



Published in final edited form as:

*Acad Med.* 2021 February 01; 96(2): 176–181. doi:10.1097/ACM.0000000000003837.

## Admissions Is Not Enough: The Racial Achievement Gap in Medical Education

**Alana C. Jones [MD-PhD trainee],**

Department of Epidemiology, University of Alabama at Birmingham School of Medicine, Birmingham, Alabama

**Alana C. Nichols, JD, MD [recent graduate],**

University of Alabama at Birmingham School of Medicine, Birmingham, Alabama.

**Carmel M. McNicholas, PhD [associate professor],**

Department of Cell, Developmental, and Integrative Biology, University of Alabama at Birmingham School of Medicine, Birmingham, Alabama

**Fatima C. Stanford, MD, MPH, MPA [assistant professor]**

Neuroendocrine Unit, Pediatric Endocrinology, Internal Medicine, and obesity medicine physician scientist, Nutrition Obesity Research Center at Harvard, Mass General Weight Center, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts

### Abstract

The achievement gap is a disparity in academic and standardized test performance that exists between White and underrepresented minority (URM) students that begins as early as preschool and worsens as students progress through the educational system. Medical education is not immune to this inequality. URM medical students are more likely to experience delayed graduation and course failure, even after accounting for science grade point average and Medical College Admission Test performance. Moreover, URM students are more likely to earn lower scores on licensing examinations, which can have a significant impact on their career trajectory, including specialty choice and residency competitiveness. After the release of preliminary recommendations from the Invitational Conference on USMLE Scoring (InCUS) and public commentary on these recommendations, the National Board of Medical Examiners and Federation of State Medical Boards announced that the United States Medical Licensing Examination (USMLE) Step 1 would transition from a 3-digit numeric score to pass/fail scoring. Given that another of InCUS's recommendations was to “minimize racial demographic differences that exist in USMLE performance,” it is paramount to consider the impact of this scoring change on URM medical students specifically. Holistic admissions are a step in the right direction of

---

Correspondence should be addressed to Alana C. Jones, Medical Scientist Training Program, University of Alabama at Birmingham, 1825 University Blvd., SHEL 121, Birmingham, AL 35294-2182; telephone: (205) 934-4092; acjones@uab.edu; Twitter: @alanacjones\_.

The authors have informed the journal that they agree that both Carmel M. McNicholas and Fatima C. Stanford completed the intellectual and other work typical of the last author.

*Other disclosures:* None reported.

*Ethical approval:* Reported as not applicable.

*Disclaimers:* The content is the sole responsibility of the authors and does not necessarily reflect the views of the NIH or AAUW.

acknowledging that URM students often travel a further distance to reach medical school. However, when residency programs emphasize USMLE performance (or any standardized test score) despite persistent test score gaps, medical education contributes to the disproportionate harm URM students face and bolsters segregation across medical specialties. This Perspective provides a brief explanation of the achievement gap, its psychological consequences, and its consequences in medical education; discusses the potential effect of the Step 1 scoring change on URM medical students; and provides a review of strategies to redress this disparity.

---

In February 2020, the cosponsors of the United States Medical Licensing Examination (USMLE) Step 1—the National Board of Medical Examiners (NBME) and Federation of State Medical Boards (FSMB)—announced that the exam would transition from reporting a numeric score to pass/fail scoring. The NBME and FSMB cited secondary uses of the 3-digit score (i.e., the overemphasis on this score and its use in screening applicants for residency interviews) as one of the reasons for this change.<sup>1</sup> This decision came about a year after the preliminary recommendations from the Invitational Conference on USMLE Scoring (InCUS) were released.<sup>2</sup> These recommendations, especially the recommendation to consider moving Step 1 to pass/fail scoring, resulted in significant feedback from stakeholders, ranging from current trainees to residency program directors to the Association of American Medical Colleges.<sup>3–6</sup> The recent Step 1 announcement has garnered a similar response; of particular note is the concern that with this change, the overemphasis that was put on Step 1 scores will simply transfer to Step 2 Clinical Knowledge (CK) exam scores.<sup>7–9</sup> However, for all of the commentary on the consideration of pass/fail scoring and its consequences, there has been little discussion on the InCUS recommendation to “minimize racial demographic differences that exist in USMLE performance”<sup>2</sup> or on the specific impact of this scoring change on underrepresented minority (URM) medical students.

