

Assessing Residents' Competency at Baseline: How Much Does the Medical School Matter?

Nathan S. Gollehon, MD
R. Brent Stansfield, PhD
Larry D. Gruppen, PhD
Lisa Colletti, MD

Hilary Haftel, MD, MHPE
James O. Woolliscroft, MD
Monica L. Lypson, MD, MHPE

ABSTRACT

Background Although there is some consensus about the competencies needed to enter residency, the actual skills of graduating medical students may not meet expectations. In addition, little is known about the association between undergraduate medical education and clinical performance at entry into and during residency.

Objective We explored the association between medical school of origin and clinical performance using a multi-station objective structured clinical examination for incoming residents at the University of Michigan Health System.

Methods Prior to assuming clinical duties, all first-year residents at the University of Michigan Health System participate in the Postgraduate Orientation Assessment (POA). This assesses competencies needed during the first months of residency. Performance data for 1795 residents were collected between 2002 and 2012. We estimated POA variance by medical school using linear mixed models.

Results Medical school predicted the following amounts of variance in performance—data gathering scores: 1.67% (95% confidence interval [CI] 0.36–2.93); assessment scores: 4.93% (95% CI 1.84–6.00); teamwork scores: 0.80% (95% CI 0.00–1.82); communication scores: 2.37% (95% CI 0.66–3.83); and overall POA scores: 4.19% (95% CI 1.59–5.35).

Conclusions The results show that residents' medical school of origin is weakly associated with clinical competency, highlighting a potential source of variability in undergraduate medical education. The practical significance of these findings needs further evaluation.

Introduction

The medical education community and accreditors have recognized that outcome- and competency-based assessment methods are important to ensuring physicians entering practice are prepared for the tasks they face.^{1–4} Despite this, there are no common mandated practices to ensure these professional competencies are achieved. Historically, there has been little agreement about essential skills and knowledge at the educational handoff between medical school and residency.^{5–7} More recently, concerns have been expressed that many graduating medical students fail to meet expectations.⁸

Prior studies suggested that variation in undergraduate medical education affects subsequent clinical competency and future practice.^{9–14} While a few studies have assessed residents' competence at baseline, to our knowledge, there are no publications quantifying the correlation of medical school of origin and matriculating residents' clinical competency.^{15–17}

Medical schools are challenged to determine individual learner competency prior to graduation,

and residency programs are responsible for quickly identifying residents' deficits in order to plan for remediation and supervision.¹⁸ A baseline assessment of resident competency is a useful tool for identifying gaps in knowledge and skills,¹⁹ and can provide a starting point for a resident's journey toward overall competence.²

In 2002 the University of Michigan Health System (UMHS) mandated that all incoming residents participate in a Postgraduate Orientation Assessment (POA). This objective structured clinical examination (OSCE) assesses baseline competency across all specialty programs that accept first-year trainees. In the current study, we analyzed data on core skills and preparedness for residency from more than a decade, and evaluated performance differences by medical school of graduation. We hypothesize that the medical school where an incoming resident has trained correlates with clinical skills in a measurable way as quantified by our OSCE.

Methods

The POA is a 10-station OSCE administered during resident orientation.^{19,20} The OSCE format allows assessment of core skills other testing modalities are not able to capture, including teamwork, learning strategies, time management, clinical reasoning, and

DOI: <http://dx.doi.org/10.4300/JGME-D-17-00024.1>

Editor's Note: The online version of this article contains the validity evidence document.

judgment.²¹ The POA was designed with the guiding principle that testing drives learning and skills acquisition. It focuses on competencies needed during the first months of residency. It is administered at UMHS, a large academic hospital system that accepts approximately 180 residents annually.

Residents were assessed in 4 groupings of core skills based on Wagner and Lypson's categorization: (1) data gathering; (2) clinical assessment; (3) team skills and procedural competence; and (4) communication.¹⁹

The project was granted educational exemption status by the UMHS Institutional Review Board.

