



Objective Test Scores Throughout Orthopedic Surgery Residency Suggest Disparities in Training Experience

Norah Foster, MD,^{‡,1} Meghan Price, BS,^{†,1} Janet Prvu Bettger, ScD,^{*} C. Rory Goodwin, MD, PhD,[†] and Melissa Erickson, MD^{*}

^{*}Department of Orthopedic Surgery, Division of Spine, Duke University Medical Center, Durham, North Carolina;

[†]Department of Neurosurgery, Division of Spine, Duke University Medical Center, Durham, North Carolina; and

[‡]Premier Health Orthopedics, Centerville, Ohio

Diversifying clinical residencies, particularly in fields that are historically dominated by majority male (M/M) cohorts, is critical to improve both the training experiences of residents and the overall physician workforce. Orthopedic surgery in particular has low numbers of females and under-represented minorities (F/URM) at all levels of training and practice. Despite efforts to increase its diversity, this field has become more homogeneous in recent years. To highlight potential barriers and disparate training environments that may contribute to this dynamic, we present 25 years' worth of institutional data on standardized exam performance throughout residency. We report that despite starting residency with standardized exam scores that were comparable to their M/M peers, F/URM orthopedic surgery residents performed progressively worse on Orthopaedic In-service Training Exams throughout residency and had lower first pass rates on the American Board of Orthopedic Surgery Part 1. Given these findings, we propose that disparate performance on standardized test scores throughout residency could identify trainees that may have different experiences that negatively impact their exam performance. Shedding light on these underlying disparities provides opportunities to find meaningful and sustained ways to develop a culture of diversity and inclusion. It may also allow for other programs to identify similar patterns within their training programs. Overall, we propose monitoring test performance on standardized exams throughout orthopedic surgery residency to identify potential disparities in training experience; further, we acknowledge that interventions to mitigate these disparities require a broad, systems wide approach and a

firm institutional commitment to reducing bias and working toward sustainable change. (J Surg Ed 78:1400–1405. © 2021 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: residency, diversity, in training exam scores, orthopedic surgery, ABOS, females, under-represented minorities

COMPETENCIES: Professionalism, Interpersonal and Communication Skills, Practice-Based Learning and Improvement

INTRODUCTION

Diversifying clinical residency programs in the United States is imperative to ensure provision of both effective care to an increasingly diverse society and wholistic training experiences to young physicians.¹ Despite several recent initiatives to increase its diversity,² orthopedic surgery continues to lag behind its medical and surgical peers with respect to representation of minorities and women in the field.^{3,4}

There is evidence to suggest that once in residency, females and underrepresented minorities (F/URM) do not have the same experience as their majority male (M/M) peers.⁵ Specifically, F/URM receive fewer opportunities to be mentored and are often subjected to implicit biases that can adversely affect feedback and assessment.^{6,7} More than half of women surveyed about their experiences in academic surgery felt that gender played a role in the career challenges and training inequities they faced.⁵ While there is an increasingly better understanding of disparate training environments and the

Correspondence: Inquiries to Melissa Erickson, MD, Department of Orthopedic Surgery, Division of Spine, Duke University Medical Center, 40 Duke Medicine Circle Clinic, Durham, NC 27710; e-mail: Melissa.erickson@duke.edu

¹ Denotes co-first authorship.

experiences of F/URM residents, there are little data on how, if at all, these disparities impact the objective outcomes by which residents are currently assessed.

Many metrics can be used to evaluate resident proficiency, performance, and training. Utilizing our institutional experience, we propose the use of trends in standardized test scores throughout residency as a “litmus test” to identify residents who not only may be struggling with standardized exams but may also be experiencing inherent disparities that affect their ability to perform on these exams. Twenty-five years of data collected at our institution from resident applications, orthopedic in-service training exams (OITE) scores and American Board of Orthopaedic Surgery (ABOS) Part 1 first-time pass rates suggest that F/URM orthopedic surgery residents may experience differences in their training environments compared to those of their M/M peers (Fig. 1). We examined the performance of all orthopedic surgery residents from our institution with graduation years of 1994 to 2018 (Table 1). We found no difference in objective test performance of F/URMs prior to starting residency (Table 2). However, despite initially comparable scores, there was a decline in objective performance by F/URM residents during residency on the relative percentile score of OITE and after residency on the ABOS Part 1 certifying examination (Fig. 1; Table 2).

