Evidence for Increasing Diversity in Graduate Medical Education: The Competence of Underrepresented Minority Residents Measured by an Intern Objective Structured Clinical Examination

Monica L. Lypson, MD
Paula T. Ross, MA
Stanley J. Hamstra, PhD
Hilary M. Haftel, MD, MHPE
Larry D. Gruppen, PhD
Lisa M. Colletti, MD

Abstract

Objective Some have commented that the limited number of underrepresented minorities (URMs) in United States' residency programs is due to a lack of qualified candidates. At the University of Michigan, an objective structured clinical examination is administered to incoming residents at the beginning of training to determine baseline competence. In this study we wanted to determine if competence differed for underrepresented minorities when compared to non-URM residents.

Method The postgraduate orientation assessment, a 10-station examination, was developed that focused specifically on the knowledge and skills needed in the first 6 to 18 weeks of training. Stations assessed competence in informed consent, aseptic technique, evidence-based medicine, diagnostic images, critical laboratory values, cross-cultural communication, and Joint Commission requirements such as surgical fire safety, pain assessment, and management. We used various assessment measures including standardized patients, computer-based testing, and multiple-choice questions.

Results Our study found no significant differences in overall mean scores between URM residents and all other residents for the 5 years during which we administered the examination, except for 2002. This stands in contrast to the consistently worse performances of URM students on USMLE Step 1 and Step 2 Clinical Knowledge. Also, URM residents did not perform better or worse than their non-URM colleagues on standardized patient stations during the course of 5 years during which the examination was administered.

Conclusions The postgraduate orientation assessment provides residency program directors with a standard format to measure initial clinical skills. When compared to incoming non-URM residents from a variety of medical schools, URM residents perform as well as other trainees. Our results may aid in the recruitment efforts of URM medical students into academic residency programs such as those at the University of Michigan.

Monica L. Lypson, MD, is Assistant Dean for Graduate Medical Education and Associate Professor, Department of Internal Medicine, at University of Michigan Medical School; **Paula T. Ross, MA,** is Research Area Specialist, Department of Health Behavior and Health Education, at University of Michigan; **Stanley J. Hamstra, PhD,** is Acting Assistant Dean, Academy for Innovation in Medical Education, and University of Ottawa Faculty of Medicine Research Director at University of Ottawa Skills and Simulation Centre, Ottawa, Ontario, Canada; **Hilary M. Haftel, MD, MHPE,** is Associate Professor of Pediatrics, Internal Medicine, and Medical Education, Associate Chair for Pediatric Education, and Pediatric Residency Program Director at University of Michigan Medical School; **Larry D. Gruppen, PhD,** is Josiah Macy Jr Professor of Medical Education and Chair, Department of Medical Education, at University of Michigan Medical School; and **Lisa M. Colletti, MD,** is Associate Dean for Graduate Medical Education and C. Gardner Child Professor of Surgery at University of Michigan Medical School, Ann Arbor, Michigan.

Corresponding author: Monica L. Lypson, MD, University of Michigan Medical School, Graduate Medical Education, 2600 Green Road No. 150, Ann Arbor, MI 48105, 734.764.3186, mlypson@umich.edu

Received March 19, 2010; revision received May 18, 2010; accepted June 21, 2010.

DOI: 10.4300/JGME-D-10-00050.1

Introduction

Underrepresented minority (URM) students (before 2003 this group included African American, Mexican American, Native American, and mainland Puerto Rican students) encounter numerous barriers as they travel the long road to becoming a physician.[1] Barriers for URM students begin to emerge before entry into medical school and include lower performance on standardized tests (eg, ACT, SAT, and MCAT) when compared to that of their Non-Hispanic White and Asian American counterparts.[2–4] Seemingly, these performance and outcome trends persist throughout medical training on measures such as basic science examinations and United States Medical Licensing Examination (USMLE) Step 1 and Step 2.[4–6] As students advance to the resident application phase, more subjective measures (eg, letters of recommendation, personal statements, and interviews) are considered as well and, to a

small degree, may counterbalance the lower academic credentials of some URM applicants.[7,8] Previous studies[4,9,10] have demonstrated the value of standardized tests in predicting future success, specifically performance on board certification examinations. These findings are often used to support the consideration of standardized tests when making critical decisions for medical school admissions, residency program placement, and subsequent practice locations or faculty appointments.