Direct comparison of POA scores across years was impractical because several POA stations were revised or eliminated over the years. We eliminated differences in station difficulty by Z-normalizing scores within each year, so means were zero and standard deviations were 1 for each station for each year. This transformation preserves information about resident skills relative to others in the same cohort, but does not allow for direct comparisons of relative skills level between years.

Each station measured 1 or more core skills. Core skills scores were computed as the average of all relevant station scores (TABLE 1), and each station was given equal weight. The overall POA scores were computed as the unweighted average of all stations, and were used as a measure of overall clinical skills.

For each core skills score, we conducted a mixed model univariate analysis (ie, hierarchical linear model) with a random intercept that varied across medical schools using restricted maximum likelihood (REML) estimation. This model estimated the percentage of resident core skills variance attributable to the residents' medical school using a post hoc Markov Chain Monte Carlo model, and generated estimates (with 95% confidence intervals) of the direction and magnitude of schools' impact on residents' scores.²² Schools where residents consistently outperform others in their cohort receive a higher estimate. These estimates are shrunk toward zero for schools with fewer students, resulting in conservative estimates. Schools were then ranked according to their performance relative to the overall mean.

We investigated the validity of normalized scores by examining the correlation pattern of core skills scores with residents' United States Medical Licensing Examination (USMLE) scores and specialty board examination scores. Correlations with USMLE scores were tested using Pearson's *r*. Specialty board scores were available for a subset of residents. These scores were Z-normalized within year due to a change in the

What was known and gap

There is concern that the skills of graduating medical students may not meet expectations at entry into residency.

What is new

A multi-year institutional study assessed the association between undergraduate medical education and clinical performance at entry into residency.

Limitations

Single institution study may limit generalizability; potential for selection bias from multiple sources.

Bottom line

Residents' medical school of origin is weakly correlated with clinical competency as measured by a standardized objective structured clinical examination.

scoring scale in 2013. Additional details about background validity evidence for the POA, including Cronbach's alpha of core skills scores, are included as online supplemental material. All analyses were performed using R version 2.11.1 (The R Foundation, Vienna, Austria).

Results

Baseline characteristics of participants were collected, including sex, race, medical degree, specialty, and USMLE Step 1 and 2 Clinical Knowledge (CK) scores (TABLE 2). During the study period 1795 residents from 139 US and 33 international medical schools participated in the POA, and results for all test takers were included in our analysis.

Medical school predicted the following amounts of variance in performance—data gathering scores: 1.67% (95% confidence interval [CI] 0.36–2.93); assessment scores: 4.93% (95% CI 1.84–6.00); teamwork scores: 0.80% (95% CI 0.00–1.82); communication scores: 2.37% (95% CI 0.66–3.83); and overall POA scores: 4.19% (95% CI 1.59–5.35; TABLE 3). All reported results were statistically significant at the $P < .05$ level.

The correlation between POA core skills and overall scores versus USMLE Step 1 and 2 scores is shown in TABLE 4. All core skills scores as well as overall scores positively correlated with Step 1 and Step 2 CK scores, with the only exception being communication scores and Step 1 scores. Core skills and overall scores tended to be more strongly correlated with Step 2 CK than Step 1.

Specialty board scores were available for 210 residents who participated in the POA. For this subset of residents, overall POA scores were significantly related to board scores ($P = .028$).

Estimates of individual school effect on the mean overall POA scores for each school were calculated using REML. This estimate reflects how much better

TABLE 1

Postgraduate Orientation Assessment Blueprint: Core Skills Mapped to Station and Accreditation Council for Graduate Medical Education (ACGME) Competency^a

