

# Comprehensive Reform and Greater Equity in Applying to Residency—Trainees' Mixed Responses to a Pass/Fail USMLE Step 1

Nishant Ganesh Kumar, MD

Matthew E. Pontell, MD

Alan T. Makhoul, BA

Brian C. Drolet, MD

## ABSTRACT

**Background** Pass/fail USMLE Step 1 score reporting may have varying implications for trainees of different demographic and training backgrounds.

**Objective** To characterize the perspectives of a diverse cohort of trainees on the impact of pass/fail Step 1 score reporting.

**Methods** In 2020, 197 US and international medical school deans and 822 designated institutional officials were invited to distribute anonymous electronic surveys among their trainees. Separate surveys for medical students and residents/fellows were developed based on the authors' prior work surveying program directors on this topic. Underrepresented in medicine (UiM) was defined in accordance with AAMC definitions. Descriptive and comparative analyses were performed, and results were considered statistically significant with  $P < .05$ .

**Results** A total of 11 633 trainees responded (4379 medical students and 7254 residents/fellows; 3.3% of an estimated 285 000 US trainees). More students favored the score reporting change than residents/fellows (43% vs 31%;  $P < .001$ ; 95% CI 0–24). Trainees identifying as UiM were more likely to favor the change (50% vs 34%;  $P < .001$ ; 95% CI 0–32) and to agree it would decrease socioeconomic disparities (44% vs 25%;  $P < .001$ ; 95% CI 0–38) relative to non-UiM trainees. Nearly twice as many osteopathic and international medical graduate students felt they would be disadvantaged compared to MD students because of pass/fail score reporting (61% vs 31%;  $P < .001$ ; 95% CI 0–60).

**Conclusions** Trainee perspectives regarding USMLE Step 1 score reporting are mixed. UiM trainees were more likely to favor the score reporting change, while osteopathic and international medical students were less in favor of the change.

## Introduction

As the number of residency applications has grown, programs have increasingly relied on United States Medical Licensing Examination (USMLE) scores to screen applicants.<sup>1–3</sup> While standardized test performance may reflect academic potential, it is also laden with inequities that disproportionately hinder residency prospects for minority groups underrepresented in medicine (UiM).<sup>4–6</sup> To decrease reliance on Step 1 scores, USMLE co-sponsors announced a transition from numerical to binary (pass/fail) score reporting as early as 2022.<sup>7</sup> In a 2020 national survey of program directors (PDs) from 30 specialties with a 45% response rate, the majority (61%) disagreed with the change compared to 15% who thought the USMLE Step 1 score reporting change was beneficial.<sup>8</sup> As to the consequences of the score reporting change on socioeconomic disparities, 14% of PDs

agreed the change would be beneficial and 43% disagreed.<sup>8</sup>

Currently, published perspectives from trainees are limited to small groups of medical students, primarily from US allopathic backgrounds. Through a national survey, we sought to understand medical student and resident/fellow perceptions regarding the change to pass/fail USMLE Step 1 score reporting and whether demographic or training factors were associated with these perceptions.

## Methods

In July and August 2020, 197 deans of accredited medical schools (142 US allopathic, 30 US osteopathic, and 25 international allopathic) and 822 designated institutional officials (DIOs) from Accreditation Council for Graduate Medical Education (ACGME)–accredited institutions were invited to distribute anonymous electronic surveys among their trainees. US medical school deans and DIOs were identified through publicly available directories published by the Association of American Medical Colleges (AAMC), the American Association of Colleges of

DOI: <http://dx.doi.org/10.4300/JGME-D-20-01511.1>

*Editor's Note: The online version of this article contains the demographics and selected survey responses from total cohort of trainees and the surveys used in the study.*

Osteopathic Medicine (AACOM), and the ACGME.<sup>9–11</sup> Email addresses were obtained from publicly available sources. Because we corresponded only with medical school deans and DIOs, and survey distribution was handled by separate officials, it was not possible to know which institutions successfully distributed the survey and to what proportion of trainees. Therefore, a response rate was estimated using the total number of trainees.