As standardized measures typically assess knowledge-based competencies, almost all medical schools have also incorporated objective structured clinical examinations (OSCEs) as a means of assessing students' clinical and communication skills.[11–13] Likewise, in our attempt to capture baseline clinical competencies, we launched a 10-station postgraduate orientation assessment (POA) OSCE.[13,14] This is administered during orientation; interns are informed of the assessment, but they are not provided with information on what is included. The POA is a formative assessment that focuses on knowledge and skills that residents would likely be required to use during the first 6 to 18 weeks of their residency and in situations without supervision. This content is different from that offered by USMLE Step 2 Clinical Skills. During the past 5 years of administering the POA, we have been able to assess initial proficiency levels on the 6 key competency domains (ie, medical knowledge, patient care, professionalism, communication skills, practice-based learning and improvement, and systems-based practice) mandated by the Accreditation Council for Graduate Medical Education.[15] The overall POA score and performance on several stations (aseptic technique, informed consent, and system compliance/surgical fire safety) have been shown to be of moderate predictive value in performance on the percentage correct score on board certification score performance among University of Michigan pediatric residency graduates (M. L. Lypson, MD, J.A. Purkiss, PhD, H. M. Haftel, MD, MHPE, unpublished data, 2009).

Standardized tests place countless URM students at a disadvantage when they apply to medical school and may impede many from entering the gateway to a career in medicine. Given the recent challenges in race-conscious versus race-blind admissions, the ability of programs to maintain the minority physician pipeline has become increasingly more difficult. In this study, we sought to determine if in fact there were differences in performance between our underrepresented minority and non-URM residents on standardized measures when compared to an OSCE. Moreover, since the state of Michigan has recently legislated that our institution can no longer consider race as a factor in admissions decisions, we endeavored to perform a retrospective analysis of the initial performance of our incoming interns. To that end, we also sought to determine if past affirmative action recruitment efforts at the University of Michigan had resulted in less capable physicians as measured by our initial intern examination.[16]

TABLE 1 **University of Michigan Incoming Residents, 2002–2006**

Year	URM No. (%)	Non-URM[a] No. (%)	Unknown	Total
2002	12 (9)	117 (87)	3	**132**
2003	9 (6)	119 (81)	19	**147**
2004	6 (4)	133 (92)	6	**145**
2005	12 (8)	132 (90)	2	**146**
2006	12 (7)	152 (92)	2	**166**
Total	**51** (7)	**653** (89)	**32**	**736**

Abbreviation: URM, underrepresented minority.

[a] The non-URM category includes students who self-identified as white, Asian American, or multiracial.

Methods

The POA was developed by the Graduate Medical Education Committee at the University of Michigan (U of M) in 2002 and focuses on providing assessment and education on the knowledge and skills needed in the unsupervised settings of the first 6 to 18 weeks of training. The examination includes the following stations: informed consent, aseptic technique, evidence-based medicine, images (x-rays), critical laboratory values, cross-cultural communication, geriatric functional assessment/pediatric-proxy history taking (specialty specific), and many Joint Commission requirements, such as surgical fire safety, pain assessment, and management. These stations were designed to cover a wide range of technical, clinical, and communication skills that closely parallel early postgraduate experiences.[17]

