Does Practice Make Perfect? Variations in OITE Performance Based on Practice Question Exposure and Proficiency

Spencer S. Schulte MD[1], Scott L. Oplinger MPA[2], Benjamin R. Childs MD[3], John C. Dunn MD[3]

1) Madigan Army Medical Center, Tacoma, WA
2) Uniformed Services University, Bethesda, MD
3)William Beaumont Army Medical Center, Ft. Bliss, TX

Background: The Orthopedic In-Training Examination (OITE) is used to assess the progression of knowledge during residency. Research has established that performance on the United States Medical Licensing Exam (USMLE) Step 1 is correlated with OITE performance, yet research into specific OITE preparation strategies is scarce.

Objective: To prospectively evaluate whether volume and performance on Orthobullets practice questions relate to OITE performance.

Design: Number of practice questions and performance of 23 orthopedic residents at a single institution were from July to November. Case log data, USMLE Step 1 and 2 scores, and prior OITE scores were also collected. The data was evaluated using univariate and multivariate regression analysis.

Main Outcome Measurements: OITE raw score.

Results: Multivariate regression found the following variables to be related to raw OITE score: year in training (PGY) (b = 7.112, p < 0.001), percent of practice questions correct (b = 0.990, p < 0.001) and Step 1 score (b = 0.624, p = 0.002). Multivariate regression for PGY OITE percentile found only mean Step 1 and 2 combined scores to be predictive (b = 1.271, p = 0.001).

Conclusion: This prospective study showed that the number of practice questions performed in the academic year prior to the annual OITE is not correlated with higher OITE performance, whereas the percentage of practice questions correct is. PGY was most strongly correlated with OITE performance.

Level of Evidence: Level II Prospective Cohort

Key Words: OITE, orthopaedic in-training examination, practice question, question bank

(J Ortho Business 2023; Volume 3, Issue:3, Pages 8-13)

INTRODUCTION

The Orthopedic In-Training Examination (OITE) has been employed since 1963 in residency programs throughout the nation to gauge resident progression throughout residency.[1-3] USMLE Step 1 score have been found to predict performance during residency, including on the OITE.[4-8] A broad array of educational approaches are used to prepare residents and allow them to perform well on the OITE. This educational framework consists of a combination of the following: directed primary literature reading plans, focused educational conferences, orthopedic case exposure, and previous OITE question review.[4,9-12] While significant research has gone into identifying best practices in training orthopaedic residents,[9,10,12-14] additional research is needed. Additionally, residency program directors (PDs) and residents disagree on which educational methods are most effective.[11] Despite differing views, both residents and PDs value the use of past OITE questions above other preparation methods.[9,11,15]

The primary purpose of this study is to prospectively evaluate whether volume and performance on practice questions has a relationship with performance on the OITE. We hypothesize that residents who perform more practice questions and get a higher percentage of practice questions correct will perform better on the OITE.

METHODS

The institutional review board deemed this study to be exempt. The 23 residents at a single orthopaedic surgery residency program utilized the same proprietary Orthobullets question bank and participated in the same weekly didactic education program. Data regarding the number of questions done by residents each month were gathered prospectively by the chief resident through Orthobullets. Variables collected retrospectively from the program director included USMLE Step 1 and 2 scores, as well as prior OITE scores and data regarding resident research activity. Step 1 and 2 scores were

available for all residents except one, who instead took the Comprehensive Osteopathic Medical Licensing Examination (COMLEX). Case log data was also obtained prospectively from the program director. The primary outcome variables were raw OITE score and PGY OITE score (OITE score standardized for a resident's year in residency). All continuous variables were compared using univariate linear regression, after which all significant variables were evaluated with multivariate linear regression. Continuous variables were evaluated using the Mann-Whitney U test. Significance was taken as p-values less than 0.05. All analysis was performed in SPSS version 27.0.1.0 by IBM Corp.

