Conclusions

There was a significant difference in ITE percentile scores between anesthesiology residents who complete a low number of Qbank questions and anesthesiology residents who complete a high number of Qbank questions. Those who completed fewer than 100 questions averaged lower than the 30th percentile. This effect was particularly pronounced for junior residents and may put those residents at greater risk of failing the BASIC Exam. Those who scored over the 90th percentile averaged using more than 1000 questions from the Qbank.

Acknowledgments

The authors thank the Department of Anesthesia, Critical Care and Pain Medicine at Beth Israel Deaconess Medical Center for their support of this study as well as the residents who were part of this study.

References

1. The American Board of Anesthesiology. Begin certification: demonstrate your commitment to excellence with the gold standard of certification in anesthesiology. https://www.theaba.org/begin-certification/. Accessed December 28, 2025.
2. Slogoff S. A history of the American board of anesthesiology certifying examinations. In: *The Wondrous Story of Anesthesia*. New York, NY: Springer; 2014;459-70.
3. McClintock JC, Gravlee GP. Predicting success on the certification examinations of the American Board of Anesthesiology. *Anesthesiology.* 2010;112(1):212-9.
4. 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.
5. Fiadjoe J, Minhaj M, Harman A. ABA update. https://www.theaba.org/wp-content/uploads/2024/11/SAAAPM-2024_Core_Web.pdf. Accessed May 5, 2025.
6. 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):653-7.
7. 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.
8. 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):775-82.
9. Nelson A, Mohammed A, An A, Traba C. Testing the effects of individual residents' retrieval practice on standardized examination scores. *Med Sci Educ.* 2024;34(3):647-52.
10. The TrueLearn Team. TrueLearn data lab: the best way to use question banks. https://truelearn.com/resource-library/truelearn-data-lab-which-test-mode-do-higher-performers-use/?srsltid=AfmBOootgbEUt6ewkmr4JfYmHnVkeBEohs1SjXghr9GIe6gvdI-UbLOZ. Accessed December 28, 2025.
11. Roediger HL, Butler AC. The critical role of retrieval practice in long-term retention. *Trends Cogn Sci.* 2011;15(1):20-7.
12. Carpenter SK, Cepeda NJ, Rohrer D, et al. Using spacing to enhance diverse forms of learning: review of recent research and implications for instruction. *Educ Psychol Rev.* 2012;24(3):369-78.
13. Sweller J. Cognitive load theory and individual differences. *Learn Individ Differ.* 2024;110:102423.
14. Stanton JD, Sebesta AJ, Dunlosky J. Fostering metacognition to support student learning and performance. *CBE Life Sci Educ.* 2021;20(2):fe3.
15. Scott NP, Martin TW, Schmidt AM, Shanks AL. Impact of an online question bank on resident in-training exam performance. *J Med Educ Curric Dev.* 2023;10:23821205231206220.

continued on next page

continued from previous page

***Daniel P. Walsh** is the Associate Residency Program Director in the Department of Anesthesia, Critical Care and Pain Medicine at Beth Israel Deaconess Medical Center, Boston, MA, the Medical Director of the Intensive Care Unit at Beth Israel Deaconess Hospital in Plymouth, MA, and an Assistant Professor of Anaesthesia at Harvard Medical School, Boston, MA; **Nanaama O'Hene** is a Staff Anesthesiologist at Jersey Shore Anesthesiology Associates, Neptune City, NJ; **Vanessa T. Wong** is a Project Manager in the Department of Anesthesia, Critical Care and Pain Medicine at Beth Israel Deaconess Medical Center, Boston, MA; **Sara E. Neves** is the Residency Program Director in the Department of Anesthesia, Critical Care and Pain Medicine at Beth Israel Deaconess Medical Center, Boston, MA, and an Assistant Professor of Anaesthesia at Harvard Medical School, Boston, MA.*

***Corresponding author:** Daniel P. Walsh, Department of Anesthesia, Critical Care and Pain Medicine, Beth Israel Deaconess Medical Center, 1 Deaconess Road, Rosenberg 660, Boston, MA 02215. Telephone (617) 754-2129, Fax (617) 754-2735*

***Email address:** Daniel P. Walsh: dpwalsh@bidmc.harvard.edu*

***Financial disclosure:** Departmental funding*

Abstract

Background: Trainees and educators spend significant effort in preparing for the anesthesiology In-Training Exam (ITE). We investigated how usage of a question bank (Qbank) among anesthesiology residents correlates to ITE performance.