Interns receive a mean score (percentage correct) for each POA station. Stations involving a standardized patient (geriatric functional assessment, pediatric proxy history taking, informed consent–blood transfusion, aseptic technique, cross-cultural communication) are scored on the basis of the standardized patient's assessment of performance, using a 3-scale scoring system (done, not done, needs improvement). Computerized stations (reading of radiographic images, evidence-based medicine, critical values, pain assessment, and surgical fire safety) use either multiple-choice or structured written response formats. The individual station scores are then averaged to derive the POA overall score. Residency program directors receive means and standard deviations of scores for all administered years in addition to those of the current class cohort. These reports provide them with formative feedback regarding the

TABLE 2 **United States Medical Licensing Examination (USMLE) Step 1 and Step 2, 2002–2006**

Year	URM, Mean (SD)	Non-URM, Mean (SD)	*t* (d)	*P* Value
USMLE Step 1				
2002	209.6 (19.1)	229.0 (20.4)	3.15 (1.0)	**.002**
2003	207.7 (18.2)	229.0 (19.1)	3.24 (1.1)	**.002**
2004	218.2 (20.5)	231.3 (17.5)	1.79 (0.7)	.076
2005	229.1 (21.5)	231.7 (16.9)	0.47 (0.2)	.636
2006	212.0 (12.7)	230.1 (17.1)	3.44 (1.1)	**<.001**
USMLE Step 2 Clinical Knowledge				
2002	208.1 (17.1)	228.7 (21.2)	2.97 (1.0)	**.004**
2003	206.3 (11.8)	231.6 (20.6)	2.96 (1.3)	**.004**
2004	216.7 (20.3)	234.2 (17.2)	2.41 (1.0)	**.018**
2005	230.8 (24.8)	233.5 (18.4)	0.42 (0.1)	.681
2006	219.2 (21.3)	233.4 (17.1)	2.61 (0.8)	**.010**

Abbreviation: URM, underrepresented minority.
Note: Bolded *P* values represent values <.05.

performance of individual interns and allow them to make within- and across-department comparisons.

The primary data used in this study were a compilation of USMLE Step 1, Step 2 Clinical Knowledge and POA scores from each of the 10 stations. Comparisons were made only in cases in which there had been no change on the assessment, or in cases in which the change was not content-related but simply cosmetic (eg, ordering of questions, test distribution). Information on the type of medical school attended (private, public, or international) was also collected as we anticipated this additional information would provide some insight on resident performance.

Independent sample *t* tests were used for all comparisons between the URM and non-URM groups in TABLES 2 and 3. Analysis of these data was granted exemption by our Institutional Review Board.

Results

Our data set comprised 736 first-year residents in 14 different specialties, representing all new residents who entered our residency programs during 2002–2006. These 736 residents were categorized as follows by self-identification: 318 (43%) were female and 417 (57%) male, and 51 (7.2%) met the 2003 Association of American Medical College classification of underrepresented minority (URM) (TABLE 1). All other students (eg, Non-Hispanic White, Asian American, and multiracial) were categorized as non-URMs. In comparing basic demographic characteristics of the national population of US medical school graduates to those of the University of Michigan population, our sample is slightly underrepresented in terms of women (46% nationally versus 43% in the U of M population), and URMs (14% nationally versus 7.2% U of M population). In our data set of 51 URM residents, 27 (53%) were from medical schools located in states with antiaffirmative action legislation (eg, California, Texas, Florida, Michigan, Georgia, and Washington).[16] We compared those residents who went to schools in antiaffirmative action states to those who did not in terms of USMLE Step 1 and Step 2 Clinical Knowledge scores, in addition to POA overall scores; there was no difference in performance between the 2 groups.

TABLE 2 provides a detailed overview of the USMLE Step 1 and Step 2 scores for our incoming residents for the years 2002–2006. These results indicate that there were significant differences between URM and non-URM residents in that URM residents scored significantly lower than non-URM residents for all years. At the U of M, all interns must pass USMLE Step 2 Clinical Skills before starting clinical work and thus, this pass/fail information was not used in our data set.