Core Skill	Station	ACGME Competency
Data gathering - History taking - Physical examination skills - Communication skills	Obtaining informed consent	Patient care - Interpersonal and communication skills - Professionalism
	Cross-cultural communications	
	Pediatric history: obtaining a history from a proxy	
	Assessing geriatric functional status, examination techniques/history	
	Assessing patient pain	
Assessment - Note writing - Order writing - Critical values and imaging - Evidence-based medicine	Order writing: emergency room orders	Patient care - Medical knowledge - Systems-based practice - Practice-based learning and improvement
	Evaluating critical laboratory values	
	Reading and evaluating images and radiation safety	
	Evidence-based medicine: formulating a clinical question, choosing the best evidence, and interpretation of the data	
Team skills/procedure competencies - Hand hygiene - Aseptic technique - Handoffs - Crisis management	Handwashing basics: knowledge and skills	Systems-based practice - Interpersonal and communication skills - Professionalism - Patient care
	Aseptic technique (mock incision and drainage)	
	Fire safety knowledge (including surgical fire and hazard management)	
	Sign-out skills: patient handoff (written and verbal)	
Communication - Communication skills - History taking	Cross-cultural communications	Interpersonal and communication skills Patient care
	Obtaining informed consent	
	Assessing geriatric functional status, examination techniques/history	
	Verbal patient handoff	
	Overall communication (cross station assessment of communication skills)	

^a Adapted with permission from Wagner D, Lyson ML. Centralized assessment in graduate medical education: cents and sensibilities. *J Grad Med Educ.* 2009;1(1):21–27.

or worse a resident from a particular school is likely to perform relative to the overall mean. Variance from the mean, which was based on medical school, ranged from 0.46 to 0.26, suggesting a modest relationship with student performance. The relative performance of the medical schools is illustrated in the FIGURE.

Discussion

We found statistically significant variance in performance attributable to medical school for the overall POA scores as well as all core skills scores except teamwork. The effect sizes were small, based on accepted interpretations of effect sizes (ie, 1% is small, 9% is medium, and 25% is large).²³ Interestingly, the magnitude of performance variance in this study was similar to prior studies of interschool variability on USMLE Step 2 CK performance.^{24,25} The results suggest that medical school of origin does correlate with clinical performance and competency,

and this correlation differs depending on the core skills assessed.

The overall POA scores can be interpreted as a measure of general resident skills, weighted toward clinically based skills. With each station score normalized within year to account for possible differences in station difficulty between years, and to equate measurement variance between stations, the mean of these normalized scores is an estimate of student ability across stations regardless of testing year.

Core skills scores, although unbiased, are noisy estimates of resident skills based on means of weakly correlated station scores with low Cronbach's alpha. It is noteworthy that school effects were still apparent, despite our conservative analysis using REML, and it is likely that more precise measures of clinical skills would find larger effects.

USMLE Step 1 is a direct test of medical knowledge, while Step 2 CK assesses students' ability to evaluate patients and apply more complex medical

TABLE 2

Characteristics of University of Michigan First-Year Residents Who Participated in Postgraduate Orientation Assessment

Characteristic	n (%)
Sex	
Male	995 (55)
Female	800 (45)
Race	
Asian/Pacific Islander	385 (21)
African American	65 (4)
Hispanic	37 (2)
White	1230 (69)
Multiple/question not answered	80 (4)
Degree ^a	
MD	1332 (97)
DO	10 (< 1)
Other (DDS) or not recorded	30 (2)
Specialty	
Primary care ^b	838 (47)
Non-primary care	959 (53)
USMLE score, mean (SD)	
Step 1	231.3 (18.4)
Step 2 CK	237.2 (19.9)

Abbreviations: MD, doctor of medicine; DO, doctor of osteopathic medicine; DDS, doctor of dental surgery; USMLE, United States Medical Licensing Examination; CK, Clinical Knowledge.

^a Degree was not recorded from 2002–2004. These results represent only data from 2005–2012.

^b Primary care defined as family medicine, internal medicine, pediatrics, and combined internal medicine–pediatrics.

concepts. USMLE Step 2 CK correlated most strongly with the assessment core skills scores, which was expected given the competencies assessed. The observed intercorrelation pattern of POA core skills scores and USMLE scores provides validity evidence for using both of these core skills scores as estimates of residents' application of medical knowledge and clinical skills. Correlations of communication, data gathering, and teamwork core skills scores with USMLE scores were low (r between 0.05 and 0.08), consistent with the findings of prior studies.^{26,27} The

TABLE 3

Percentage of Variance in Performance by Core Skill and Overall Score Attributable to Medical School of Training

Core Skill	Variance, %	95% CI
Data gathering	1.67	0.36–2.93
Assessment	4.93	1.84–6.00
Teamwork	0.80	0.00–1.82
Communication	2.37	0.66–3.83
Overall score	4.19	1.59–5.35

Abbreviation: CI, confidence interval.