Recently, in a commentary discussing this very issue, McDade and colleagues acknowledged that pass/fail scoring does not solve the problem that leads to USMLE score differences between URM and White medical students.<sup>10</sup> Unless medical education undertakes proactive interventions, future residency selection systems risk reinforcing the same disparities that currently place URM medical students at a particular disadvantage in competitiveness for residency, even with the removal of Step 1 scores as a screening tool. In this Perspective, we provide a brief explanation of these racial demographic differences, also known as the achievement gap, their psychological consequences, and their consequences in medical education; discuss the potential effect of the Step 1 scoring change on URM students; and provide a review of strategies to redress the achievement gap.

## A Brief Explanation of the Achievement Gap

The achievement gap is a disparity in academic and standardized test performance that exists between White and URM students that begins as early as preschool and worsens as students progress through the educational system.<sup>11</sup> The gap is especially pronounced in mathematics and science.<sup>12</sup> Even though it has been about 65 years since the landmark *Brown v Board of Education* decision, public schools remain highly segregated across the country, and in some states, school segregation has actually worsened since the 1960s.<sup>13</sup> The expanding

achievement gap, driven in part by persistent residential segregation that impacts school funding, results in fewer academic resources for Black and Latinx children.<sup>14</sup> For example, these funding disparities result in higher student-teacher ratios and fewer science and mathematics college preparatory courses in predominantly Black and Latinx schools.<sup>15-18</sup> The unequal quality of K-12 education then results in lower college admissions test scores and lower likelihood of obtaining an undergraduate degree in science, technology, engineering, or mathematics.<sup>19</sup> In many ways, the achievement gap is a resource gap, but these racial disparities persist even in high-quality, well-resourced schools.<sup>20</sup>

## Psychological Consequences of the Achievement Gap

The achievement gap also carries a psychological impact that can further exacerbate the existing disparity. Stereotype threat, the fear that an individual's performance will be judged in light of negative stereotypes about their social group (e.g., race or gender category) can actually worsen academic performance.<sup>21</sup> Thus, as students begin to internalize negative messages that their social group is academically inferior, their performance declines.<sup>22</sup> In a 2011 study, Owens and Massey examined associations between URM student academic performance and externalization and internalization psychological pathways among ~4,000 college freshmen.<sup>23</sup> The externalization pathway was defined as URMs believing that White students perceived them as less intelligent, whereas the internalization pathway was defined as URMs themselves believing that they were less intelligent than White students. The study showed significant relationships between academic underperformance and both internalization and externalization of negative stereotypes.

Moreover, the impact of stereotype threat also extends to standardized test performance. In low threat test conditions, participants were told they were completing a "series of puzzles," and Black undergraduate students performed just as well their White classmates on this cognitive ability test.<sup>24</sup> However, in standard and high threat conditions, in which students were told they were completing an IQ test, Black students earned lower scores on the same exam. Another study of female students in science, technology, engineering, or mathematics undergraduate programs in Germany showed that internalization of gender stereotypes negatively impacted students' self-concept even though they earned good grades in their courses.<sup>25</sup> The consequences of stereotype threat on coursework and standardized test performance are far-reaching, and students with intersecting oppressions (e.g., Black women) may be especially at risk.

While both stereotype threat and impostor syndrome relate to negative stereotypes, impostor syndrome is defined as an emergent identity that represents how individuals view themselves.<sup>26</sup> In a study on impostorism among third-year medical students, (1) women experienced more severe impostor syndrome than men; (2) impostor syndrome significantly correlated with perceived stress; and (3) impostorism was associated with lower Step 1 scores among males.<sup>27</sup> As it relates to URM students' mental health in secondary and postsecondary education, McClain and colleagues observed that both impostor syndrome and "minority status stress" were negative independent predictors of psychological well-being among Black students at a predominantly White university.<sup>26</sup> More research is needed in this area, but these findings suggest that URM medical students may also suffer from

additional psychological effects, such as stereotype threat and impostor syndrome, that exacerbate the achievement gap. Considering the high-stakes nature of medical school, in which both URM and non-URM students may experience impostor syndrome, it is important to acknowledge and address how psychosocial factors also contribute to academic performance.

