Orthopedic in-training examination question metrics and resident test performance

John Mason DePasse,[1] Jack Haglin,[2] Adam E.M. Eltorai,[2] Mary K. Mulcahey,[3] Craig P. Eberson,[4] Alan H. Daniels,[5]

[1]Department of Orthopaedics, Warren Alpert Medical School, Brown University, Providence, RI; [2]Brown University, Providence, RI; [3]Department of Orthopaedics, Drexel University College of Medicine, Philadelphia, PA; [4]Division of Pediatrics, Department of Orthopaedics, Warren Alpert Medical School, Brown University, Providence, RI; [5]Department of Orthopaedics, Division of Spine Surgery, Warren Alpert Medical School, Brown University, Providence, RI, USA

Abstract

First administered in November 1963, the orthopedic in-training examination (OITE) is now distributed to more than 4000 residents in over 20 countries and has become important for evaluation of resident fund of knowledge. Several studies have assessed the effect of didactic programs on resident performance, but only recently has it become possible to assess detailed test-taking metrics such as time spent per question. Here, we report the first assessment of resident OITE performance utilizing this full electronic dataset from two large academic institutions.

Full 2015 OITE score reports for all orthopedic surgery residents at two institutions were anonymized and compiled. For every question answered by each resident, the resident year, question content or domain, question result (correct or incorrect), and answer speed were recorded. Data were then analyzed to determine whether resident year, result, or domain affected answer speed and whether performance in each subspecialty domain varied based on resident year in training.

Data was available for 46 residents and 12,650 questions. Mean answer speed for questions answered correctly, 54.0±48.1 s, was significantly faster than for questions answered incorrectly, 72.2±61.2 s (P<0.00001). When considering both correct and incorrect answers, PGY-1s were slower than all other years (P<0.02). Residents spent a mean of nearly 80 seconds on foot and ankle and shoulder and

elbow questions, compared to only 40 seconds on basic science questions (P<0.05).

In education, faster answer speed for questions is often considered a sign of mastery of the material and more confidence in the answer. Though faster answer speed was strongly associated with correct answers, this study demonstrates that answer speed is not reliably associated with resident year. While answer speed varies between domains, it is likely that the majority of this variation is due to question type as opposed to confidence. Nevertheless, it is possible that in domains with more tiered experience such as shoulder, answer speed correlates strongly with resident year and percentage correct.

Introduction

The orthopedic in-training examination (OITE) was first administered to residents in November 1963, representing the first in-training examination for any specialty.[1] Since then, the OITE has been validated as a tool for assessing resident knowledge and has been shown to correlate with resident performance.[2,3] OITE scores have also been shown to correlate with passing the American Board of Orthopaedic Surgery (ABOS) Part 1 certifying examination.[4,5] Thus, the OITE continues to be an important metric for evaluation of resident fund of knowledge and resident development.

Several studies have investigated resident performance on the OITE. Recent studies have found that both a weekly subspecialty didactic conferences and a weekly literature reading program improved resident OITE performance.[6,7] Similarly, a didactic curriculum designed to improve OITE scores has been shown to significantly increase scores.[8] In order to assist with the development of these programs, multiple studies have analyzed the ten domains of the OITE, reviewing question categories and sources referenced for individual domains.[9-11]

Although some studies have reported overall resident performance, no studies have assessed resident test-taking in detail. In November 2009, the OITE was administered electronically for the first time to 4300 orthopaedic residents in 20 countries.[1] As the electronic test has improved, data is now easily captured on resident test-taking ability. In this investigation, we assessed resident test-taking question metrics through evaluation of the time each resident spent on every question for all residents at two large academic institutions.

Materials and Methods

Full 2015 OITE score reports for all orthopedic surgery residents at two institutions were anonymized and compiled. For every question answered by each resident, the resident year, question content or domain, question result (correct or incorrect), and answer speed were recorded.

Data were then analyzed to determine whether resident year, result, or domain affected answer speed. Answer speed was also assessed for correct and incorrect answers by resident year. Furthermore, question result and answer speed were also analyzed by resident year for each of the ten domains to determine whether performance in each subspecialty domain varied based on resident year in training. All analyses were performed for pooled data from two institutions. Statistical analyses were performed using Microsoft Excel (Microsoft Corporation, Redmond, WA, USA), with a value for statistical significance of P<0.05.

Results

2015 OITE score reports were obtained for 46 residents: six postgraduate year (PGY)-1s and ten each of PGY-2s, PGY-3s, PGY-4s, and PGY-5 s. The OITE consists of 275 questions, and thus data was available for 12,650 questions.