TABLE 4

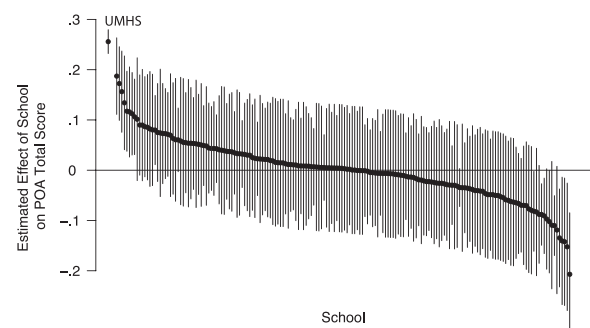
Results of Statistical Correlation Analysis Comparing USMLE Step 1 and Step 2 CK Scores Versus Core Skill and Overall Score

Test	Core Skill	r	P
Step 1	Data gathering	0.07	< .010
	Assessment	0.24	< .0001
	Teamwork	0.05	< .05
	Communication	0.04	NS
	Overall score	0.19	< .0001
Step 2 CK	Data gathering	0.10	< .0001
	Assessment	0.28	< .0001
	Teamwork	0.08	< .005
	Communication	0.07	< .010
	Overall score	0.24	< .0001

Abbreviations: USMLE, United States Medical Licensing Examination; CK, Clinical Knowledge; NS, not significant.

positive correlation between overall POA scores and specialty board scores provides additional validity evidence for the POA as a measure of clinical knowledge and understanding.

As predicted, all correlations were positive, and the strongest correlations were between assessment core skills scores and residents' USMLE Step 2 CK scores; Step 1 scores were the next strongest. Correlation patterns were appropriate for the other core skills scores: communication core skills scores correlated best with Step 2 CK scores and data gathering correlated best with Step 1 scores. While communication and data gathering core skills scores strongly correlated with each other, their distinct patterns of correlation with USMLE scores indicate that they serve as estimates of 2 correlated but dissociable skills. Teamwork core skills scores correlated most

**FIGURE**

Individual School Performance

Abbreviations: UMHS, University of Michigan Health System; POA, Postgraduate Orientation Assessment.

Note: Estimated variance in performance for individual medical schools. The median estimate of school effect is indicated by a dark dot, with 95% confidence intervals indicated by the gray error bars. Zero on the y-axis represents mean performance for all schools.

strongly with USMLE Step 2 CK scores and least strongly with Step 1 scores.

The overall scores demonstrated validity evidence by correlating most strongly with USMLE Step 2 CK scores, yet also with USMLE Step 1 scores. Taken together, residents' overall POA scores appear to be useful measures of resident clinical skills.

Our data underestimated school effects because our analysis and variance estimates are inherently conservative given the use of REML. The intrayear normalization of scores eliminates all variance due to yearly differences in cohort skills.

There is some measurement error in our data, as each year has a somewhat different set of stations. Nonetheless, measured constructs are the same from year to year, and stations were changed with the goal of improving the measurement of core skills. The commonalities between years will likely overshadow the differences.

This study has several limitations. It was conducted at a single institution with generally stringent selection criteria, reducing the ability to generalize to other sites. It is likely that there were regional effects. While incoming residents came from a geographically diverse set of medical schools, proportionally more came from nearby schools. There was also likely selection bias at several levels. Each program selected residents based on its own set of criteria, and letters of recommendation and medical school grades likely affected who was offered an interview and ultimately entered the programs. Program directors also may have considered the prestige of the medical school of origin when ranking candidates' applications.²⁸ Finally, prospective residents' perceptions of programs at the University of Michigan likely played a role in their decision-making. Graduates of the University of Michigan Medical School performed substantially better on the POA than their peers. Assessment of these students is almost certainly skewed by familiarity with the assessment, as several POA stations are used in a fourth-year medical student OSCE. We did not collect data on the curricula used at individual schools, so a direct analysis of the effect of specific curricula was not possible. While USMLE Clinical Skills data exist and could provide additional validity for our results, these results are not currently available for analysis.

