

Impact of a training program on the performance of graduating Canadian residents on a national urology exam: Results of the last 20 years

Naji J. Touma, MD, FRCS(C)¹; Darren T. Beiko, MD, FRCS(C)¹; Andrew E. MacNeily, MD, FRCS(C)²; Michael J. Leveridge, MD, FRCS(C)¹

¹Department of Urology, Queen's University, Kingston, ON; ²Department of Urologic Sciences, University of British Columbia, Vancouver, BC; Canada

Cite as: *Can Urol Assoc J* 2018 July 31; Epub ahead of print.

<http://dx.doi.org/10.5489/cuaj.5386>

Published online July 31, 2018

Abstract

Introduction: Many factors impact the performance of graduating residents on certification exams. It is thought that most factors are related to the individual candidate's ability, motivation, and work ethic. Less understood, however, is whether a training program has any impact on the preparation and performance of its graduates on certification exams. We present 20 years of results of a national preparatory exam that all graduating residents complete about three months before the Royal College of Physicians and Surgeons of Canada (RCPSC) qualifying urology exam. This exam, known colloquially as QUEST, aims to simulate the RCPSC exam with written and oral components. We aimed to analyze the impact of a training program on the performance of its residents.

Methods: A retrospective review of exam results from 1997–2016 was conducted. During that time, 495 candidates from all 12 Canadian urology training programs undertook the exam. The performance of graduating residents from each individual program was grouped together for any given year. The different programs were anonymized, as the aim of this study is to assess the impact of a training program and not to rate the different programs. Statistical analysis using one-way ANOVA was conducted.

Results: All training programs fall within one standard deviation of the mean for the written component, the oral component, and the overall score. The residents of four training programs had statistically better scores than the overall mean of the written component. The residents of three out of these four training programs also had statistically better scores than the overall mean of the oral component and the overall results of the exam.

Conclusions: Most Canadian training programs prepare their residents adequately for this simulated certification exam in urology. However, there are some training programs that consistently prepare graduating residents to outperform their peers.

Introduction

Certification by the Royal College of Physicians and Surgeons of Canada (RCPSC) into the specialty of urology requires the successful completion of a five-year residency at an accredited institution and the successful sitting of a comprehensive exam at the end of training. (1) The content and the results of this certifying exam are not disclosed beyond a pass or fail result. The training programs do not receive feedback about the performance of their trainees compared to other programs. While protecting the privacy of individual candidates, this practice restricts the ability of a training program to gather important information about its own performance. The occasional resident failure may prompt a program to re-evaluate its pedagogic methodology. However, it is often difficult to ascertain the exact role of the training program in this negative outcome.

Knowledge acquisition is a complex process that happens over time. The level of knowledge attained is multifactorial and likely involves student specific abilities such as motivation, work ethic, and study behaviours. (2,3) The impact of a urology training program on the level of resident preparation is not known. We hypothesize that a training milieu is an independent predictor of a resident's exam performance.

The RCPSC is in the process of overhauling postgraduate medical education in Canada with the goal of establishing "competence by design" (CBD) as the new model for residency training. (4) A cornerstone of establishing a successful CBD program involves defining clear targets known as entrustable professional activities (EPAs) for attaining competency and creating the assessment tools to measure them. Before undertaking such major change within training programs, it would be helpful to define the state of knowledge about the performance of these programs under the current time-based model.

The purpose of this study is to evaluate the performance of graduating Canadian urology residents on a national preparatory exam over a 20-year period. We will seek to determine the impact a training program on the performance of its residents.

Methods

Since 1997 the department of urology at Queen's University has hosted a national preparatory exam that all Canadian graduating urology residents complete about 3 months before the RCPSC exam. The Queen's Urology Examination Skills Training, known colloquially as QUEST, aims to simulate the RCPSC exam with a written and an oral component. The written component consists of short answer questions (SAQ) and the oral component is an objective structured clinical exam (OSCE) composed of eight stations. Given its content and proximity to the RCPSC exam, QUEST reflects the state

of knowledge of soon-to-be graduates and constitutes a good assessment of the CanMEDS Medical Expert role.