Mean answer speed for questions answered correctly was 54.0±48.1 seconds, which was significantly faster than answer speed for incorrect answers, 72.2±61.2 s (P<0.00001). Mean answer speed differed between resident years (P=0.0006); PGY-1s

averaged 66.0±52.3 s, PGY-2s averaged 60.1±48.8 s, PGY-3s averaged 62.1±54.0 s, PGY-4s averaged 58.9±53.4 s, and PGY-5s averaged 61.1±61.9 s. Post-hoc Tukey analysis revealed that PGY-1s were slower than all other years (P<0.02), and PGY-3s were slower than PGY-4s (P=0.03). No other comparisons were significant (P>0.15). These answer speeds did not strongly correlate with resident year (r=-0.65) or the overall percentage correct of each resident year (r=-0.73) (Figure 1).

Mean answer speed also differed between resident years for both correct (P=0.008) and incorrect answers (P=0.019). For questions answered correctly, post-hoc analysis demonstrated that PGY-1s were slower than all other classes (P<0.004) except the PGY-3s (P=0.17), who were in turn slower than PGY-2s and PGY-5s (P<0.05). Answer speeds for correct answers did not correlate with resident year (r=-0.69) or percentage correct of each resident year (r=-0.66) (Figure 1).

For questions answered incorrectly, post-hoc analysis revealed that PGY-5s were slower than PGY-2s and PGY-4s (P<0.02), but no other comparisons were significant (P>0.14). Answer speeds for incorrect answers did not correlate with resident year (r=0.51) or percentage correct for each resident year (r=0.31).

Answer speed varied by domain from nearly 80 s for shoulder and foot and ankle questions to 40 s for basic science questions (P<0.0001). Answer speed did not correlate with percentage correct in each domain (r=-0.50) (Table 1).

Answer speed was also analyzed by resident year within each of the ten domains, and significant differences were noted only within pediatrics (P=0.02) and shoulder and elbow (P=0.01). In pediatrics, answer speed and percentage correct of each resident year did not correlate (r=-0.51), but in shoulder, there was a strong correlation (r=-0.84) (Table 2). Within the eight other domains, there were no significant differences in answer speed between years (P>0.24).

Discussion

As of 2015, electronic measurements of the time that each resident spends on each question are available and provide valuable data for assessing resident performance on the OITE. This data allows for the first analysis of OITE answer speed, which has previously been impossible to study. Though OITE answer speed has not previously been examined, it is not new in education; when attending and other teaching staff ask questions, they assess for the ease and alacrity with which medical students and residents are able to respond. Answer speed is considered a measure of confidence; a quick, direct answer indicates a level of mastery in fund of knowledge.

In this investigation we assessed answer speed for 12,650 OITE questions from two institutions. A faster answer was very strongly associated with a correct answer for all test-takers at both institutions. Though occasionally residents may answer quickly even when they do not understand the question, the data suggest that the majority of quick responses are correct ones. We hypothesized that senior residents would answer more quickly, as they may
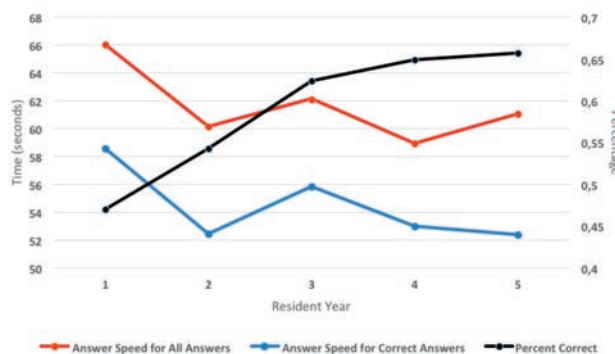

Figure 1. Mean answer speed and percent correct by resident year.

Table 1. Mean answer speed and percentage correct by orthopedic domain.

	Answer speed	Percentage correct
Shoulder and elbow	79.2±67.5	54.4
Foot and ankle	78.5±50.8	45.1
Hand	65.1±49.6	45.1
Sports	63.7±50.4	63.4
Joints	62.7±52.3	53.1
Trauma	62.6±59.0	63.0
Spine	57.5±67.0	72.2
Oncology	55.9±44.0	62.3
Pediatrics	55.7±49.0	71.9
Basic science	41.0±37.5	57.9
Correlation	r=-0.50	

Table 2. Mean answer speed and percentage correct by resident year for pediatrics and shoulder and elbow domains.