RESULTS

The mean number of practice questions completed per resident per month was 934±471 for a mean total of 2722±1299 completed questions per resident between July 1 and the OITE. The month with the highest average practice questions completed per resident was October (586±311), the last full month before the OITE. Residents on average chose the correct answer 80.1±9.1% of the time. Residents logged an average of 194±104 cases from July 1 to the date of the OITE. The average Step 1 score was 244±11.7. The mean of each resident's Step 1 and Step 2 scores was 246±9.9 (Table 1). Among the participants with a Step 1 score below 240 (8 out of the 22 residents), 8 out of 8 scored above the 32nd percentile within their PGY on the 2020 OITE, and 6 out of 8 scored above the 50th percentile.

The average raw score on the 2020 OITE for the residents in this study was 172±20 for a mean class percentile rank of 70th percentile. The mean increase in raw score from 2019 to 2021 was 6.2±6.7 points (Table 1).

In univariate regression, PGY (b = 10.027, p < 0.001), percent of practice questions correct (b = 1.405, p = 0.002), Step 1 score (0.904, p = 0.012), number of cases logged (0.135, p < 0.001), and average number of questions logged per month (0.021, 0.016) were all significantly correlated with increased 2021 OITE raw score. However, after multivariate regression to control for collinearity only PGY (7.112, p < 0.001), percent of practice questions correct (0.990, p < 0.001), and Step 1 score (0.624, 0.002) were correlated with

Table 1: Descriptive and Outcome Statistics

Practice questions performed	N	Min	Max	Mean	Std. Dev
July	23	0	971	392	309
August	23	0	1378	575	368
September	23	137	2733	709	589
October	23	98	3015	1210	805
November	23	201	2392	847	536
Per month	23	156	1507	934	471
Total, July to OITE	23	780	7535	3735	1863
Percent of practice questions correct	23	59	98	80.1	9.1
OITE Scores					
2020, raw	23	135	207	172	20
Change in raw score, 2019 to 2020	18	-10	23	6.2	6.7
2020 percentile rank	23	34	99	70	20
Surgical cases logged, July to OITE	23	33	401	194	104
USMLE scores					
Step 1	23	215	263	244	11.7
Mean of Step 1 and 2	23	219	264	246	9.9

*OITE: Orthopedic In-Training Examination, USMLE: United States Medical Licensing Examination

improved 2021 OITE raw score (Table 2). The regression captured 84% of score variation when plotted against the actual scores, as seen in Figure 1 (R^2= 0.84).

In univariate regression, Step 1 score (β=0.830, p=0.027), mean Step 1 and 2 scores (β=1.271, p=0.001), and percent of practice questions correct (β=0.958, p=0.045) were each correlated with higher 2021 PGY percentile score. However, multivariate regression revealed that only the mean of Step 1 and 2 scores (β=1.271, p=0.001) was correlated with a higher 2021 PGY percentile score (Table 3).

DISCUSSION

Our hypothesis that OITE performance would be correlated with performance on practice questions was supported. Our hypothesis that OITE performance would be influenced by the number of practice questions performed was not supported. For OITE raw score, multivariate regression

Table 2: 2020 OITE raw score linear regression.

	Single variable			Multivariate		
	b	R sq	p	b	R sq	p
PGY	10.027	.520	<0.001	7.112		<0.001
Percent correct	1.405	.398	0.002	0.990	0.838	<0.001
USMLE Step 1 score	0.904	.274	0.012	0.624		0.002
Cases since July	0.135	.498	<0.001	-		
Questions per month	0.021	.247	0.016	-		
Total questions	0.005	.126	0.096	-		
Papers published	2.141	.085	0.177	-		
Questions - November	-0.010	0.015	0.574	-		
Questions - October	0.018	0.081	0.188	-		

*PGY: Post Graduate Year, USMLE: United States Medical Licensing Examination, OITE: Orthopedic In-Training Examination

Figure 1: Scatter plot of regression calculated OITE raw score (regression OITE score = -79.9 + (7.112) * PGY + (0.99) * Percent of practice questions correct + (0.624) * Step 1 score) vs actual scores (R^2= 0.84)