Between 1997 and 2016, 495 graduating residents from all 12 Canadian urology training programs undertook the exam. The overall scores for the exam were determined by weighting the SAQ and OSCE marks equally. The results of the SAQ and OSCE components as well as the overall scores were then collated and grouped by residency training program. The training programs were then anonymized as the goal was not to rate the different training programs, but rather evaluate whether a particular training milieu impacts a trainee's level of knowledge and examination skills.

Statistical analysis

The data was evaluated and determined to adequately fit a normal distribution and hence, suitable to carry out parametric analysis utilizing the one-way ANOVA. A post-hoc analysis, using Tukey's test, comparing the mean scores between the different programs was then carried out.

Results

Plots of the SAQ, OSCE, and overall scores were all determined to fit reasonable normal distributions. There was also a moderate correlation between SAQ and OSCE scores confirming the internal validity of the exam with a R^2 coefficient of 0.33. (Figure 1) The overall, SAQ, and OSCE marks were found to fit within 1 standard deviation from the respective means for all 12 training programs.

A one-way ANOVA was conducted examining the impact of a training program on the SAQ, OSCE, and overall scores. ANOVA showed a significant effect of training program on SAQ scores ($F(12,482) = 5.488$, $p = .000$. (Figure 2)) as well as for OSCE scores ($F(12,482) = 8.319$, $p = .000$. (Figure 3) and overall scores ($F(12,482) = 8.049$, $p = .000$ (Table 1)).

A Tukey's honestly significant difference (HSD) post hoc test was then conducted for the SAQ, OSCE, and overall scores confirming that statistical differences indeed existed between certain programs.

Discussion

Over 20 years, there was a moderate correlation between candidates' SAQ scores and OSCE scores. This confirms that both parts of the exam reflect a candidate's knowledge base. The data from the first 5 years of QUEST (1997 to 2001) has previously been shown to be a moderate predictor of performance on the RCPSC exam and the American Board of Urology part 1 exam. (5)

Specialty accreditation of a urology training program by the RCPSC requires compliance with certain standards to achieve competence in graduating residents. From an academic perspective, programs offer didactic teaching sessions, grand rounds, and

multidisciplinary rounds among the tools to ensure knowledge transfer. However, there is no standardized and uniform curriculum between programs. Academic teaching is considered complimentary to the resident's effort to ensure preparedness. Indeed, 73% of Canadian chief urology residents report studying more than 25 hours a week in anticipation of the Royal College exam. (2) Until this study, it was unclear whether a training program has a significant impact on a resident's exam performance. It would appear that all programs prepare their residents adequately for exams as the results for all programs fall within 1 standard deviation of the mean. However, the residents of a few select Canadian programs indeed outperform their peers on board exams.

There are many factors that can improve a program's preparation of its own residents. For instance, a formalized curriculum has been shown to correlate with performance on United States Medical Licensing Examination (USMLE) scores and otolaryngology training examination scores. (6,7) An educational program of weekly assigned reading, followed by weekly examinations prepared and administered by the program director results in improvements in the 5-year pass rates on the American Board of Surgery (ABS) qualifying and certifying exams. (8) It is therefore possible that the variability in the formal teaching between training programs could account for some of the discrepancy.

Another factor that affects board exam outcomes is the performance on in-training examinations. This has been shown across a broad spectrum of specialties. (9-13) Monthly mock orals and simulated exams have been shown to improve outcomes on the ABS certifying exam. (14,15) These exams give trainees real-time feedback about their progress and may allow for rehearsal and development of familiarity with high stakes exams. It is therefore possible that training programs that institute frequent and rigorous exams during training may imbue their residents with a potential advantage.