TABLE 3 demonstrates that URM and non-URM interns perform similarly on individual POA stations; there was limited evidence for a difference in performance between the URM residents and the other residents. With the exception of 3 occasions in the study period (testing in 2002, 2003, 2006), URM interns performed significantly worse on pain assessment and surgical fire safety (when adjusted for the use of multiple independent *t* tests). Similarly, URM and non-URM interns had similar overall scores on the POA for all years except 2002 (when α is adjusted for multiple *t* tests, not shown).

TABLE 3 **POA Examination Results: URM Residents Compared to Non-URM Residents**

Aseptic Technique

Year	URM, Mean (SD)	Non-URM, Mean (SD)	*t* (d)	*P* Value
2002	87.9 (6.2)	89.2 (9.7)	0.44 (0.1)	.659
2003	76.5 (14.9)	75.8 (13.6)	0.14 (0.1)	.889
2004	72.9 (11.6)	71.6 (12.8)	0.24 (0.1)	.808
2005	73.5 (15.6)	74.6 (13.4)	0.26 (0.1)	.794
2006	64.8 (14.0)	66.4 (13.3)	0.40 (0.1)	.693

Evidence-Based Medicine

Year	URM, Mean (SD)	Non-URM, Mean (SD)	*t* (d)	*P* Value
2002	58.6 (17.7)	70.1 (15.3)	2.33 (0.7)	**.022**
2003	68.3 (19.0)	66.7 (16.2)	0.24 (0.1)	.811
2004	56.0 (13.2)	68.2 (16.1)	1.50 (0.8)	.137
2005	66.3 (7.8)	71.6 (16.9)	0.93 (0.3)	.356
2006	63.5 (23.6)	68.9 (15.7)	0.99 (0.3)	.327

Imaging

Year	URM, Mean (SD)	Non-URM, Mean (SD)	*t* (d)	*P* Value
2002	57.4 (20.6)	67.6 (16.0)	2.05 (0.6)	**.042**
2003 (pediatric imaging)	36.1 (12.7)	44.4 (12.7)	1.08 (0.7)	.289
2003 (adult imaging)	75.0 (13.9)	69.3 (16.2)	0.84 (0.4)	.406
2004	75.9 (15.6)	73.5 (12.8)	0.45 (0.2)	.656
2005	66.7 (9.1)	67.7 (12.2)	0.28 (0.1)	.778
2006	65.0 (13.8)	65.3 (11.2)	0.08 (0.0)	.931

Pain Assessment and Management

Year	URM, Mean (SD)	Non-URM, Mean (SD)	*t* (d)	*P* Value
2002	87.0 (11.8)	89.5 (8.8)	0.93 (0.3)	.353
2003	89.6 (6.2)	88.5 (7.4)	0.45 (0.2)	.654
2004	N/A	N/A	N/A	N/A
2005	77.3 (4.8)	78.1 (7.5)	0.37 (0.1)	.715
2006	73.5 (9.9)	78.7 (7.2)	2.36 (0.7)	**.020**

System Compliance

Year	URM, Mean (SD)	Non-URM, Mean (SD)	*t* (d)	*P* Value
2002	89.2 (5.1)	85.1 (11.2)	1.25 (0.4)	.213
2003	73.3 (26.7)	87.3 (12.1)	2.99 (1.1)	**.003**
2004	96.3 (9.1)	95.6 (9.8)	0.18 (0.7)	.859
2005	89.8 (11.1)	95.1 (8.4)	2.02 (0.6)	**.046**
2006	83.3 (12.1)	78.9 (12.8)	1.15 (0.3)	.253

Abbreviation: URM, underrepresented minority.
Note: Bolded *P* values represent values <.05.
No test results available for that administration = N/A.

An exception to this is in the system compliance station, where the non-URM residents performed significantly better in 1 of the 5 years of administration. When taken together, the results from all other individual stations show no clear differences in performance between the 2 groups, suggesting that the URM residents performed no differently than the non-URM residents on the POA, despite scoring significantly lower on the USMLE Steps 1 and 2.