## Consequences of the Achievement Gap in Medical Education

Medical education is not immune to the achievement gap. Lucey and Saguil have previously described in detail how structural racism impacts Medical College Admission Test (MCAT) score differences across racial groups, and thus disproportionately disadvantages URM applicants.<sup>28</sup> Although some students manage to achieve admission to medical school despite society-level disadvantages, matriculation does not erase the effects of decades of educational inequity. URM students are more likely than their White and Asian classmates to arrive to medical school with lower MCAT scores and science grade point averages.<sup>29</sup> Black and Latinx medical students are also more likely to receive lower scores on and/or fail licensing exams compared to their White counterparts. In an analysis performed by the NBME, Black, Hispanic, and Asian students were significantly more likely to receive lower Step 1, Step 2 CK, and Step 3 scores than White students.<sup>30</sup> These associations persisted even after accounting for science grade point average, MCAT score, and previous USMLE performance. Further, URM students make up nearly half of all medical students who initially fail the Step 1 exam.<sup>31</sup>

This trend has persisted for decades, as a study of medical students in 1992 showed that URM status independently predicted academic difficulty, including delayed graduation and course failure even after accounting for science grade point average and MCAT score.<sup>32</sup> Moreover, URM students are more likely to receive lower clerkship grades and are less likely to be inducted into the Alpha Omega Alpha academic honor society than their non-URM peers.<sup>33,34</sup> Clerkship grades are primarily based on shelf exam scores—which have score gaps that are similar to those seen on the USMLE exams—and narrative evaluations, which have been shown to be influenced by evaluator bias.<sup>34,35</sup> These gaps must be understood in the context of unequal resources throughout primary and secondary education. If, because of the structural consequences of racism, URM students are more likely to start from behind their peers in elementary school and never gain adequate resources to catch up (in addition to facing implicit and explicit bias in subjective evaluations<sup>34</sup>), then their worse academic performance at subsequent stages of education should not be surprising.

## Potential Effect of Step 1 Transition to Pass/Fail on URM Medical Students

While the overall effects of a pass/fail grading system for Step 1 seem promising, it is important to consider that, for URM medical students, this decision may actually have unforeseen negative impacts on their medical education. McDade and colleagues recently authored a commentary with compelling opinions on the potential positive overall outcomes and alluded to some of these potential negative impacts for URM students, such as simply shifting the focus from Step 1 to Step 2 CK scores.<sup>10</sup> As a component of the National Resident Matching Program's 2018 Program Director Survey, residency program directors

were asked to indicate and rate factors they used when selecting applicants to interview.<sup>36</sup> Of the 33 factors provided, the USMLE Step 1/Comprehensive Osteopathic Medical Licensing Examination of the United States (COMLEX-USA) Level 1 score was the most cited factor at 94%, with an average importance rating of 4.1 out of 5. Other factors cited by at least 70% of program directors included the following: letters of recommendation in the specialty, Medical Student Performance Evaluations (MSPEs), USMLE Step 2 CK/COMLEX-USA Level 2 score, personal statement, clerkship grades, any failed attempt of a USMLE/COMLEX-USA exam, and class ranking/quartile. As with USMLE scores, URM medical students have been found to be at a persistent disadvantage in most of these categories.  
30,31,33-35

While the Step 1 exam was designed to determine if an individual meets the minimum proficiency in the basic sciences needed to obtain medical licensure (likewise, the Step 2 CK exam determines this for the clinical sciences), as noted above, these exam scores are considered to be among the most important factors for residency interview selection.<sup>36</sup> Furthermore, 64% of program directors indicated using a target Step 1 score and 40% used a target Step 2 CK score when considering which applicants to invite for interviews.<sup>36</sup> Residency programs continue to use score cutoffs, even though the data have long supported that this method is disproportionately biased against women and URM applicants.<sup>37-40</sup> If (because of the upcoming Step 1 changes) the weighting of factors used to determine residency interviews transfers to the other highest cited factors, such as Step 2 CK score and class ranking/quartile, then medical education risks perpetuating the same disparities through a new method.