Further studies should focus on greater clarification of institutional and curricular characteristics that may contribute to variability. Additional work is also needed on approaches to use identified performance differences to guide subsequent educational interventions during residency.

Conclusion

Our results suggest that residents' medical school of origin is weakly correlated with clinical competency as measured by a standardized OSCE.

References

1. Swing SR, Clyman SG, Holmboe ES, et al. Advancing resident assessment in graduate medical education. *J Grad Med Educ*. 2009;1(2):278–286.
2. Nasca TJ, Philibert I, Brigham T, et al. The next GME accreditation system—rationale and benefits. *N Engl J Med*. 2012;366(11):1051–1156.
3. No authors listed. CanMEDS 2000: extract from the CanMEDS 2000 project societal needs working group report. *Med Teach*. 2000;22(6):549–554.
4. Liaison Committee on Medical Education. Scope and purpose of accreditation. <http://lcme.org/about>. Accessed July 5, 2017.
5. Scott CS, Barrows HS, Brock DM, et al. Clinical behaviors and skills that faculty from 12 institutions judged were essential for medical students to acquire. *Acad Med*. 1991;66(2):106–111.
6. Bass EB, Fortin AH 4th, Morrison G, et al. National survey of clerkship directors in internal medicine on the competencies that should be addressed in the medicine core clerkship. *Am J Med*. 1997;102(6):564–571.
7. Drafting Panel for Core Entrustable Professional Activities for Entering Residency. *Core Entrustable Professional Activities for Entering Residency (CEPAER)*. Washington, DC: Association of American Medical Colleges; 2013.
8. Moercke AM, Eika B. What are the clinical skills levels of newly graduated physicians? Self-assessment study of an intended curriculum identified by a Delphi process. *Med Educ*. 2002;36(5):472–478.
9. Crobett EC, Whitcomb M. *The AAMC Project on the Clinical Education of Medical Students: Clinical Skills Education*. Washington, DC: Association of American Medical Colleges; 2004.
10. Ripkey DR, Swanson DB, Case SM. School-to-school differences in step 1 performance as a function of curriculum type and use of step 1 in promotion/graduation requirements. *Acad Med*. 1998;73(suppl 10):16–18.
11. Hecker K, Violato C. How much do differences in medical schools influence student performance? A longitudinal study employing hierarchical linear modeling. *Teach Learn Med*. 2008;20(2):104–113.
12. Levy AR, Tamblyn RM, McLeod PJ, et al. The effect of physicians' training on prescribing beta-blockers for secondary prevention of myocardial infarction in the elderly. *Ann Epidemiol*. 2002;12(2):86–89.
13. Monette J, Tamblyn RM, McLeod PJ, et al. Do medical education and practice characteristics predict