Separate analogous survey instruments were designed for medical students and residents/fellows.<sup>12</sup> Survey items were developed by the authors (one medical student, one resident, one fellow, and one attending) based on prior experience sampling PDs regarding their perceptions of pass/fail Step 1 score reporting.<sup>8</sup> Survey questions utilized 3-item Likert scales. Instruments were pilot tested with a diverse group from each target population (medical students: 4 White, 1 African American, 1 American Indian or Alaska Native, and 2 Asian American; residents/fellows: 3 White, 2 African American, 2 Asian American, and 1 respondent who self-identified as “Other”). Internal consistency reliability of each instrument was assessed using Cronbach’s alpha. The final student and resident/fellow instruments can be found in supplemental data.

UiM trainees were defined as those who identified as African American, Mexican American, and Native American (ie, American Indians, Alaska Natives, and Native Hawaiians).<sup>13</sup> To contextualize the proportion of UiM trainees in our cohort relative to the national proportion of UiM trainees, these numbers are also reported using national estimates.

The primary outcome of interest was to assess student and resident/fellow sentiment toward the change by asking if “Changing the USMLE Step 1 to pass/fail is a good idea.” Additionally, we assessed for potential improvements in socioeconomic disparities by asking if “Changing the USMLE Step 1 to pass/fail will decrease socioeconomic disparities in the residency application process.” Free-text responses were solicited to identify reforms that trainees believed would improve the residency application process. Lastly, trainee opinions regarding Step 2 Clinical Knowledge (CK) score reporting were examined.

Comparative analyses were performed using 2-tailed Fisher’s exact test, and differences were considered statistically significant with a *P* value less than .05. Additionally, 95% CIs are reported for differences in subgroup responses.<sup>14</sup> As this was exploratory research without prior hypotheses, 8 comparisons were made without adjusting for multiple comparisons, and the *P* level remained set at .05 for each comparison. Free text survey responses were

reviewed by the authors to identify salient suggestions regarding residency application reforms.

Both surveys received exemption from the Vanderbilt University Institutional Review Board.

## Results

Both student and resident/fellow survey instruments demonstrated internal consistency (Cronbach’s alpha 0.76 and 0.91, respectively). After invitation, 37 deans (19%) and 266 (32%) DIOs agreed to survey distribution. In total, 11 633 trainees (4379 medical students and 7254 residents/fellows) responded, of which 9303 were US allopathic or osteopathic trainees. With an estimated 232 000 allopathic and 53 000 osteopathic medical students and residents in the United States, our response rate approximates 3.3% of all US trainees.<sup>15,16</sup> Nationally, nearly 8.5% of DO medical students (~2300) and 21% of MD medical students (~20 000) identify as UiM.<sup>17,18</sup> In our national sample, 14% (~600) of medical students identified as UiM.

Overall, 36% of respondents (4146 of 11 609) agreed that the score reporting change was a good idea and 27% (3133 of 11 589) agreed the change would decrease socioeconomic disparities in the residency application process (TABLE and online supplementary data). More medical students agreed with the score reporting change than residents/fellows (1891 of 4374 [43%] vs 2255 of 7235 [31%]; *P* < .001; 95% CI 0–24).

UiM trainees were more likely to agree with pass/fail score reporting compared to non-UiM trainees (493 of 986 [50%] vs 3653 of 10 623 [34%]; *P* < .001; 95% CI 0–32). Additionally, UiM trainees were more likely to agree that pass/fail score reporting would decrease socioeconomic disparities associated with Step 1 compared to non-UiM trainees (433 of 983 [44%] vs 2700 of 10 606 [25%]; *P* < .001; 95% CI 0–38). A greater proportion of UiM trainees agreed with pass/fail Step 2 CK than non-UiM trainees (366 of 980 [37%] vs 2468 of 10 562 [23%]; *P* < .001; 95% CI 0–28).