Because the achievement gap affects multiple aspects of medical education, simply shifting the focus from USMLE scores to these other academic factors will not eliminate disparities in residency applicant screening. URM medical students are more likely to receive lower grades than their White peers in their third-year clerkships. In a study of ~3,000 medical students, White students were twice as likely as URM students to receive the highest attainable grade across every core clerkship.<sup>41</sup> In another study of ~1,100 medical students at the University of Washington, Low and colleagues observed that, even after adjusting for Step 1 scores and gender, both URM and Asian students received lower final clerkship grades than their White counterparts.<sup>34</sup> They also found that URM medical students received less favorable MSPE summary words. That is, whereas 27% of all students received “outstanding” reviews, only 3% of URM students were rated “outstanding” on the MSPE. Moreover, in a 2018 analysis of differences in narrative language in student evaluations at 2 medical schools, Rojek and colleagues observed that URM and women students were significantly more likely than non-URM and male students to be described by their personal attributes (e.g., pleasant) rather than by competency-related attributes (e.g., knowledgeable).<sup>42</sup> The importance of narrative language in student evaluations cannot be overstated, as these evaluations are the basis of clerkship grades and are quoted in the MSPE and recommendation letters, all of which are core components of the residency application.

URM medical students’ struggle does not end upon completion of third-year clerkships, as they typically take the Step 2 CK exam immediately following these core clerkships. The Step 2 CK exam correlates to clinical principles, and some research suggests that it may be

an appropriate predictor of residency performance and passing specialty board exams.<sup>43-45</sup> Yet, it is important to remember that standardized tests scores are influenced by environmental and nonacademic factors, such as socioeconomic status, and may not assess clinical performance equally across all populations.<sup>46-48</sup> Analyses from multiple residency programs have illustrated that despite lower Step 1 and Step 2 CK scores, URM interns and residents are just as clinically competent as their non-URM peers.<sup>40,49,50</sup> Even if the Step 2 CK exam is a better predictor of career outcomes than the Step 1 exam, medical educators must be cautious not to assign value to residency applicants in a way that reinforces bias. When test scores are overemphasized despite known persistent disparities, medical education contributes to the disproportionate harm URM students face and bolsters segregation across medical specialties.

Given the history of inequities in science and mathematics education and the current weight of Step 1 on career trajectory, including specialty choice and residency competitiveness, the elimination of numeric scoring is a step in the right direction. However, the factors that contribute to Step 1 score gaps also influence Step 2 CK scores and clerkship grades. Therefore, early interventions for URM students remain vital. The InCUS recommendations and Step 1 announcement demonstrate that there is an urgent need to address racial achievement gaps in undergraduate and graduate medical education. In the next section, we will discuss programs that seek to redress these inequities before the residency application cycle.

## Strategies to Address the Achievement Gap

The achievement gap was created and reinforced by decades of structural disparities, and it persists at every socioeconomic level.<sup>51</sup> We do not expect medical educators to bear the burden of repairing the entire society, nor do we assume that all URM students will perform poorly in medical school. However, it is important to understand the context in which a significant proportion of URM students arrive to medical school. Namely, it is the responsibility of medical education institutions to be strategic in addressing one of the largest obstacles that URM medical students are likely to face: standardized test performance.

The creation of interventions for at-risk students, who are much more likely to be URM students, is not new to medical education. Many of these programs, ranging from prematriculation programs to programs to improve test-taking skills, have shown gains in academic performance for these students.<sup>52-64</sup> However, few institutions have implemented a comprehensive overhaul of medical curricula. One institution that has done this is the University of Texas Medical Branch at Galveston, which implemented an integrated preclinical medical curriculum that emphasized problem-based learning, early identification and intervention for potentially at-risk students, and opportunities for deliberate practice through an internal question bank provided by the administration, among other measures.<sup>57,58</sup> These global changes in the curriculum resulted in a 14-point increase in mean Step 1 scores overall, but Black students experienced a 21-point mean score increase and a 94% decrease in Step 1 failures.<sup>59</sup> Integrative approaches that are targeted and proactive (such as the one used by the University of Texas Medical Branch at Galveston) improve outcomes for

all students, but the results for URM students in particular suggest that these types of approaches may be a potential means of closing the achievement gap.