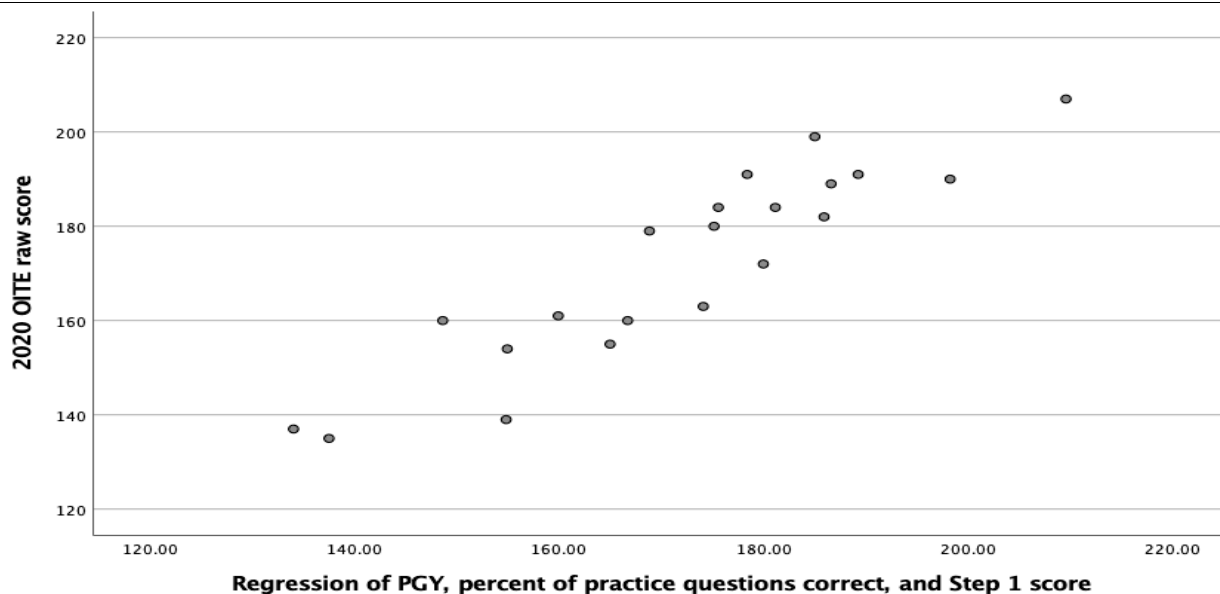Table 3: 2020 OITE Post Graduate Year (PGY) percentile linear regression.

	Single variable			Multivariate		
	b	R sq	p	b	R sq	p
USMLE Step 1 and 2 mean	1.271	0.408	0.001	1.271	0.408	0.001
Percent correct	0.958	0.186	0.045	-		
USMLE Step 1 score	0.830	0.222	0.027	-		
Papers published	-1.119	0.022	0.469	-		
Total questions	-0.002	0.018	0.542	-		
Cases since July	0.023	0.014	0.596	-		
Questions - November	-0.006	0.005	0.739	-		
Questions per month	0.002	0.001	0.864	-		
Questions - October	0.002	0.001	0.869	-		

*PGY: Post Graduate Year, USMLE: United States Medical Licensing Examination, OITE: Orthopedic In-Training Examination

analysis identified PGY as the primary factor, while the percent of practice questions correct and Step 1 score were also positively correlated. For OITE PGY score, the multivariate model identified only a single related variable: the mean of an individual's Step 1 and 2 scores.

Our study found a correlation between OITE practice question performance (percent of questions correct) and OITE performance. This is valuable information to PDs because it is data that can be tracked throughout the year. Low OITE scores are used to identify academically at-risk residents; however, OITE scores are only available once annually. Our results suggest that low practice question performance relative to one's peers may also be used to identify potentially at-risk residents. Unfortunately, our data does not indicate a specific target performance on practice questions for a given level of training. Nevertheless, the relationship between practice question performance and OITE performance further supports current research that has found the Orthobullets question bank to be a useful OITE-preparation resource.[16,17]

Prior research has correlated PGY with OITE raw score performance.[5,16,18-21] This fact, in conjunction with the research conducted by Karlen et al. (adding 3 months of orthopedic training to the intern year increased OITE scores) supports the assumption that more time in training equates to better performance on the OITE.[20] While current research suggests that work hour numbers are not correlated to OITE scores,[22,23] an overall increase in OITE performance has been correlated with overall time in training (PGY). Our research further supports this conclusion.