In addition to curricular and in-training exams, some non-academic characteristics have also been shown to predict board exam performance. For instance, the number of on-call shifts in the six months prior to an exam has been shown to inversely impact performance. (9) Some common attributes of high performing American general surgery programs include: rotating at a level 1 trauma center, spending more time with simulation, and, offering numerous fellowship programs. (16)

Beyond board exam performance, the data is mixed about the impact of a training program on patient outcomes. For instance, an Obstetrics training program was found to be a predictor of the maternal complication rates of its graduates. (17) Conversely, the surgical performance of bypass surgeons was not associated with graduation from any particular medical school, residency, or fellowship program. (18)

There are several limitations to our data that must be considered. Firstly, training programs change substantially over a 20-year period. Several program directors would have been at the helm. The turnover of faculty is likely to be substantial. This natural

renewal in personnel would inevitably bring fresh perspectives and overhaul of academic curricula. Secondly, it could be that different residents tend to gravitate to programs that suit their skills, natural abilities, personalities and career aspirations. For instance, a program's reputation for research opportunities or clinical exposure may attract different residents with different abilities for studying or board exam performance. A prime example of a resident's characteristics influencing a training program's results would be the admission of international medical graduates (IMG). Over the years, some programs have had a policy of consistently welcoming IMGs and others not. Although not analyzed here, it has been shown that these residents tend to underperform on board exams. (19) This data should be interpreted with caution. It provides a snapshot of exam performance. A number of other qualities of competent graduating urologists are not measured here. Such qualities include surgical skills and clinical acumen that may have substantially more impact on real world performance and patient outcomes. Other qualities that are not measured include the intrinsic roles of CanMEDS beyond the medical expert role. In addition, the performance of residents on the different topics examined is not analyzed as the breakdown of the exam content does not exist over a 20-year period. However, this data is being collected prospectively and will be reported in the future to provide a more fulsome picture of the trainees' performance.

Conclusion

Most Canadian training programs prepare their residents adequately for this simulated certification exam in urology. However, there are some training programs that consistently prepare graduating residents that outperform their peers. Further research to elucidate the qualities that make a training program successful in academically preparing its residents is needed.

References

1. Specialty Training Requirements in Urology. www.royalcollege.ca; accessed March 2018
2. Skinner TAA, Ho L, Touma NJ: Study habits of Canadian urology residents: Implications for development of a competence by design curriculum. *Can Urol Assoc J.* 2017; 11(3-4) 83-87
3. Burk-Rafel J, Santen SA, Purkiss J: Study behaviors and USMLE Step 1 Performance: Implications of a Student Self-Directed Parallel Curriculum. *Acad Med.* 2017; 92 S67-S74
4. Competence by Design: ebook. www.royalcollege.ca; accessed March 2018
5. MacNeily AE, Baverstock RJ, Cole G, et al: Quantitative assessment of a new preparatory tool for board certification in urology. *BJU Int* 2004, 93(4); 558-61
6. Redmann AJ, Tawfik KO, and Myer CM 3rd: The impact of a resident-run review curriculum and USMLE score on the Otolaryngology in-service exam. *Int J. Pediatr Otolaryngol* 2018, 104; 25-28
7. Reh DD, Ahmed A, Li R et al: A learner-centered educational curriculum improves resident performance on the otolaryngology training examination. *Laryngoscope* 2014, 124(10); 2262-7
8. De Virgilio C, Chan T, Kaji A et al: Weekly assigned reading and examinations during residency, ABSITE performance, and improved pass rates on the American Board of Surgery Examinations. *J Surg Educ.* 2008, 65(6); 499-503
9. Brateanu A, Yu C, Kattan MW et al: A nomogram to predict the probability of passing the American Board of Internal Medicine examination. *Med Educ Online* 2012; 17: 18810
10. Indik JH, Duhigg LM, McDonald FS et al: Performance on the Cardiovascular In-Training Examination in relation to the ABIM Cardiovascular Disease Certification Examination. *J Am Coll Cardiol* 2017; 69(23); 2862-8
11. Grabovsky I, Hess BJ, Haist SA et al: The relationship between performance on the Infectious Diseases In-Training and Certification Examinations. *Clin Infect Dis* 2015 60(5): 677-83
12. Lohr KM, Clauser A, Hess BJ et al: Performance on the adult rheumatology in-training examination and relationship to outcomes on the rheumatology certification examination. *Arthritis Rheumatol* 2015, 67(11), 3082-90
13. Kempainen RR, Hess BJ, Addrizzo-Harris DJ et al: Pulmonary and critical care in-service training examination score as a predictor of board certification examination performance. *Ann Am Thorac Soc* 2016, 13(4); 481-8
14. Guzman E, Bakakhani A, Maker VK: Improving outcomes on the ABS certifying examination: can monthly mock orals do it? *J Surg Ed.* 2008, 65(6); 441-444