Although there has been extensive research on the efficacy of preclinical interventions,<sup>52-59</sup> studies on the impact of supplemental programs for the clinical curriculum and Step 2 CK exam are lacking. Given that racial gaps persist on both Step 1 and Step 2 CK, the absence of programming after the preclinical curriculum is particularly concerning in light of the Step 1 scoring change. Simply put, if institutions see a need for intervention programs to prepare students for Step 1, that need persists throughout medical education. These interventions must be both comprehensive and longitudinal to fully close the achievement gap for URM students.

## Conclusions

The achievement gap between White and URM students exists across all educational areas and levels in the United States and is not limited to standardized test performance. Importantly, medical education is not exempt. The impact of centuries of structural racism on the achievement gap is complex. While holistic admissions are a step in the right direction of acknowledging that URM students often travel a further distance to reach medical school, medical education cannot forget about these barriers once biochemistry lectures begin. A holistic admissions process that is not married to holistic medical education is simply not enough.

Here, we specifically discussed the racial achievement gap and its consequences in medical education. Not only do URM students cope with the psychological challenges often associated with being a URM in higher education, but they also face a climate that places undue emphasis on standardized test scores. And while the NBME and FSMB have decided to no longer report numeric scores for Step 1, the achievement gap will persist without deliberate, structural change. Academic interventions should be proactive and not delayed until URM and at-risk students are already failing. Evaluation measures during clerkships should be standardized, and implicit bias training is imperative. Residency programs should consider the consequences of using *any* standardized test score as the primary screening tool for extending interview offers. Policymakers should advocate for educational equity in pre-K through postsecondary education.

Admittedly, this area is difficult to address, but it is important that medical education affords *all students an equal opportunity to excel*. If medical education is to be truly committed to diversifying the medical profession (and all of its specialties), it must develop and implement specific, actionable strategies to close the achievement gap in medical education that begin during the preclinical years and extend throughout the curriculum.

## Acknowledgments:

The authors wish to thank Dr. Marquita N. Hicks (Wake Forest School of Medicine), Mr. Justin Bailey (University of Alabama at Birmingham School of Medicine), and Dr. Linda H. Godley (Alcorn State University School of Nursing) for their assistance and insights during the course of our research.

*Funding/Support:* This work was supported by the National Institutes of Health (NIH)—National Institute of General Medical Sciences Medical Scientist Training Program T32GM00836 (A.C.J.), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) P30 DK040561 (F.C.S.), and NIDDK L30 DK118710 (F.C.S.)—and the American Association of University Women (AAUW) Selected Professions Fellowship (A.C.N.).