Methods: After gathering the number of Qbank questions completed by residents and their ITE percentile scores for 3 years, we categorized the number of questions completed by each resident each year as follows: 0–100, 101–500, 501–1000, 1001–2000, >2000. We categorized ITE score percentile for each resident each year as follows: 0–10, 11–25, 26–50, 51–75, 76–90, >90. We then performed a nonparametric regression analysis of (1) ITE score on the different categories of number of questions completed and (2) number of questions completed on the different categories of ITE score.

Results: ITE scores for residents who completed >500 questions were significantly higher than for residents who completed ≤500 questions. Those who completed <100 questions averaged <30th percentile on the ITE. The number of questions completed by residents scoring >10th percentile was significantly higher than residents scoring ≤10th percentile. Residents scoring >90th percentile on the ITE completed significantly more practice questions than other residents, averaging >1000 questions.

Conclusions: ITE scores were significantly different between residents who completed a low number of questions and residents who completed a high number of questions. Further work should investigate other factors that may contribute to exam performance.

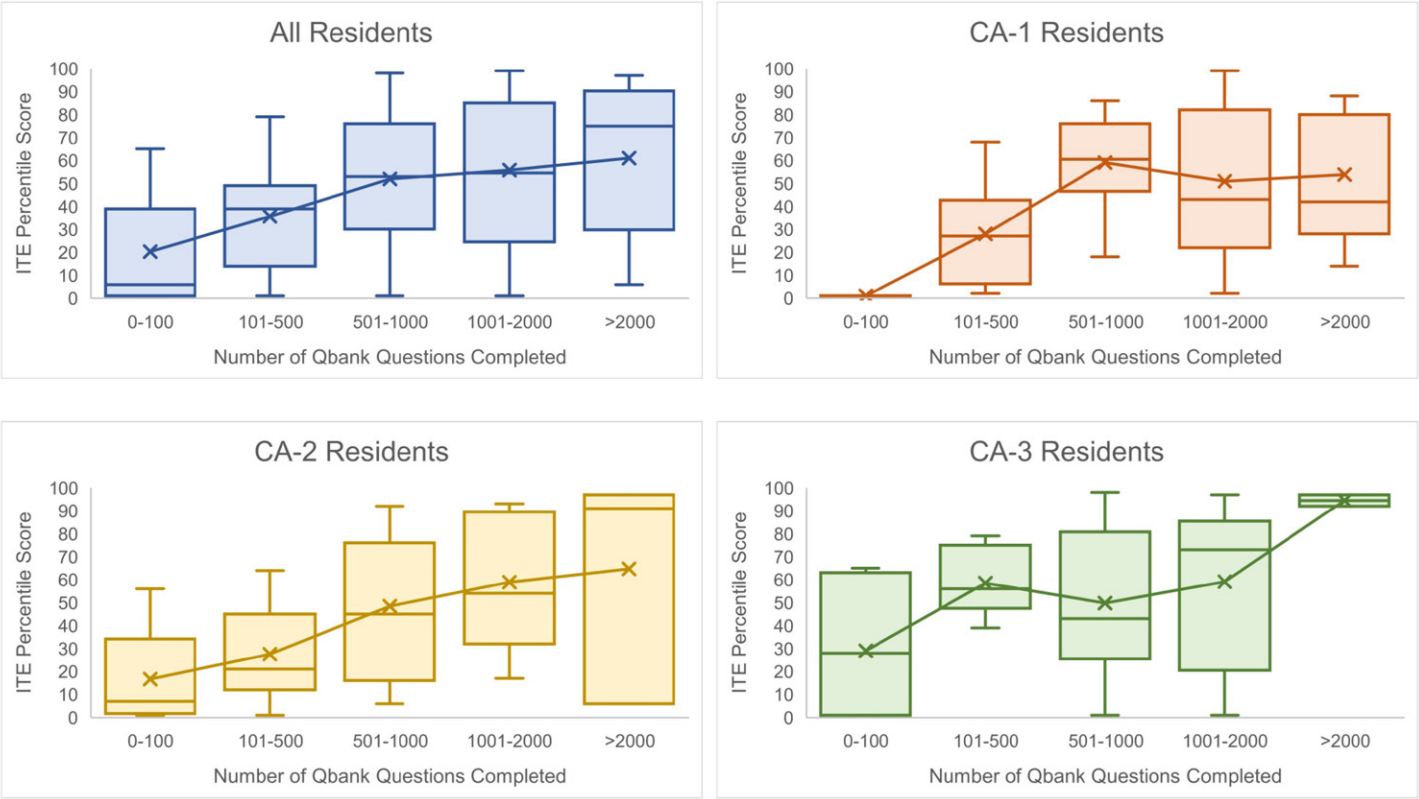
Keywords: question bank, graduate medical education, exam preparation, anesthesiology resident