	Pediatrics		Shoulder and elbow	
	Answer speed	Percentage correct	Answer speed	Percentage correct
PGY-1	63.6±43.2	61.0	95.8±111.6	39.9
PGY-2	52.2±42.4	66.0	78.1±53.4	46.1
PGY-3	57.7±56.3	74.6	81.6±61.4	55.7
PGY-4	51.0±42.3	76.6	76.3±60.3	61.3
PGY-5	57.0±55.7	76.9	71.1±55.2	63.5
Correlation	r =-0.51	r =-0.84		

PGY, postgraduate year.

have a larger fund of knowledge and thus be more confident in their answers. Our data do not support this hypothesis. Though interns were slower than other classes, the chief residents were not faster than the PGY-2s. When assessing answer speed only for questions that were answered correctly, there is still no correlation between resident year and answer speed, though the interns were slower than all other classes except the PGY-3s. It is interesting to note that when assessing answer speed for questions that were answered incorrectly, the PGY-5s were the slowest. The PGY5s were the fastest for questions answered correctly, though this relationship reached significance only when compared to the PGY-1s and PGY-3s. This finding suggests that the most senior residents may be spending more time considering difficult questions than more junior residents.

Residents spent almost 80 seconds on shoulder and foot and ankle questions, which contrasts sharply to the 40 seconds they spent on basic science questions. As we hypothesized, answer speed was fastest for basic science questions, which are often knowledge-level fact questions. In contrast, shoulder and foot and ankle questions took residents longest, which is likely due to those questions frequently necessitating review of advanced imaging. However, spine and tumor also both frequently require imaging review, so it is unclear why shoulder and foot and ankle demanded more time.

Within shoulder, answer speed differed between the resident years and correlated with the percentage of correct answers. With each additional year in residency, residents answered faster and performed better. This contrasts with pediatrics, where PGY-2s and PGY-4s answered most quickly, and answer speed did not correlate with percentage correct. One possible contributing fac-tor could be that for both institutions, shoulder and elbow exposure to junior residents is limited. Thus, seniors have a more significant advantage in shoulder.

This study has several potential limitations. Only 2 institutions were assessed, and although they are relatively large programs, data from other programs, including programs from a variety of geographic locations may be beneficial. Additionally, only 2015 data was assessed. Future investigations should include a larger data set.

Conclusions

This study revealed that faster answer speed was strongly associated with correct answers, though answer speed is not reliably associated with resident year. While answer speed varies between domains, it is likely that the majority of this variation is due to question type as opposed to confidence. Nevertheless, it is possible that in domains with more tiered experience, such as in shoulder, answer speed correlates strongly with resident year and percentage correct. Further research with additional test offerings would help confirm these findings.

References

1. Marsh JL, Hruska L, Mevis H. An electronic orthopaedic in-training examination. J Am Acad Orthop Surg 2010;18:589-96.
2. Buckwalter JA, Schumacher R, Albright JP, Cooper RR. The validity of orthopaedic in-training examination scores. J Bone Joint Surg Am 1981;63:1001-6.
3. Spitzer AB, Gage MJ, Looze CA, et al. Factors associated with successful performance in an orthopaedic surgery residency. J Bone Joint Surg Am 2009;91:2750-5.
4. Dougherty PJ, Walter N, Schilling P, et al. Do scores of the USMLE Step 1 and OITE correlate with the ABOS Part I certifying examination?: a multicenter study. Clin Orthop 2010;468:2797-802.
5. Herndon JH, Allan BJ, Dyer G, et al. Predictors of success on the American Board of Orthopaedic Surgery examination. Clin Orthop 2009;467:2436-45.
6. Franklin CC, Bosch PP, Grudziak JS, et al. Does a weekly didactic conference improve resident performance on the pediatric domain of the orthopaedic in-training examination? J Pediatr Orthop 2016; 37:149-53.
7. Weglein DG, Gugala Z, Simpson S, Lindsey RW. Impact of a weekly reading program on orthopedic surgery residents' in-training examination. Orthopedics 2015;38:e387-e93.
8. Klena JC, Graham JH, Lutton JS, et al. Use of an integrated, anatomic-based, orthopaedic resident education curriculum: a 5-year retrospective review of its impact on orthopaedic in-training examination scores. J Grad Med Educ 2012;4:250-3.
9. Sheibani-Rad S, Arnoczky SP, Walter NE. Analysis of the basic science section of the orthopaedic in-training examination. Orthopedics 2012;35: e1251-55.
10. Mesfin A, Farjoodi P, Tuakli-Wosornu YA, et al. An analysis of the orthopaedic in-training examination rehabilitation section. J Surg Educ 2012;69:286-91.
11. Osbahr DC, Cross MB, Taylor SA, et al. An analysis of the shoulder and elbow section of the orthopedic in-training examination. Am J Orthop Belle Mead NJ 2012;41:63-8.