## References

1. United States Medical Licensing Examination. Change to pass/fail score reporting for Step 1. <https://www.usmle.org/incus/#decision>. Accessed September 29, 2020.
2. United States Medical Licensing Examination. Summary report and preliminary recommendations from the Invitational Conference on USMLE Scoring (InCUS), March 11-12, 2019. [https://www.usmle.org/pdfs/incus/incus\\_summary\\_report.pdf](https://www.usmle.org/pdfs/incus/incus_summary_report.pdf). Accessed September 29, 2020.
3. Rozenshtein A, Marx MV. The USMLE Step 1 pass/fail reporting proposal: The APDR position. *Acad Radiol*. 2019;26(10):1400–1402. [PubMed: 31383545]
4. De Haan JB, Markham T, Ghebremichael S. Considerations of a resident recruitment committee on the USMLE Step 1 examination. *Acad Med*. 2019;94(7):922–923.
5. Swails JL, Omowunmi A, Stoll BJ. The conundrum of the United States Medical Licensing Examination score reporting structure. *JAMA*. 2019;322(7):605–606.
6. Weiner S Should the USMLE be pass/fail? Published August 13, 2019. <https://www.aamc.org/news-insights/should-usmle-be-pass/fail>. Accessed September 29, 2020.
7. Chaudhry HJ, Katsufarakis PJ, Tallia AF. The USMLE Step 1 decision: An opportunity for medical education and training. *JAMA*. 2020;323(20):2017–2018. [PubMed: 32142115]
8. Lin GL, Nwora C, Warton L. Pass/fail score reporting for USMLE Step 1: An opportunity to redefine the transition to residency together. *Acad Med*. 2020;95:1308–1311. [PubMed: 32379144]
9. Willett LL. The impact of a pass/fail Step 1—A residency program director’s view. *N Engl J Med*. 2020;382:2387–2389. [PubMed: 32558466]
10. McDade W, Vela MB, Sánchez JP. Anticipating the impact of the USMLE Step 1 pass/fail scoring decision on underrepresented-in-medicine students. *Acad Med*. 2020;95:1318–1321. [PubMed: 32379143]
11. Jencks C, Phillips M, eds. *The Black-White Test Score Gap*. Washington, DC: Brookings Institution Press; 2011.
12. Musu-Gillette L, de Brey C, McFarland J, Hussar W, Sonnenberg W, Wilkinson-Flicker S. Status and trends in the education of racial and ethnic groups 2017. Washington, DC: U.S. Department of Education; 2017 <https://nces.ed.gov/pubs2017/2017051.pdf>. Accessed September 29, 2020.
13. Orfield G, Frankenberg E, Ee J, Ayscue JB. *Harming our common future: America’s segregated schools 65 years after Brown*. Published May 10, 2019. <https://www.civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/harming-our-common-future-americas-segregated-schools-65-years-after-brown>. Accessed September 29, 2020.
14. Rearden SF. School segregation and racial academic achievement gaps. *RSF Russell Sage Found J Soc Sci*. 2016;2(5):34–57.
15. Mosteller F The Tennessee study of class size in the early school grades. *Future Child*. 1995;5(2):113–127. [PubMed: 8528684]
16. Fry R The high schools Hispanics attend: Size and other key characteristics. Washington, DC: Pew Hispanic Center; 2005 <https://www.pewtrusts.org/en/research-and-analysis/reports/2005/11/01/the-high-schools-hispanics-attend-size-and-other-key-characteristics>. Accessed September 29, 2020.
17. Snyder TD, de Brey C, Dillow SA. *Digest of education statistics 2017*, 53rd edition. Washington, DC: U.S. Department of Education; 2019:259 <https://nces.ed.gov/pubs2018/2018070.pdf>. Accessed September 29, 2020.
18. U.S. Department of Education Office for Civil Rights. 2015–16 Civil Rights Data Collection: STEM course taking. Washington, DC: U.S. Department of Education ; 2018 <https://www2.ed.gov/about/offices/list/ocr/docs/stem-course-taking.pdf>. Accessed September 29, 2020.
19. College Board. 2018 SAT Suite of Assessments Annual Report. New York, NY: College Board; 2018 <https://reports.collegeboard.org/pdf/2018-total-group-sat-suite-assessments-annual-report.pdf>. Accessed September 29, 2020.