Fewer US osteopathic (DO) and international medical graduate (IMG) trainees agreed with the change compared to MD trainees (1190 of 4112 [29%] vs 2806 of 7120 [39%]; *P* < .001; 95% CI 0–20). DO and IMG trainees were less likely to agree that the score reporting change would decrease socioeconomic disparities in the application process (2097 of 4002 [52%] vs 2935 of 7111 [41%]; *P* < .001; 95% CI 0–22). Nearly twice as many DO and IMG students agreed they would be disadvantaged by their training backgrounds relative to US MD students (620 of 1024 [61%] vs 1012 of 3282

**TABLE**  
Survey Responses From Medical Student and Resident and Fellow Subgroups

Trainee Groups	Changing Step 1 to Pass/Fail:						Step 2 CK Should Be Reported as:					
	Is a Good Idea			Will Decrease Socioeconomic Disparities in the Residency Application Process			Pass/Fail	3-Digit Numeric	Quartiles			
	Agree	Neutral	Disagree	Agree	Neutral	Disagree						
Total cohort (n = 11 633)	35.7% (4146/11 609)	20.1% (2328/11 609)	44.2% (5135/11 609)	27% (3133/11 589)	27.6% (3198/11 589)	45.4% (5258/11 589)	24.6% (2834/11 542)	53.9% (6225/11 542)	21.5% (2483/11 542)			
Medical students (n = 4379)	43.2% (1891/4374)	20.2% (884/4374)	36.6% (1599/4374)	34.9% (1522/4367)	25.5% (1112/4367)	39.7% (1733/4367)	28.8% (1250/4343)	47.6% (2067/4343)	23.6% (1026/4343)			
UIM (n = 597)	57.3% (342/597)	18.9% (113/597)	23.8% (142/597)	48.8% (291/596)	22.1% (132/596)	29.0% (173/596)	39.3% (233/593)	37.6% (223/593)	23.1% (137/593)			
Non-UIM (n = 3782)	41.0% (1549/3777)	20.4% (771/3777)	38.6% (1457/3777)	32.6% (1231/3771)	26.0% (980/3771)	41.4% (1560/3771)	27.1% (1017/3750)	49.2% (1844/3750)	23.7% (889/3750)			
US MD (n = 3288)	45.9% (1507/3286)	20.6% (677/3286)	33.5% (1102/3286)	37.4% (1228/3281)	26.2% (859/3281)	36.4% (1194/3281)	29.2% (952/3263)	46.2% (1507/3263)	24.6% (804/3263)			
US DO (n = 857)	31.7% (303/957)	28.0% (268/957)	40.3% (386/957)	26.9% (230/856)	22.7% (194/856)	50.5% (432/856)	29.6% (252/850)	49.6% (422/850)	20.7% (176/850)			
IMG (n = 168)	29.2% (49/168)	18.5% (31/168)	52.4% (88/168)	24.0% (40/167)	24.6% (41/1670)	51.5% (86/167)	17.3% (29/168)	63.7% (107/168)	19.0% (32/168)			
Residents/fellows (n = 7254)	31.2% (2255/7235)	20.0% (1444/7235)	48.9% (3536/7235)	22.3% (1611/7222)	28.9% (2086/7222)	48.8% (3525/7222)	22.0% (1584/7199)	57.8% (4158/7199)	20.2% (1457/7199)			
UIM (n = 392)	38.8% (151/389)	20.3% (79/389)	40.9% (159/389)	36.7% (142/387)	17.3% (67/387)	46.0% (178/387)	34.4% (133/387)	48.1% (186/387)	17.6% (68/387)			
Non-UIM (n = 6862)	30.7% (2104/6846)	19.9% (1365/6846)	49.3% (3377/6846)	21.5% (1469/6835)	29.5% (2019/3347)	49.0% (3347/6835)	21.3% (1451/6812)	58.3% (3972/6812)	20.4% (1389/6812)			
US MD (n = 3842)	33.9% (1299/3384)	21.1% (808/3834)	45.0% (1727/3834)	23.4% (897/3830)	31.1% (1192/3830)	45.5% (1741/3830)	22.9% (876/3823)	53.8% (2055/3823)	23.3% (892/3823)			
US DO (n = 1316)	33.9% (445/1312)	21.0% (275/1312)	45.1% (592/1312)	21.6% (283/1311)	28.4% (372/1311)	50.0% (656/1311)	26.3% (344/1307)	52.6% (688/1307)	21.0% (275/1307)			
IMG (n = 1677)	23.5% (393/1675)	16.8% (282/1675)	59.7% (1000/1675)	20.5% (342/1668)	24.2% (403/1668)	55.3% (923/1668)	16.3% (271/1661)	70.3% (1168/1661)	13.4% (222/1661)			