Declining relative performance in ITE percentile rankings across years of training is not unique to orthopedics⁸; however, the OITE percentile attrition we observed in F/URM residents was specific to the field.^{9,10} In fact in other areas of medicine, there appears to be no differences in objective measures of medical knowledge between sex and race throughout residency training.⁹

Therefore, while other factors result in declining scores, the divergent patterns observed between M/M and F/URM residents suggest that there could be inherent differences in the training experiences of F/URM residents in orthopedic surgery. Given our observations, we believe it is critical that programs study trends in their own data and intervene when possible to offset negative training experiences and lower attrition rates.

WORKFORCE DIVERSITY

Despite evidence that both physicians and patients generally benefit from a more diverse provider workforce,¹¹⁻¹⁵ both women and URMs report lower levels of career success and higher attrition rates both during and after residency.¹⁶ While the causes of this are likely multifactorial, the higher rates of burnout and perceived sexual discrimination^{17,18} likely play a role for female residents. These factors are exacerbated by a gross under-representation of female surgeons which limits mentorship opportunities that could impact some of these negative experiences.^{19,20} Similar challenges are documented for URM representation at the student, trainee, and professional level in academic surgery.²¹ Recent US Census data for orthopedic surgery indicates that orthopedic surgery is the least diverse surgical subspecialty with respect to race and confirms that diversity is not increasing.²² Furthermore, the experiences of URM in orthopedic programs are not always positive. Fourteen percent of URM orthopedic residents perceived racial harassment, and 8% felt they had experienced racial discrimination. Both female and URM orthopedic residents reported a lowered sense of personal accomplishment.²³

OPPORTUNITIES TO INTERVENE

While there is skepticism regarding the correlation of performance on standardized test scores to clinical acumen, there is significant evidence supporting the correlation of ABOS pass rates with various other standardized measures including OITE scores and USMLE scores prior to residency.²⁴⁻²⁶ Thus, there is an emphasis on these metrics in both selecting and training residents.²⁷ It is also important to note that the imminent change in USMLE Step 1 grading to Pass/Fail may result in programs relying more heavily on other objective metrics such as first year OITE scores at the beginning of residency. Secondly, objective measures of performance may serve as an opportunity to measure effectiveness of interventions, within and across programs, aimed to

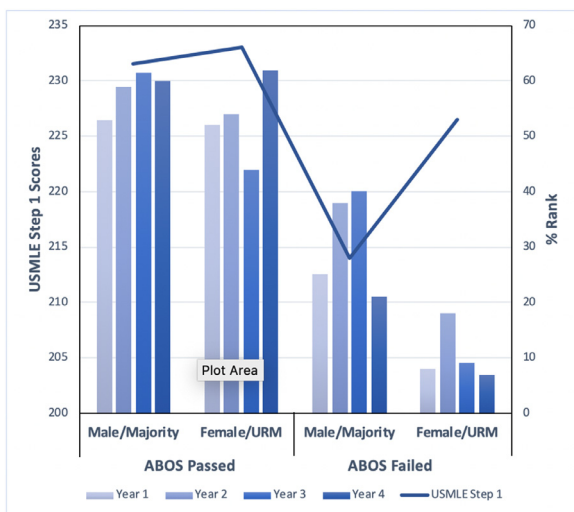


FIGURE 1. USMLE1 (line) and OITE percentile rank by year (bars) for male/majority and female/underrepresented minority who did and did not pass ABOS exams.

TABLE 1. Population Characteristics Overall and for the Male/Majority Compared With Female/Underrepresented Minorities Prior to Orthopaedic Residency

Resident Characteristics	Overall, n = 197	Male/Majority, n = 163	Female/URM, n = 34	p Value
Age at program start	27.1 (26.4-28.5)	27.6 (26.5-28.6)	27.4 (26.3-27.8)	0.62
Undergrad STEM degree	170 (86.3%)	139 (85.3%)	31 (91.2%)	0.52
Collegiate athlete	93 (47.2%)	75 (46.0%)	18 (52.9%)	0.58
Armed forces	7 (3.6%)	9 (5.5%)	0	0.36
Prior graduate degree	24 (12.2%)	20 (12.3%)	4 (11.8%)	0.93
Private medical school	112 (56.9%)	102 (62.6%)	10 (29.4%)	<0.01
AOA status	84 (42.6%)	72 (44.2%)	12 (35.3%)	0.58

address improved equity and culture of orthopedic residency programs.