In particular, there were no discernable differences in performance on stations involving a standardized patient. We had hypothesized that URM students would perform better on the cultural communication and informed consent stations, but this did not prove to be the case.

Discussion

Despite poorer performance on standardized tests and unequal academic credentials, our first-year URM interns performed similarly to their Non-Hispanic White and Asian American counterparts on a single examination of clinical performance. These findings support previous research that suggests that the value of the Medical College Admission Test (MCAT) and undergraduate grade point averages lies in the prediction of preclinical knowledge among medical students, and these tests may not necessarily predict successful performance in residents or practicing physicians.[3,18] These findings are especially critical in light of the recent abandonment of many affirmative action policies that had historically increased the minority physician pipeline.[19]

The POA is an orientation-based OSCE, which may at times uncover serious deficiencies in knowledge or skills, while providing residency program directors with a reliable tool to measure initial clinical skills. Furthermore, in order to limit the impact of these deficits, all participants are provided with remediation materials immediately after the assessment. This ensures instantaneous access to the correct policies, procedures, and clinical information imperative to engaging in safe and effective medical practice, as participants will soon find themselves in very similar settings.

We speculate that lower overall scores of URM residents on the POA in 2002 may be the result of the timing of the examination. The POA was first introduced in 2002, which was 1 year before the implementation of USMLE Step 2 Clinical Skills. At this time, only a limited number of schools had a final comprehensive clinical skills assessment required for graduation.[20] Although it remains unclear whether our URM residents hailed from such schools, after the Step 2 Clinical Skills was implemented URM residents' overall POA scores reached levels similar to those of the non-URM group. As previously stated, we developed the POA to address clinical issues seen in the first several weeks to months as a new physician. It is a test of practical clinical skill that may have more practical translation to actual bedside behaviors. The clinical skills assessed during the USMLE clinical skills examination and the POA may measure different aspects of clinical acumen than those assessed in traditional standardized tests (eg, Step 1 and Step 2 Clinical Knowledge). It is also possible that our URM resident selection process improved in the years subsequent to 2002. Additionally, the consequences of stereotype threat among undergraduate and medical students may in fact contribute to poorer performance on standardized tests such as Step 1 and Step 2 Clinical Knowledge. URM resident performance on the POA may not be affected by this, owing to the lack of any preconceived notions on performance for this test.[21,22]

We originally hypothesized that the URM residents would perform better on some of the standardized patient stations, most notably the cross-cultural communication and informed consent stations. This did not bear out, and in fact the findings support the argument that all physicians need training in the area of cultural competency and health disparities.

We must note the limitations of this study. First, the sample includes a relatively small sampling of URM residents at 1 institution; however, on average, we recruit from 60 different medical schools annually. We fully recognize that the small number of URM interns limits the power of the study in its ability to generalize to larger populations. Also, the residency programs at the University of Michigan consider themselves to be highly selective, which may place them at an advantage when selecting the most qualified URM students in the country; however, this would be true of all our residents. These factors, taken either alone or together, may also limit the generalizability of our findings. One could argue that the POA has limited a priori data regarding construct validation, which may increase the likelihood of nonsignificant group differences. Finally, the combined educational and assessment stations (eg, fire and pain assessment) could present a new modality experience for some incoming interns who typically may have had difficulty with a time-limited, traditional, multiple-choice framework. This testing format may reinforce difficulties URM students have in navigating single best-answer test items.[23]

Conclusion

When compared to non-URM residents from a variety of medical schools, incoming URM residents perform as well as other trainees. These results highlight the multidimensional skill set involved in obtaining sufficient medical skills (eg, knowledge, clinical skills, communication) and will hopefully encourage more schools to broaden their view of URM applicants.

References

1 American Association of Medical Colleges. Underrepresented in medicine definition. Available at: http://www.aamc.org/meded/urm/start.htm. Accessed June 19, 2008.