A problem with asking whether the number of questions correlates with OITE performance is that a PGY1 can perform the same number of questions as a PGY5, yet we would still expect the PGY5 to have a higher raw score than the PGY1. If the number of questions performed were correlated with PGY OITE percentile we would expect a PGY5 who performed more practice questions to perform better than another PGY5 who performed fewer practice questions. In both univariate and multivariate analysis this was not the case. It is likely that a single-year study with a single cohort of residents does not have adequate power to detect a relationship between the number of questions performed and OITE performance. We recommend that a multi-year study be performed to better define the relationship between the number of practice questions performed and OITE performance.

Another notable result of this study was the lack of correlation found between the number of cases logged and OITE performance. Though case numbers showed a correlation with OITE performance in univariate analysis, in multivariate analysis the relationship disappeared. Current research does not agree on the relationship between resident workload and resident performance. Bohm et al. suggest there is no relationship between work hours and OITE performance,[24] whereas Laporte et al. found a positive relationship between those reporting the highest number of cases and those scoring the highest on the OITE.[11] Resident recall and discipline are involved with reporting both work hours and caseloads. It has previously been established that there is variability in case log practices among orthopaedic residents.[25] We suggest that practice variability could be an inherent limitation in attempting to establish relationships with these variables.

One strength of this study is its prospective evaluation of practice question data. The only data obtained retrospectively were the residents' USMLE Step 1 and 2 scores, 2019 OITE scores, and number of papers published. Prospectively gathering data limits bias. Though efforts were made during the study period to maintain a low profile, participants of the study were not blinded, and it is possible that participants exhibited the Hawthorne effect. The introduction of this bias may have obscured the relationship between practice question performance and OITE performance.

A second strength is the heterogeneity of USMLE Step 1 scores of the residents involved with the study. Residents in this study had scores above and below the 240 raw score threshold previously used by many programs to screen applicants. However, all study participants with a Step 1 score below 240 scored above the 32nd percentile within their PGY on the 2020 OITE, and 6 out of 8 scored above the 50th percentile. In a prior survey, it was determined that the 32nd percentile was the mean threshold at which PDs consider

their residents at risk.[26] Though our study did reinforce the relationship between Step 1 scores and OITE performance established in the literature,[4,5,7,14,27] none of our residents who scored below 240 achieved OITE scores indicating they were at risk of failure. With the USMLE Step 1 transition to pass/fail, residency directors will need to rely on other criteria when assessing applicants. Additional research focused on the correlation between USMLE Step 2 scores and OITE performance is warranted.

Potential confounders which were not directly assessed include the use of alternate review materials, time available to study, and priority of preparation. An adequate volume of knowledge is essential to both practice question success and OITE performance. While this research demonstrates that practice question correctness is positively correlated with OITE performance, it does not address how to improve performance on practice question banks.

Our study is limited because we focused on a single institution. The generalizability of the findings at our program may not be appropriate to other programs. This study was performed at an orthopaedic surgery training program with an established weekly academics curriculum where residents regularly score at or above the 90th percentile among programs nationwide. It is possible that our program's curriculum provides a ceiling effect and makes discrimination among our test-takers difficult. At programs with a less structured OITE preparation curriculum, discretionary study materials (e.g. practice questions) may have a more profound effect on OITE performance.

Additionally, our study is limited by a short data collection window. Given that question metrics were only collected from July through November, our study captures final preparation but may miss initial preparation conducted by residents who begin their review earlier. Additional years of data collection, with collection occurring throughout the year, may show trends that our single-year prospective study was unable to detect. For example, it is possible that relationships exist between the number of questions performed in prior years and current year OITE scores. Additional multi-year studies are needed.