The means to identify and address potential causes of worsening objective performance on OITE and lower first pass rates on ABOS part 1 for F/URM residents are multifold. Beyond evidence-based interventions focused on study strategies and knowledge acquisition,^{17,18} it is important to consider the residency training environment for F/URMs. Factors known to exacerbate disadvantages for F/URMs include stereotype threats, implicit biases, and fewer opportunities for feedback, higher level training experiences, and effective mentorship.²⁸ These factors can interfere with learning,^{29,30} performance on assessments, and well-being.^{28,31-33} These negative impacts are often heightened when learners are more aware of how they differ from those around them.^{34,35} This is particularly relevant for female and URM orthopedic surgery residents who have reported awareness of how their sex and/or race affects both the residency selection process³⁶ and experience in residency.³⁷

Implicit biases can also negatively impact the everyday experience, feedback and training milestone acquisitions

for residents.^{7,38} Although various initiatives in medical training are effective in reducing implicit bias,^{6,39} there is increasing evidence that these strategies may result in short-lived rather than sustained change.⁴⁰ For example, providing female surgical residents with positive affirmation to combat the potential of stereotype threat successfully prevented a decline in performance ratings compared with residents who did not receive this intervention.³⁵ Similar interventions implemented at various levels of academic learning for females and URMs have shown to improve both performance and testing scores.⁴¹ However, despite these initial positive results, it is important to acknowledge that they are short-term benefits and their sustainability has not been appropriately assessed. While these strategies are a start, time and sustained effort will be required to definitively change deeply entrenched disparities.

Unlike identification of a need for remediation in testing strategies, ways to address differences in training experiences that may be fueled by inherent biases are more amorphous and thus are more daunting to address. However, it is clear that it is time to seriously take on these challenges at all levels. Mitigating disparities

TABLE 2. USMLE I, OITE Ranks, and ABOS Performance Overall, and for the Male/Majority Compared With Female/Underrepresented Minority Residents

Resident Objective Outcomes	Overall, n = 197	Male/Majority, n = 163	Female/URM, n = 34	Mean Difference (95% CI), p Value
USMLE	232.3 [232 (219-242)]	230.3 [231 (219-242)]	241.9 [235 (217-242)]	-11.13 (-37.0 to 14.73), 0.39
OITE				
Year 1 % rank	50.4 [50 (30-72)]	51.6 [51 (31-72.5)]	44.9 [45.5 (18.8-68.2)]	10.6 (-5.0 to 26.3), 0.18
Year 2 % rank	55.8 [57 (36-75)]	57.0 [59 (38.5-75.5)]	50.1 [48.5 (27.2-70)]	3.0 (-10.6 to 16.7), 0.65
Year 3 % rank	53.9 [56 (31-79)]	55.7 [60 (33-79)]	45.4 [43.5 (20.8-72.2)]	10.21 (-4.5 to 24.9), 0.17
Year 4 % rank	55.8 [57 (33.8-79)]	56.8 [58 (34-78.8)]	51.0 [45.5 (27.2-78.8)]	1.62 (-12.9 to 16.2), 0.82
ABOS boards pass (%)	186 (94.4)	157 (96.3)	29 (85.3)	$\chi^2 = 4.6, p = 0.03$

TABLE 3. Quick Recommendations for Strategies to Increase Support for Females and Under-Represented Minorities Throughout Orthopedic Surgery Training

Quick Recommendations for Improvement

Mentorship	<ul style="list-style-type: none">• Assign all residents a mentorship team comprised of diverse faculty starting at the beginning of residency• Establish formalized metrics for the engagement and mentorship expected of this team (goal setting, number of expected meetings)• Set up inter-resident mentoring opportunities• For institutions with low numbers of diverse faculty members, provide residents resources to connect with more diverse colleagues in the field who would be willing to provide mentorship at other institutions
Educational planning/coaching	<ul style="list-style-type: none">• Mentorship teams will be responsible for helping review residency expectations and assess whether resident is meeting them appropriately• Help with career planning, study resources, research planning• Regularly review the performance of residents on standardized exams and other metrics with a plan to intervene if a resident is struggling or not meeting expected metrics
Cultural shifts	<ul style="list-style-type: none">• Set top-down expectations that faculty and residents will not tolerate overtly discriminatory behavior and other forms of bias• Have an anonymous, effective reporting system for transgressions with clear, consistent follow up
Recruitment	<ul style="list-style-type: none">• Incorporate a framework for inclusive GME recruitment strategies (Gonzaga et al 2020)

within medical training requires a systems wide approach and a firm commitment to both reducing bias and achieving enduring change. Specific examples of sustained culture change include effectively addressing stereotype threat and implicit bias as potential obstacles for certain residents, engaging faculty and learners at all levels in active awareness of potential discrimination, and building stronger pathways for leadership, professional, and mentoring advancements for diverse F/URM (Table 3).^{32,42,43} For these initiatives to be successful, it is critical to start with an honest introspective review to identify program-specific factors that may be impeding improvements and change.