continued on next page

continued from previous page

Figures

Figure 1. *Exam scores of residents for each category of number of question bank (Qbank) questions completed. Overall, in-training exam (ITE) percentile scores of residents who completed more than 500 Qbank questions were significantly higher than ITE percentile scores of residents who completed fewer than 500 Qbank questions. For CA-1s, ITE percentile scores of residents who completed more than 500 Qbank questions were significantly higher than ITE percentile scores of residents who completed fewer than 500 Qbank questions. They were also significantly higher for residents who completed 101–500 Qbank questions than for residents who completed 0–100 Qbank questions. For CA-2s, ITE percentile scores of residents who completed 501–2000 Qbank questions were significantly higher than ITE percentile scores of residents who completed fewer than 500 Qbank questions. For CA-3s, ITE percentile scores of residents who completed greater than 2000 Qbank questions were significantly higher than ITE percentile scores of residents in all other categories.*

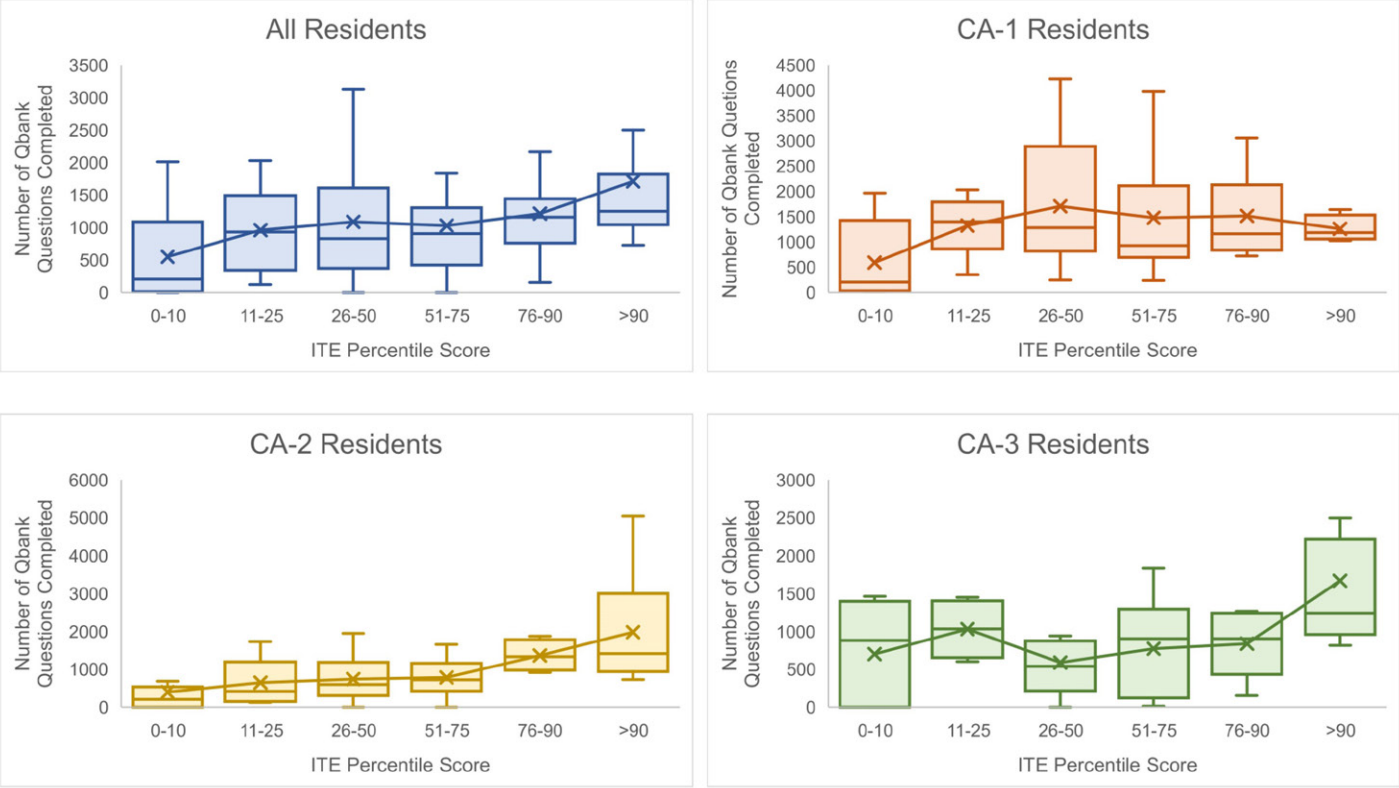


continued on next page

continued from previous page

Figures

Figure 2. *Number of question bank (Qbank) questions completed by residents for each category of exam scores. Overall, the number Qbank questions completed by residents who scored from the 11th to greater than the 90th percentile were significantly higher than the number of Qbank questions completed by residents who scored below the 10th percentile. They were also significantly higher for residents who scored greater than 90th percentile than those who scored between the 11th and 75th percentiles. For CA-1s, the number of Qbank questions completed by residents who scored from the the 11th to greater than the 90th percentile were significantly higher than residents who scored in the 10th percentile or lower. For CA-2s, the number of Qbank questions completed by residents in who scored greater than the 75th percentile were significantly higher than those who scored 75th percentile or lower, but they were not significantly different between those who scored between the 76th and 90th percentiles and those who scored greater than the 90th percentile. For CA-3s, the number of Qbank questions completed by residents who scored greater than the 90th percentile was significantly higher than those who scored in the 10th percentile or less and those who scored between the 26th and 50th percentiles and between the 51st and 75th percentiles.*



continued on next page

continued from previous page

Tables

Table 1. *Nonparametric Series Regression of In-Training Exam (ITE) Percentile Score on Number of Question Bank (Qbank) Questions Completed*