CONCLUSION

In conclusion, our data supports PGY, percent of OITE practice questions answered correctly, and USMLE Step I score as correlated with increased OITE performance. This model provides an additional tool to be employed by PDs in identifying and developing remedial training within their programs.

REFERENCES

1. Mankin HJ. The orthopaedic in-training examination (OITE). Clin Orthop Relat Res. 1971 Mar-Apr;75:108-16. PMID: 5554614.
2. Miyamoto RG Jr, Klein GR, Walsh M, Zuckerman JD. Orthopedic surgery residents' study habits and performance on the orthopedic in-training examination. Am J Orthop (Belle Mead NJ). 2007 Dec;36(12):E185-8. PMID: 18264562.
3. Risner B, Nyland J, Crawford CH 3rd, Roberts CS, Johnson JR. Orthopaedic in-training examination performance: a nine-year review of a residency program database. South Med J. 2008 Aug;101(8):791-6. PMID: 18622350.
4. Crawford CH 3rd, Nyland J, Roberts CS, Johnson JR. Relationship among United States Medical Licensing Step I, orthopedic in-training, subjective clinical performance evaluations, and american board of orthopedic surgery examination scores: a 12-year review of an orthopedic surgery residency program. J Surg Educ. 2010 Mar-Apr;67(2):71-8. PMID: 20656602.
5. Dougherty PJ, Walter N, Schilling P, Najibi S, Herkowitz H. Do scores of the USMLE Step 1 and OITE correlate with the ABOS Part I certifying examination?: a multicenter study. Clin Orthop Relat Res. 2010 Oct;468(10):2797-802. PMID: 20352386
6. Herndon JH, Allan BJ, Dyer G, Jawa A, Zurakowski D. Predictors of success on the American Board of Orthopaedic Surgery examination. Clin Orthop Relat Res. 2009 Sep;467(9):2436-45. PMID: 19557490
7. Klein GR, Austin MS, Randolph S, Sharkey PF, Hilibrand AS. Passing the Boards: can USMLE and Orthopaedic in-Training Examination scores predict passage of the ABOS Part-I examination? J Bone Joint Surg Am. 2004 May;86(5):1092-5. PMID: 15118058.
8. Swanson D, Marsh JL, Hurwitz S, DeRosa GP, Holtzman K, Bucak SD, Baker A, Morrison C. Utility of AAOS OITE scores in predicting ABOS Part I outcomes: AAOS exhibit selection. J Bone Joint Surg Am. 2013 Jun 19;95(12):e84. PMID: 23783215.
9. Raman T, Alrabaa RG, Sood A, Maloof P, Benevenia J, Berberian W. Does Residency Selection Criteria Predict Performance in Orthopaedic Surgery Residency? Clin Orthop Relat Res. 2016 Apr;474(4):908-14. PMID: 25940336.
10. Dyrstad BW, Pope D, Milbrandt JC, Beck RT, Weinhoeft AL, Idusuyi OB. Predictive measures of a resident's performance on written Orthopaedic Board scores. Iowa Orthop J. 2011;31:238-43. PMID: 22096449
11. Camp CL, Degen RM, Hanssen AD, Dines JS, Karam MD, Turner NS. Residents and Program Director Perspectives Often Differ on Optimal Preparation Strategies and the Value of the Orthopedic In-Training Examination. J Surg Educ. 2018 Jan-Feb;75(1):164-170. PMID: 28673803.
12. Franklin CC, Bosch PP, Grudziak JS, Dede O, Ramirez RN, Mendelson SA, Ward WT, Brooks M, Kenkre T, Lubahn JD, Deeney VF, Roach JW. Does a Weekly Didactic Conference Improve Resident Performance on the Pediatric Domain of the Orthopaedic In-Training Examination? J Pediatr Orthop. 2017 Mar;37(2):149-153. PMID: 26866645.
13. LaPorte DM, Marker DR, Seyler TM, Mont MA, Frassica FJ. Educational resources for the Orthopedic In-Training Examination. J Surg Educ. 2010 May-Jun;67(3):135-8. PMID: 20630421.
14. Taylor BC, Fowler TT, Dimitris C. Achieving educational excellence: a strategic initiative to enhance orthopedic resident academic performance. J Surg Educ. 2011 May-Jun;68(3):162-6. PMID: 21481797.
15. Janney CF, Kunzler D, Safavi PS, Panchbhavi V. Orthopaedic In-Training Examination Resources and Residency Training for the Foot and Ankle Domain. Foot Ankle Spec. 2019 Apr;12(2):146-152. PMID: 29707970.
16. Boody B, Johnson P, Pugely A, Miller D, Geller J, Payne W, Boegener J, Schafer M, Beal M. Validation of a Web-Based Curriculum for Resident Education in Orthopedic Surgery. J Surg Educ. 2016 Nov-Dec;73(6):1060-1065. PMID: 27453538.