20. Griffin K, Allen W. Mo' money, mo' problems? High-achieving black high school students' experiences with resources, racial climate, and resilience. *J Negro Educ.* 2006;75(3):478–494.
21. Spencer SJ, Logel C, Davies PG. Stereotype threat. *Annu Rev Psychol.* 2016;67:415–437. [PubMed: 26361054]
22. Cadinu M, Maass A, Rosabianca A, Kiesner J. Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychol Sci.* 2005;16(7):572–578. [PubMed: 16008792]
23. Owens J, Massey DS. Stereotype threat and college academic performance: A latent variables approach. *Soc Sci Res.* 2011;40(1):150–166. [PubMed: 23950616]
24. Brown RP, Day EA. The difference isn't black and white: Stereotype threat and the race gap on Raven's Advanced Progressive Matrices. *J Appl Psychol.* 2006;91(4):979–985. [PubMed: 16834521]
25. Ertl B, Luttenberger S, Paechter M. The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front Psychol.* 2017;8:703. [PubMed: 28567022]
26. McClain S, Beasley ST, Jones B, Awosogba O, Jackson S, Cokley K. An examination of the impact of racial and ethnic identity, impostor feelings, and minority status stress on the mental health of Black college students. *J Multicult Counsel Dev.* 2016;44(2):101–117.
27. Levant B, Villwock JA, Manzardo AM. Impostorism in American medical students during early clinical training: Gender differences and intercorrelating factors. *Int J Med Educ.* 2020;11:90–96. [PubMed: 32356519]
28. Lucey CR, Saguil A. The consequences of structural racism on MCAT scores and medical school admissions: The past is prologue. *Acad Med.* 2020;95:351–356. [PubMed: 31425184]
29. Association of American Medical Colleges. Table A-18: MCAT scores and GPAs for applicants and matriculants to U.S. medical schools by race/ethnicity, 2019-2020. [https://www.aamc.org/system/files/2019-10/2019\\_FACTS\\_Table\\_A-18.pdf](https://www.aamc.org/system/files/2019-10/2019_FACTS_Table_A-18.pdf). Accessed October 19, 2020.
30. Rubright JD, Jodoin M, Barone MA. Examining demographics, prior academic performance, and United States Medical Licensing Examination scores. *Acad Med.* 2019;94(3):364–370. [PubMed: 30024473]
31. Andriole DA, Jeffe DB. A national cohort study of U.S. medical school students who initially failed Step 1 of the United States Medical Licensing Examination. *Acad Med.* 2012;87(4):529–536. [PubMed: 22361789]
32. Huff KL, Fang D. When are students most at risk of encountering academic difficulty? A study of the 1992 matriculants to U.S. medical schools. *Acad Med.* 1999;74:454–460. [PubMed: 10219232]
33. Boatright D, Ross D, O'Connor P, Moore E, Nunez-Smith M. Racial disparities in medical student membership in the Alpha Omega Alpha honor society. *JAMA Intern Med.* 2017;177(5):659–665. [PubMed: 28264091]
34. Low D, Pollack SW, Liao ZC, et al. Racial/ethnic disparities in clinical grading in medical school. *Teach Learn Med.* 2019;31(5):487–496. [PubMed: 31032666]
35. Teherani A, Hauer KE, Fernandex A, King TE Jr, Lucey C. How small differences in assessed clinical performance amplify large differences in grades and awards: A cascade with serious consequences for students underrepresented in medicine. *Acad Med.* 2018;93(9):1286–1292. [PubMed: 29923892]
36. National Resident Matching Program. Results of the 2018 NRMP Program Director Survey. Washington, DC: National Resident Matching Program; 2018:3 <https://www.nrmp.org/wp-content/uploads/2018/07/NRMP-2018-Program-Director-Survey-for-WWW.pdf>. Accessed September 29, 2020.
37. Spector AR, Railey KM. Reducing reliance on test scores reduces racial bias in neurology residency recruitment. *J Natl Med Assoc.* 2019;111(5):471–474. [PubMed: 30935681]
38. 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. [PubMed: 11739053]