CONCLUSION

Gender and racial diversity in orthopedic residencies remain low. Significant work is required to better understand why certain residents may have disparate performance as they progress through training. Our experience is meant to raise awareness of a pattern that could exist in other training programs and the need to examine trends that can be shifted to improve diversity, equity and inclusion. While we propose monitoring test performance as one approach to identifying potential disparities and opportunities to intervene, more research is required to address how best to identify and mitigate these disparities. Direct attention is needed to promote positive residency experiences and prevent future attrition

of F/URM residents from fields such as orthopedic surgery that would greatly benefit from their presence.

REFERENCES

1. Blakemore LC, Hall JM, Biermann JS. Women in surgical residency training programs. *J Bone Joint Surg Am.* 2003;85:2477-2480.
2. Van Heest AE, Fishman F, Agel J. A 5-year update on the uneven distribution of women in orthopaedic surgery residency training programs in the United States. *J Bone Joint Surg Am.* 2016;98:e64.
3. Day CS, Lage DE, Ahn CS. Diversity based on race, ethnicity, and sex between academic orthopaedic surgery and other specialties: a comparative study. *J Bone Joint Surg Am.* 2010;92:2328-2335.
4. Adelani MA, Harrington MA, CO Montgomery. The distribution of underrepresented minorities in U.S. Orthopaedic Surgery Residency Programs. *J Bone Joint Surg Am.* 2019;101:e96.
5. Seemann NM, Webster F, Holden HA, et al. Women in academic surgery: why is the playing field still not level? *Am J Surg.* 2016;211:343-349.
6. Chadwick AJ, Baruah R. Gender disparity and implicit gender bias amongst doctors in intensive

care medicine: a “disease” we need to recognise and treat. *J Intensive Care Soc.* 2020;21:12–17.

7. Klein R, Julian KA, Snyder ED, et al. Gender bias in resident assessment in graduate medical education: review of the literature. *J Gen Intern Med.* 2019;34:712–719.
8. Garibaldi RA, Subhiyah R, Moore ME, Waxman H. The in-training examination in internal medicine: an analysis of resident performance over time. *Ann Intern Med.* 2002;137:505–510.
9. Sulistio MS, Khera A, Squiers K, et al. Effects of gender in resident evaluations and certifying examination pass rates. *BMC Med Educ.* 2019;19:10.
10. Pico K, Gioe TJ, Vanheest A, Tatman PJ. Do men outperform women during orthopaedic residency training? *Clin Orthop Relat Res.* 2010;468:1804–1808.
11. Saha S, Guiton G, Wimmers PF, Wilkerson L. Student body racial and ethnic composition and diversity-related outcomes in US medical schools. *JAMA.* 2008;300:1135–1145.
12. Cooper-Patrick L, Gallo JJ, Gonzales JJ, et al. Race, gender, and partnership in the patient-physician relationship. *JAMA.* 1999;282:583–589.
13. Butler PD, Longaker MT, Britt LD. Major deficit in the number of underrepresented minority academic surgeons persists. *Ann Surg.* 2008;248:704–711.
14. Jimenez RL. Barriers to minorities in the orthopaedic profession. *Clin Orthop Relat Res.* 1999; 44–50.
15. Tsugawa Y, Jena AB, Figueroa JF, Orav EJ, Blumenthal DM, Jha AK. Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. *JAMA Intern Med.* 2017;177:206–213.
16. Gifford E, Galante J, Kaji AH, et al. Factors associated with general surgery residents’ desire to leave residency programs: a multi-institutional study. *JAMA Surg.* 2014;149:948–953.
17. 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;36:E185–E188.
18. Drake SM, Qureshi W, Morse W, Baker-Genaw K. A time-efficient web-based teaching tool to improve medical knowledge and decrease ABIM failure rate in select residents. *Med Educ Online.* 2015;20:29221.
19. Zhuge Y, Kaufman J, Simeone DM, Chen H, Velazquez OC. Is there still a glass ceiling for women in academic surgery? *Ann Surg.* 2011;253:637–643.
20. Neumayer L, Kaiser S, Anderson K, et al. Perceptions of women medical students and their influence on career choice. *Am J Surg.* 2002;183:146–150.
21. Abelson JS, Symer MM, Yeo HL, et al. Surgical time out: our counts are still short on racial diversity in academic surgery. *Am J Surg.* 2018;215:542–548.
22. Ramirez RN, Franklin CC. Racial diversity in orthopedic surgery. *Orthop Clin North Am.* 2019;50:337–344.
23. Sargent MC, Sotile W, Sotile MO, et al. Managing stress in the orthopaedic family: avoiding burnout, achieving resilience. *J Bone Joint Surg Am.* 2011;93:e40.
24. Turner NS, Shaughnessy WJ, Berg EJ, Larson DR, Hanssen AD. A quantitative composite scoring tool for orthopaedic residency screening and selection. *Clin Orthop Relat Res.* 2006;449:50–55.
25. 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;474:908–914.
26. Spitzer AB, Gage MJ, Looze CA, Walsh M, Zuckerman JD, Egol KA. Factors associated with successful performance in an orthopaedic surgery residency. *J Bone Joint Surg Am.* 2009;91:2750–2755.
27. Schrock JB, Kraeutler MJ, Dayton MR, McCarty EC. A Cross-sectional analysis of minimum USMLE step 1 and 2 criteria used by orthopaedic surgery residency programs in screening residency applications. *J Am Acad Orthop Surg.* 2017;25:464–468.
28. Burgess DJ, Joseph A, van Ryn M, Carnes M. Does stereotype threat affect women in academic medicine. *Acad Med.* 2012;87:506–512.
29. Lievens F, Sackett PR. The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *J Appl Psychol.* 2012;97:460–468.
30. Schmader T, Johns M, Forbes C. An integrated process model of stereotype threat effects on performance. *Psychol Rev.* 2008;115:336–356.
31. Milam LA, Cohen GL, Mueller C, Salles A. Stereotype threat and working memory among surgical residents. *Am J Surg.* 2018;216:824–829.
32. Burgess D. Can a brief values affirmation help achieve gender parity within the surgical