17. Krueger CA, Shakir I, Fuller BC. Prevalence of answers to orthopaedic in-training examination questions in 3 commonly used orthopedic review sources. Orthopedics. 2012 Sep;35(9):e1420-6. PMID: 22955412.

18. Froelich J, Milbrandt JC, Allan DG. Impact of the 80-hour workweek on surgical exposure and national in-training examination scores in an orthopedic residency program. J Surg Educ. 2009 Mar-Apr;66(2):85-8. PMID: 19486871.

19. Karlen AI, Solberg EJ, Quanbeck DS, Van Heest AE. Orthopaedic Surgery Residency Rotations and Correlation With Orthopaedic In-Training Examination Performance. J Surg Educ. 2018 Sep-Oct;75(5):1325-1328. PMID: 29449163.

20. Kragh JF Jr, Robbins J, Ward JA. Faculty turnover and resident in-training examination and board scores. Am J Orthop (Belle Mead NJ). 2007 Dec;36(12):E180-4. PMID: 18264561.

21. Strauss EJ, Markus DH, Kingery MT, Zuckerman J, Egol KA. Orthopaedic Resident Burnout Is Associated with Poor In-Training Examination Performance. J Bone Joint Surg Am. 2019 Oct 2;101(19):e102. PMID: 31577687.

22. Bohm KC, Hill BW, Braman JP, Ly TV, Van Heest AE. Orthopedic Residency: Are Duty Hours Predictive of Performance? J Surg Educ. 2016 Mar-Apr;73(2):281-5. PMID: 26774934.

23. Salazar D, Schiff A, Mitchell E, Hopkinson W. Variability in Accreditation Council for Graduate Medical Education Resident Case Log System practices among orthopaedic surgery residents. J Bone Joint Surg Am. 2014 Feb 5;96(3):e22. PMID: 24500594.

24. Camp CL, Degen RM, Turner NS, Hanssen AD, Karam MD, Dines JS. Orthopedic In-Training Examination: A Performance Review Based on Program- and Resident-Specific Characteristics. J Surg Educ. 2017 Jul-Aug;74(4):754-761. PMID: 28343951.

25. Carmichael KD, Westmoreland JB, Thomas JA, Patterson RM. Relation of residency selection factors to subsequent orthopaedic in-training examination performance. South Med J. 2005 May;98(5):528-32. PMID: 15954509.

26. Kreitz T, Verma S, Adan A, Verma K. Factors Predictive of Orthopaedic In-training Examination Performance and Research Productivity Among Orthopaedic Residents. J Am Acad Orthop Surg. 2019 Mar 15;27(6):e286-e292. PMID: 30252788.

27. Atesok KI, Hurwitz SR, Egol KA, Ahn J, Owens BD, Crosby LA, Pellegrini VD Jr. Perspective: Integrating research into surgical residency education: lessons learned from orthopaedic surgery. Acad Med. 2012 May;87(5):592-7. doi: 10.1097/ACM.0b013e31824d2b57. PMID: 22450176.