Residents	Nonparametric Series Regression Base Level (Number of Questions)	Comparison (Number of Questions)	Average Effect (95% Confidence Interval)	*p*-value
All (n = 155)	0-100	101-500 (n = 31) vs 0-100 (n = 15)	15.34 (0.762 to 29.926)	.039
		501-1000 (n = 35) vs 0-100 (n = 15)	31.64 (16.381 to 46.895)	<.001[b]
		1001-2000 (n = 58) vs 0-100 (n = 15)	35.48 (20.859 to 50.095)	<.001[b]
		>2000 (n = 16) vs 0-100 (n = 15)	40.60 (21.002 to 60.206)	<.001[b]
	101-500	501-1000 (n = 35) vs 101-500 (n = 31)	16.29 (3.962 to 28.626)	.010[b]
		1001-2000 (n = 58) vs 101-500 (n = 31)	20.13 (8.600 to 31.666)	.001[b]
		>2000 (n = 16) vs 101-500 (n = 31)	25.26 (7.838 to 42.683)	.004[b]
	501-1000	1001-2000 (n = 58) vs 501-1000 (n = 35)	3.84 (-8.536 to 16.214)	.543
		>2000 (n = 16) vs 501-1000 (n = 35)	8.97 (-9.025 to 26.957)	.329
	1001-2000	>2000 (n = 16) vs 1001-2000 (n = 58)	5.13 (-12.325 to 22.580)	.565
CA-1 (n = 54)	0-100	101-500 (n = 8) vs 0-100 (n = 2)	27 (10.959 to 43.041)	.001[a]
		501-1000 (n = 10) vs 0-100 (n = 2)	58 (44.784 to 71.216)	<.001[a]
		1001-2000 (n = 23) vs 0-100 (n = 2)	49.91 (36.517 to 63.309)	<.001[a]
		>2000 (n = 11) vs 0-100 (n = 2)	52.82 (36.863 to 68.773)	<.001[a]
	101-500	501-1000 (n = 10) vs 101-500 (n = 8)	31 (10.216 to 51.784)	.003[a]
		1001-2000 (n = 23) vs 101-500 (n = 8)	22.91 (2.014 to 43.812)	.032[a]
		>2000 (n = 11) vs 101-500 (n = 8)	25.82 (3.194 to 48.443)	.025[a]
	501-1000	1001-2000 (n = 23) vs 501-1000 (n = 10)	-8.09 (-26.905 to 10.731)	.400
		>2000 (n = 11) vs 501-1000 (n = 10)	-5.18 (-25.900 to 15.536)	.624
	1001-2000	>2000 (n = 11) vs 1001-2000 (n = 23)	2.91 (-17.928 to 23.738)	.785
CA-2 (n = 52)	0-100	101-500 (n = 15) vs 0-100 (n = 6)	10.93 (-8.173 to 30.039)	.262
		501-1000 (n = 11) vs 0-100 (n = 6)	31.79 (7.307 to 56.269)	.011[a]
		1001-2000 (n = 17) vs 0-100 (n = 6)	42.27 (20.755 to 63.794)	<.001[a]
		>2000 (n = 3) vs 0-100 (n = 6)	48 (-4.145 to 100.145)	.071
	101-500	501-1000 (n = 11) vs 101-500 (n = 15)	20.85 (0.383 to 41.326)	.046*
		1001-2000 (n = 17) vs 101-500 (n = 15)	31.34 (14.524 to 48.159)	<.001[a]
		>2000 (n = 3) vs 101-500 (n = 15)	37.07 (-13.321 to 87.454)	.149
	501-1000	1001-2000 (n = 17) vs 501-1000 (n = 11)	10.49 (-12.254 to 33.227)	.366
		>2000 (n = 3) vs 501-1000 (n = 11)	16.21 (-36.449 to 68.873)	.546
	1001-2000	>2000 (n = 3) vs 1001-2000 (n = 17)	5.73 (-45.626 to 57.077)	.827

continued on next page

continued from previous page

Tables

Residents	Nonparametric Series Regression Base Level (Number of Questions)	Comparison (Number of Questions)	Average Effect (95% Confidence Interval)	*p*-value
CA-3 (n = 49)	0-100	101-500 (n = 8) vs 0-100 (n = 7)	29.5 (6.891 to 52.109)	.011[a]
		501-1000 (n = 14) vs 0-100 (n = 7)	20.71 (-5.316 to 46.745)	.119
		1001-2000 (n = 18) vs 0-100 (n = 7)	30.11 (4.095 to 56.128)	.023
		>2000 (n = 2) vs 0-100 (n = 7)	65.5 (45.022 to 85.978)	<.001[a]
	101-500	501-1000 (n = 14) vs 101-500 (n = 8)	-8.79 (-28.198 to 10.627)	.375
		1001-2000 (n = 18) vs 101-500 (n = 8)	0.61 (-18.782 to 20.004)	.951
		>2000 (n = 2) vs 101-500 (n = 8)	36 (25.111 to 46.889)	<.001[a]
	501-1000	1001-2000 (n = 18) vs 501-1000 (n = 14)	9.40 (-13.895 to 32.689)	.429
		>2000 (n = 2) vs 501-1000 (n = 14)	44.79 (27.904 to 61.668)	<.001[a]
	1001-2000	>2000 (n = 2) vs 1001-2000 (n = 18)	35.39 (18.529 to 52.249)	<.001[a]

[a]Significant at α = 0.05.