- profession? The promise and limitations of “wise” interventions. *J Grad Med Educ*. 2016;8:439–441.
- 33.** Gomez E, Wright JG. A simple strategy to reduce stereotype threat for orthopedic residents. *Can J Surg*. 2014;57:E19–E24.
- 34.** Fassiotto M, Hamel EO, Ku M, et al. Women in academic medicine: measuring stereotype threat among junior faculty. *J Womens Health (Larchmt)*. 2016;25:292–298.
- 35.** Salles A, Mueller CM, Cohen GL. Exploring the relationship between stereotype perception and residents’ well-being. *J Am Coll Surg*. 2016;222:52–58.
- 36.** Bajaj G, Carmichael KD. What attributes are necessary to be selected for an orthopaedic surgery residency position: perceptions of faculty and residents. *South Med J*. 2004;97:1179–1185.
- 37.** Zern NK, Shalhub S, Wood DE, Calhoun KE. Association of sex with perceived career barriers among surgeons. *JAMA Surg*. 2019;154:1155–1158.
- 38.** Hansen M, Schoonover A, Skarica B, Harrod T, Bahr N, Guise JM. Implicit gender bias among US resident physicians. *BMC Med Educ*. 2019;19:396.
- 39.** Slattengren AH, Prasad S, Bury DC, et al. Erratum: a better approach to the diagnosis of PE. *J Fam Pract*. 2019;68:340.
- 40.** Devine PG, Forscher PS, Austin AJ, Cox WT. Long-term reduction in implicit race bias: a prejudice habit-breaking intervention. *J Exp Soc Psychol*. 2012;48:1267–1278.
- 41.** Cohen GL, Garcia J, Purdie-Vaughns V, Apfel N, Brzustoski P. Recursive processes in self-affirmation: intervening to close the minority achievement gap. *Science*. 2009;324:400–403.
- 42.** Okike K, Utuk ME, White AA. Racial and ethnic diversity in orthopaedic surgery residency programs. *J Bone Joint Surg Am*. 2011;93:e107.
- 43.** Gonzaga AMR, Appiah-Pippim J, Onumah CM, Yialamas MA. A framework for inclusive graduate medical education recruitment strategies: meeting the ACGME standard for a diverse and inclusive workforce. *Acad Med*. 2020;95:710–716.