2 Cohen JJ. The consequences of premature abandonment of affirmative action in medical school admissions. *JAMA*. 2003;289(9):1143–1149.
3 Gonzalez P, Stoll B. *The Color of Medicine: Strategies for Increasing Diversity in the U.S. Physician Workforce*. Boston, MA: Community Catalyst; 2002.
4 Reede JY. Predictors of success in medicine. *Clin Orthop Relat Res*. 1999;362:72–77.
5 Edmond MB, Deschenes JL, Eckler M, Wenzel RP. Racial bias in using USMLE Step 1 scores to grant internal medicine residency interviews. *Acad Med*. 2001;76(12)1253–1256.
6 Tekian A, Han Y, Hruska L, Krainik AJ. Do underrepresented minority medical students differ from nonminority students in problem-solving ability? *Teach Learn Med*. 2001;13(2):86–91.
7 Bernstein AD, Jazrawi LM, Elbeshbeshy B, Della Valle CJ, Zuckerman JD. Orthopaedic resident-selection criteria. *J Bone Joint Surg*. 2002;84A(11):2090–2096.
8 Metro DG, Talarico JF, Patel RM, Wetmore AL. The resident application process and its correlation to future performance as a resident. *Anesth Analg*. 2005;100:502–505.
9 Julian ER. Validity of the medical college admission test for predicting medical school performance. *Acad Med*. 2005;80(10):910–917.
10 Reiter H, Maccoon K. A compromise method to facilitate under-represented minority admissions to medical school. *Adv Health Sci Educ Theory Pract*. 2007;12(2):223–237.
11 Duffy FA, Gordon GH, Whelan G, Cole-Kelly K, Frankel R. Assessing competence in communication and interpersonal skills: The Kalamazoo II report. *Acad Med*. 2004;79(6):495–507.
12 Yudkowsky R, Downing SM, Sandlow LJ. Developing an institution-based assessment of resident communication and interpersonal skills. *Acad Med*. 2006;81(12):1115–1122.
13 Tetzlaff JE. Assessment of competency in anesthesiology. *Anesthesiology*. 2007;106(4): 812–825.
14 Lypson ML, Frohna JG, Gruppen LD, Woolliscroft JO. Assessing residents' competencies at baseline: identifying the gaps. *Acad Med*. 2004;79(6):564–570.
15 Accreditation Council for Graduate Medical Education. Common program requirements: general competencies. Available at: http://www.acgme.org/outcome/comp/GeneralCompetenciesStandards21307.pdf. Accessed June 11, 2008.
16 Long MC, Tienda M. Winners and losers: changes in Texas University admissions post-Hopwood. *Educ Eval Policy Anal*. 2008;30(3):255–280.
17 Cerilli GJ, Merrick HW, Staren ED. Objective structured clinical examination technical skill stations correlate more closely with postgraduate year level than do clinical skill stations. *Am Surg*. 2001;67(4):326–327.
18 Petersdorf RG, Turner KS, Nickens HW, Ready T. Minorities in medicine: past, present, and future. *Acad Med*. 1990;65(11):663–670.
19 Carlisle DM, Gardner JE, Liu H. The entry of underrepresented minority students into US medical schools: an evaluation of recent trends. *Am J Public Health*. 1998;88(9):1314–1318.
20 Barzansky B, Etzel SI. Educational programs in US medical schools, 2002-2003. *JAMA*. 2003;290(9):1190–1196.
21 Woolf K, Haq I, McManus IC, Higham J, Dacre J. Exploring the underperformance of male and minority ethnic medical students in first year clinical examinations. *Adv Health Sci Educ Theory Pract*. 2008;13(5):607–616.
22 Burgess DJ, Warren J, Phelan S, Dovidio J, van Ryn M. Stereotype threat and health disparities: what medical educators and future physicians need to know. *J Gen Intern Med*. 2010;25(suppl 2):169–177.
23 Frierson HT Jr. Impact of an intervention program on minority medical students' National Board Part I performance. *J Natl Med Assoc*. 1984;76(12):1185–1190.