[b]Significant at α = 0.0125.

continued on next page

continued from previous page

Tables

Table 2. *Nonparametric Series Regression of Number of Question Bank (Qbank) Questions Completed on In-Training Exam (ITE) Percentile Score*

Residents	Nonparametric Series Regression Base Level (ITE Percentile Score)	Comparison (ITE Percentile Score)	Average Effect (95% Confidence Interval)	p-value
All (n = 155)	0-10	11-25 (n = 22) vs 0-10 (n = 23)	412.81 (35.866 to 789.759)	.032[a]
		26-50 (n = 42) vs 0-10 (n = 23)	540.13 (133.382 to 946.879)	.009[a]
		51-75 (n = 28) vs 0-10 (n = 23)	479.24 (50.057 to 908.418)	.029[a]
		76-90 (n = 20) vs 0-10 (n = 23)	664.08 (263.928 to 1064.233)	.001[a]
		> 90 (n = 20) vs 0-10 (n = 23)	1164.53 (584.556 to 1744.505)	<.001[a]
	11-25	26-50 (n = 42) vs 11-25 (n = 22)	127.32 (-268.667 to 523.303)	.529
		51-75 (n = 28) vs 11-25 (n = 22)	66.43 (-352.569 to 485.419)	.756
		76-90 (n = 20) vs 11-25 (n = 22)	251.27 (-137.938 to 640.475)	.206
		> 90 (n = 20) vs 11-25 (n = 22)	751.72 (179.241 to 1324.196)	.010[a]
	26-50	51-75 (n = 28) vs 26-50 (n = 42)	-60.89 (-506.888 to 385.102)	.789
		76-90 (n = 20) vs 26-50 (n = 42)	123.95 (-294.186 to 542.086)	.561
		> 90 (n = 20) vs 26-50 (n = 42)	624.4 (31.875 to 1216.925)	.039[a]
	51-75	76-90 (n = 20) vs 51-75 (n = 28)	184.84 (-255.145 to 624.831)	.410
		> 90 (n = 20) vs 51-75 (n = 28)	685.293 (77.150 to 1293.436)	.027[a]
	76-90	> 90 (n = 20) vs 76-90 (n = 20)	500.45 (-87.567 to 1088.467)	.095

continued on next page

continued from previous page

Tables

Residents	Nonparametric Series Regression Base Level (ITE Percentile Score)	Comparison (ITE Percentile Score)	Average Effect (95% Confidence Interval)	p-value
CA-1 (n = 54)	0-10	11-25 (n = 8) vs 0-10 (n = 7)	733.55 (49.712 to 1417.396)	.036[a]
		26-50 (n = 17) vs 0-10 (n = 7)	1118.66 (317.303 to 1920.024)	.006[a]
		51-75 (n = 10) vs 0-10 (n = 7)	891.33 (-42.141 to 1824.798)	.061
		76-90 (n = 8) vs 0-10 (n = 7)	925.18 (124.648 to 1725.709)	.024[a]
		> 90 (n = 4) vs 0-10 (n = 7)	670.43 (54.895 to 1285.962)	.033[a]
	11-25	26-50 (n = 17) vs 11-25 (n = 8)	385.11 (-298.647 to 1068.867)	.270
		51-75 (n = 10) vs 11-25 (n = 8)	157.78 (-676.917 to 992.467)	.711
		76-90 (n = 8) vs 11-25 (n = 8)	191.63 (-491.160 to 874.410)	.582
		> 90 (n = 4) vs 11-25 (n = 8)	-63.13 (-515.039 to 388.789)	.784
	26-50	51-75 (n = 10) vs 26-50 (n = 17)	-227.34 (-1160.743 to 706.072)	.633
		76-90 (n = 8) vs 26-50 (n = 17)	-193.49 (-993.944 to 606.973)	.636
		> 90 (n = 4) vs 26-50 (n = 17)	-448.24 (-1063.674 to 167.204)	.153
	51-75	76-90 (n = 8) vs 51-75 (n = 10)	33.85 (-898.845 to 966.545)	.943
		> 90 (n = 4) vs 51-75 (n = 10)	-220.9 (-1000.613 to 558.813)	.579
	76-90	> 90 (n = 4) vs 76-90 (n = 8)	-254.75 (-869.108 to 359.608)	.416