15. London D, Awad M: The impact of an advanced certifying examination simulation program on the American Board of Surgery certifying examination passage rates. *J Am Coll surg.* 2014, 219(2); 280-4
16. Bankhead-Kendall B, Slama E, et al: Common attributes of high/low performing general surgery programs as they relate to QE/CE pass rates. *Am J Surg* 2016, 212(6); 1248-50
17. Asch DA, Nicholson S, Srinivas S, et al: Evaluating obstetrical residency programs using patient outcomes. *JAMA* 2009; 302(12): 1277-83
18. Hartz AJ, Kuhn EM, Pulido J: Prestige of training programs and experience of bypass surgeons as factors in adjusted patient mortality rates. *Med Care* 1999, 37(1); 93-103
19. Karnik A, Venkatesh B, Angelico D: Analysis of performance and predictors of success in the final fellowship examination of the college of intensive care medicine. *Crit Care Resusc* 2015, 17(1); 47-50

Figures and Tables

Fig. 1. Correlation between SAQ and OSCE scores.

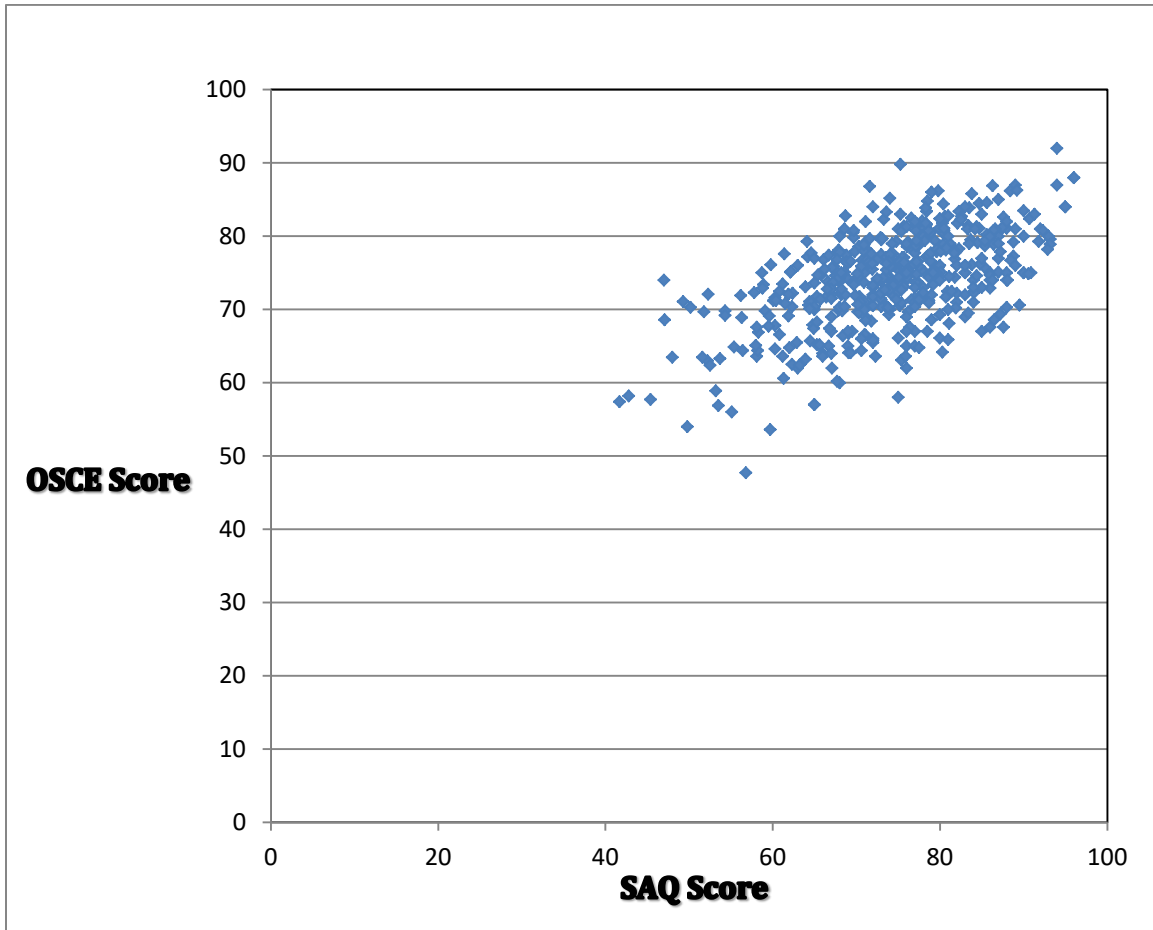


Fig. 2. Performance of different programs on the SAQ exam.

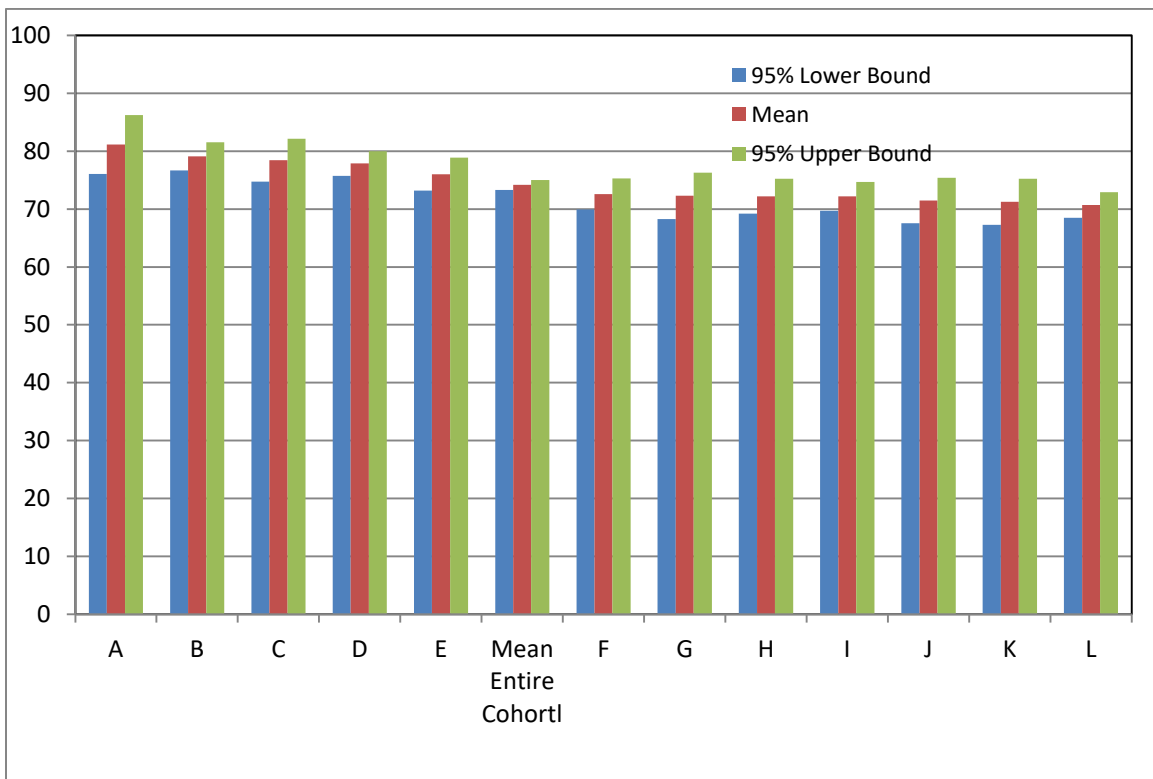


Fig. 3. Performance of different programs on the OSCE exam.