39. Williams M, Kim EJ, Pappas K, et al. The impact of United States Medical Licensing Exam (USMLE) Step 1 cutoff scores on recruitment of underrepresented minorities in medicine: a retrospective cross-sectional study. *Health Sci Rep.* 2020;3(2):e2161. [PubMed: 32318628]
40. Fernandez C, Lopez BL, Kushner M, Leiby BE, Den RB. Overemphasis of Step 1 scores may affect application pool diversity in radiation oncology. *Pract Radiat Oncol.* 2020;10(1):e3–e7. [PubMed: 31369886]
41. Lee KB, Vaishnavi SN, Lau SK, Andriole DA, Jeffe DB. “Making the grade”: Noncognitive predictors of medical students’ clinical clerkship grades. *J Natl Med Assoc.* 2007;99(10):1138–1150. [PubMed: 17987918]
42. Rojek AE, Khanna R, Yim JW, et al. Differences in narrative language in evaluations of medical students by gender and under-represented minority status. *J Gen Intern Med.* 2019;34(5):684–691. [PubMed: 30993609]
43. Sharma A, Schauer DP, Kelleher M, Kinnear B, Sall D, Warm E. USMLE Step 2 CK: Best predictor of multimodal performance in an internal medicine residency. *J Grad Med Educ.* 2019;11(4):412–419. [PubMed: 31440335]
44. Lee M, Vermillion M. Comparative values of medical school assessments in the prediction of internship performance. *Med Teach.* 2018;40(12):1287–1292. [PubMed: 29390938]
45. Radabaugh CL, Hawkins RE, Welcher CM, et al. Beyond the United States Medical Licensing Examination score: Assessing competence for entering residency. *Acad Med.* 2019;94(7):983–989. [PubMed: 30920448]
46. Sedlacek WE. Multiple choices for standardized tests. *Priorities.* 1998;10:1–21. <http://williamsedlacek.info/publications/articles/multiple1.pdf>. Accessed September 29, 2020.
47. Jencks C. America’s next achievement test: Closing the Black-White test score gap. *Am Prospect.* 1998;40:44–53.
48. Giordano C, Hutchinson D, Peppler R. A predictive model for USMLE Step 1 scores. *Cureus.* 2016;8(9):e769. [PubMed: 27738569]
49. Lypson ML, Ross PT, Hamstra SJ, Haftel HM, Gruppen LD, Colletti LM. Evidence for increasing diversity in graduate medical education: The competence of underrepresented minority residents measured by an intern objective structured clinical examination. *J Grad Med Educ.* 2010;2(3):354–359. [PubMed: 21976083]
50. Poon S, Nellans K, Rothman A, et al. Underrepresented minority applicants are competitive for orthopaedic surgery residency programs, but enter residency at lower rates. *J Am Acad Orthop Surg.* 2019;27(21):e957–e968. [PubMed: 30614894]
51. Lubienski S. A closer look at Black-White mathematics gaps: Intersections of race and SES in NAEP achievement and instructional practices data. *J Negro Educ.* 2002;71(4):269–287.
52. Williams MT. Pre-matriculation program at the University of South Florida College of Medicine. *Acad Med.* 1999;74(4):397–399. [PubMed: 10219221]
53. Schneid SD, Apperson A, Laiken N, Mandel J, Kelly CJ, Brandle K. A summer prematriculation program to help students in medical school. *Adv Health Sci Educ.* 2019;23:499–511.
54. Wilson WA, Henry MK, Ewing G, et al. A prematriculation intervention to improve the adjustment of students to medical school. *Teach Learn Med.* 2011;23(3):256–262. [PubMed: 21745061]
55. Kosbuski AW, Whitney A, Skildum A, Prunuske A. Development of an interdisciplinary pre-matriculation program designed to promote medical students’ self efficacy. *Med Educ Online.* 2017;22(1):1272835. [PubMed: 28178916]
56. Winston KA, van der Vleuten CPM, Scherpbier AJJA. An investigation into the design and effectiveness of a mandatory cognitive skills programme for at-risk medical students. *Med Teach.* 2010;32(3):236–243. [PubMed: 20218839]
57. Asimakis GK, Ainsworth MA, Aronson JF, Frye AW, Lieberman SA, Rabek JP. Evolution of student assessment following implementation of an integrated medical curriculum: Contribution to improved education outcomes. *Med Sci Educ.* 2011;21(2):181–189.
58. Lieberman SA, Frye AW, Thomas L, Rabek JP, Anderson GD. Comprehensive changes in the learning environment: Subsequent Step 1 scores of academically at-risk students. *Acad Med.* 2008;83(10 Suppl):S49–S52. [PubMed: 18820500]

59. Lieberman SA, Ainsworth MA, Asimakis GK, et al. Effects of comprehensive education reforms on academic success in a diverse student body. *Med Educ.* 2010;44:1232–1240. [PubMed: 21070343]
60. Yu F, Liu YH. The comparative effects of student question-posing and question-answering strategies on promoting college students' academic achievement, cognitive and metacognitive strategies use. *J Educ Psychol.* 2008;31(3):25–52.
61. Walsh J, Harris B, Tayyaba S, Harris D, Smith P. Student-written single-best answer questions predict performance in finals. *Clin Teach.* 2016;13:352–356. [PubMed: 26450346]
62. Abdolhussein S, Aslami M, Bijanzadeh M. The effect of question generation activity on students' learning and perception. *J Adv Med Educ Prof.* 2018;6(2):70–77. [PubMed: 29607334]
63. Bottomley S, Denny P. A participatory learning approach to biochemistry using student authored and evaluated multiple-choice questions. *Biochem Mol Biol Educ.* 2011;39(5):352–361. [PubMed: 21948507]
64. Gooi ACC, Sommerfeld CS. Medical school 2.0: How we developed a student-generated question bank using small group learning. *Med Teach.* 2014;37(10):892–896. [PubMed: 25310243]