- inappropriate prescribing of sedative-hypnotics for the elderly? *Acad Med*. 1994;69(suppl 10):10–12.
14. Asch DA, Nicholson S, Srinivas S, et al. Evaluating obstetrical residency programs using patient outcomes. *JAMA*. 2009;302(12):1277–1283.
 15. Stachnik TJ, Simons RC. A comparison of DO and MD student performance. *J Med Educ*. 1977;52(11):920–925.
 16. Remmen R, Derese A, Scherpier A, et al. Can medical schools rely on clerkships to train students in basic clinical skills? *Med Educ*. 1999;33(8):600–605.
 17. Dickson GM, Chesser AK, Keene Woods N, et al. Self-reported ability to perform procedures: a comparison of allopathic and international medical school graduates. *J Am Board Fam Med*. 2013;26(1):28–34.
 18. Angus S, Vu TR, Halvorsen AJ, et al. What skills should new internal medicine interns have in July? A national survey of internal medicine residency program directors. *Acad Med*. 2014;89(3):432–435.
 19. Wagner D, Lypson ML. Centralized assessment in graduate medical education: cents and sensibilities. *J Grad Med Educ*. 2009;1(1):21–27.
 20. Lypson ML, Frohna JG, Gruppen LD, et al. Assessing residents' competencies at baseline: identifying the gaps. *Acad Med*. 2004;79(6):564–570.
 21. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226–235.
 22. Gelfand AE, Smith AF. Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc*. 1990;410(85):398–409.
 23. Cohen J, Cohen P, West SG, et al. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 3rd ed. New York, NY: Lawrence Erlbaum Associates; 2003.
 24. Cuddy MM, Swanson DB, Dillon GF, et al. A multilevel analysis of the relationships between selected examinee characteristics and United States Medical Licensing Examination Step 2 Clinical Knowledge performance: revisiting old findings and asking new questions. *Acad Med*. 2006;81(suppl 10):103–107.
 25. Case SM, Ripkey DR, Swanson DB. The relationship between clinical science performance in 20 medical schools and performance on Step 2 of the USMLE licensing examination. 1994–95 Validity Study Group for USMLE Step 1 and 2 Pass/Fail Standards. *Acad Med*. 1996;71(suppl 1):31–33.
 26. Cohen ER, Barsuk JH, Moazed F, et al. Making July safer: simulation-based mastery learning during intern boot camp. *Acad Med*. 2013;88(2):233–239.
 27. Dong T, Saguil A, Artino AR Jr, et al. Relationship between OSCE scores and other typical medical school performance indicators: a 5-year cohort study. *Mil Med*. 2012;177(suppl 9):44–46.
 28. National Resident Matching Program. Results of the 2012 NRMP Program Director Survey. <http://www.nrmp.org/wp-content/uploads/2013/08/programresultsbyspecialty2012.pdf>. Accessed July 5, 2017.



Nathan S. Gollehon, MD, is Assistant Professor of Neonatology and Associate Pediatrics Clerkship Director, Department of Pediatrics, University of Nebraska Medical Center; **R. Brent Stansfield, PhD**, is Assistant Professor, Department of Learning Health Sciences, University of Michigan Medical School; **Larry D. Gruppen, PhD**, is Professor, Department of Learning Health Sciences, and Director, Master of Health Professions Education Program, University of Michigan Medical School; **Lisa Colletti, MD**, is Professor Emeritus, Department of Surgery, University of Michigan Medical School; **Hilary Haftel, MD, MHPE**, is Professor of Pediatrics and Communicable Diseases, Departments of Internal Medicine and Learning Health Sciences, and Associate Chair of Pediatric Education, University of Michigan Medical School; **James O. Woolliscroft, MD**, is Lyle C. Roll Professor of Medicine, Departments of Internal Medicine and Learning Health Sciences, and former Dean, University of Michigan Medical School; and **Monica L. Lypson, MD, MHPE**, is Director of Medical and Dental Education, Office of Academic Affiliations, Department of Veteran Affairs, and Adjunct Professor of Internal Medicine and Learning Health Sciences, University of Michigan Medical School.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

These results were presented as a poster at the Accreditation Council for Graduate Medical Education Annual Educational Conference, National Harbor, Maryland, February 27–March 2, 2014; as an oral presentation at the Association of American Medical Colleges Learn Serve Lead Annual Meeting, Seattle, Washington, November 11–15, 2016; and as an abstract in the 2016 Research in Medical Education supplement to *Academic Medicine*.

The authors would like to thank Paula Ross, PhD, University of Michigan Medical School, for her editorial assistance with this article, and Trina O'Boyle for her work in collecting and organizing the data used for this project.

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government.

Corresponding author: Nathan S. Gollehon, MD, University of Nebraska Medical Center, Department of Pediatrics, 8200 Dodge Street, Omaha, NE 68114-4113, 402.955.6140, ngolleho@unmc.edu

Received January 10, 2017; revisions received April 5, 2017, and June 2, 2017; accepted June 6, 2017.