continued on next page

continued from previous page

Tables

Residents	Nonparametric Series Regression Base Level (ITE Percentile Score)	Comparison (ITE Percentile Score)	Average Effect (95% Confidence Interval)	p-value
CA-2 (n = 52)	0-10	11-25 (n = 10) vs 0-10 (n = 9)	246.42 (-325.331 to 818.176)	.398
		26-50 (n = 13) vs 0-10 (n = 9)	342.22 (-194.092 to 878.537)	.211
		51-75 (n = 8) vs 0-10 (n = 9)	382.10 (-175.279 to 939.474)	.179
		76-90 (n = 4) vs 0-10 (n = 9)	963.97 (397.231 to 1530.713)	.001[a]
		> 90 (n = 8) vs 0-10 (n = 9)	1583.47 (474.298 to 2692.647)	.005*
	11-25	26-50 (n = 13) vs 11-25 (n = 10)	95.8 (-410.466 to 602.066)	.711
		51-75 (n = 8) vs 11-25 (n = 10)	135.68 (-392.851 to 664.201)	.615
		76-90 (n = 4) vs 11-25 (n = 10)	717.55 (179.157 to 1255.943)	.009[a]
		> 90 (n = 8) vs 11-25 (n = 10)	1337.05 (242.089 to 2432.011)	.017[a]
	26-50	51-75 (n = 8) vs 26-50 (n = 13)	39.88 (-450.096 to 529.846)	.873
		76-90 (n = 4) vs 26-50 (n = 13)	621.75 (121.152 to 1122.348)	.015[a]
		> 90 (n = 8) vs 26-50 (n = 13)	1241.25 (164.370 to 2318.13)	.024[a]
	51-75	76-90 (n = 4) vs 51-75 (n = 8)	581.88 (58.775 to 1104.975)	.029[a]
		> 90 (n = 8) vs 51-75 (n = 8)	1201.38 (113.852 to 2288.898)	.030[a]
	76-90	> 90 (n = 8) vs 76-90 (n = 4)	619.5 (-472.852 to 1711.852)	.266

continued on next page

continued from previous page

Tables

Residents	Nonparametric Series Regression Base Level (ITE Percentile Score)	Comparison (ITE Percentile Score)	Average Effect (95% Confidence Interval)	p-value
CA-3 (n = 49)	0-10	11-25 (n = 4) vs 0-10 (n = 7)	329.64 (-270.356 to 929.642)	.282
		26-50 (n = 12) vs 0-10 (n = 7)	-111.36 (-682.944 to 460.230)	.703
		51-75 (n = 10) vs 0-10 (n = 7)	72.44 (-550.555 to 695.441)	.820
		76-90 (n = 8) vs 0-10 (n = 7)	135.89 (-431.792 to 703.578)	.639
		> 90 (n = 8) vs 0-10 (n = 7)	969.39 (47.865 to 1890.921)	.039[a]
	11-25	26-50 (n = 12) vs 11-25 (n = 4)	-441 (-908.551 to 26.551)	.065
		51-75 (n = 10) vs 11-25 (n = 4)	-257.2 (-786.368 to 271.968)	.341
		76-90 (n = 8) vs 11-25 (n = 4)	-193.75 (-656.523 to 269.023)	.412
		> 90 (n = 8) vs 11-25 (n = 4)	639.75 (-221,125 to 1500.625)	.145
	26-50	51-75 (n = 10) vs 26-50 (n = 12)	183.8 (-312.920 to 680.520)	.468
		76-90 (n = 8) vs 26-50 (n = 12)	247.25 (-178.039 to 672.539)	.255
		> 90 (n = 8) vs 26-50 (n = 12)	1080.75 (239.431 to 1922.069)	.012[a]
	51-75	76-90 (n = 8) vs 51-75 (n = 10)	63.45 (-428.775 to 555.675)	.801
		> 90 (n = 8) vs 51-75 (n = 10)	896.95 (19.891 to 1774.009)	.045[a]
	76-90	> 90 (n = 8) vs 76-90 (n = 8)	833.5 (-5.174 to 1672.173)	.051

[a]Significant at $\alpha = 0.05$.