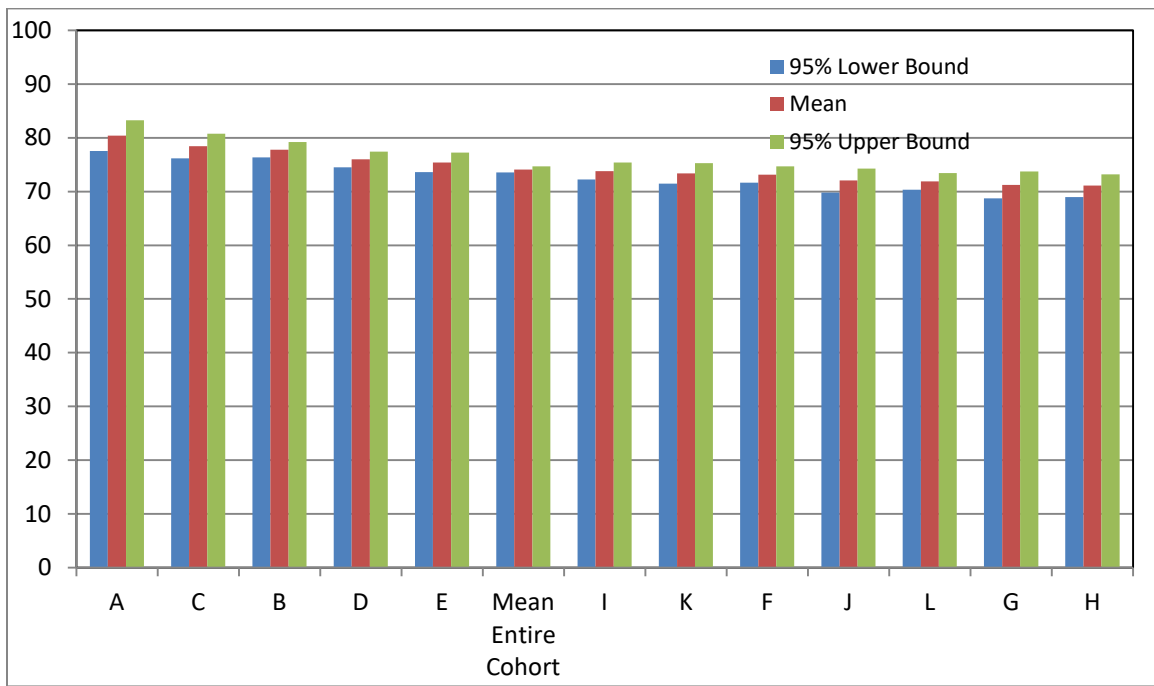


Table 1. Overall marks on the exam by different training program

Training program	n	Mean	Standard deviation	Standard error	95% confidence interval for mean	
					Lower bound	Upper bound
					A	20
B	45	78.4	5.4	0.8	76.8	80.0
C	20	78.4	5.1	1.1	76.0	80.8
D	67	76.9	6.4	0.8	75.4	78.5
E	37	75.5	6.0	1.0	73.5	77.5
F	52	72.8	6.4	0.9	71.1	74.6
G	30	72.2	7.9	1.4	69.2	75.1
H	38	71.7	7.3	1.2	69.3	74.1
I	44	73.0	5.5	0.8	71.3	74.6
J	30	71.7	7.7	1.4	68.9	74.6
K	29	72.3	6.7	1.2	69.8	74.8
L	72	71.3	7.1	0.8	69.6	73.0
Unaffiliated candidates	11	69.3	8.1	2.4	63.8	74.7
Total	495	74.1	7.2	0.3	73.5	74.8