Abbreviations: UIM, underrepresented in medicine; IMG, international medical graduate.

Note: The first row of each section reports the responses of the total medical students and residents and fellows respectively, with standard percentages over the total respondents for each question. The total cohort was further broken down to UIM, Non-UIM, US MD, US DO, and IMG subgroups and responses were reported in each of the corresponding rows.

[31%];  $P < .001$ ; 95% CI 0–60; TABLE). DO/IMG students were also more likely to agree with being unable to distinguish themselves on the basis of merit without a numeric Step 1 score (678 of 1023 [66%] vs 1665 of 3282 [51%];  $P < .001$ ; 95% CI 0–30).

A common free-text suggestion to improve equity in the residency application process was capping the number of interview invitations an applicant may accept. Many respondents also suggested increasing transparency (eg, disclosure of score and medical school rank cutoffs) and standardization of application components (eg, Medical Student Performance Evaluation and letters of recommendation). See TABLE and online supplementary data for trainee opinions on USMLE Step 2 CK reporting.

## Discussion

This survey of allopathic and osteopathic students' as well as graduate medical trainees' perceptions of the upcoming change of USMLE Step 1 score reporting to pass/fail found that about one-third of students and trainees agree with the change. UiM trainees were more likely to support this change, whereas DO and IMG trainees were less likely to support the change in comparison with US allopathic trainees.

The effect of pass/fail Step 1 score reporting on UiM trainees remains an open question. While some have argued that pass/fail reporting will help address racial disparities in examination performance,<sup>5</sup> others have suggested that the adoption of substitute heuristics will increase bias in the residency selection process.<sup>19</sup> Our findings support the notion that pass/fail score reporting will have a positive impact on trainees from underrepresented backgrounds. UiM trainees were significantly more likely to support the change compared to non-UiM trainees. Moreover, UiM trainees were more likely to favor additional reforms aimed at reducing emphasis on standardized testing, such as pass/fail score reporting of Step 2 CK.

The score reporting change has prompted concerns in the IMG community. Because Step 1 numeric scores served as a standard metric of medical knowledge that could be compared among applicants regardless of medical school background, many IMGs are concerned they will be unable to distinguish themselves from US graduates.<sup>20</sup> As it stands, IMGs believe they will have a more difficult time securing residency positions.<sup>21</sup> With Step 2 CK as the last remaining standardized test with numeric score reporting, many IMGs have expressed undue stress.<sup>21</sup> This potential consequence contradicts one

of the initial reasons for the score reporting change, which was to decrease the negative impact of Step 1 numeric scores on student well-being.<sup>7,22</sup> These perspectives may explain why the majority (59%) of IMGs in our study disagreed with the score reporting change.

Osteopathic students face similar challenges to IMG students. Lin et al<sup>22</sup> describe how, for students of osteopathic and international backgrounds, the loss of Step 1 numeric scores represents the loss of a key metric to ensure applications from these non-allopathic students are reviewed. Similarly, in a single institution cross-sectional survey of 215 medical students, osteopathic students were more likely to report that the pass/fail transition will negatively impact their residency match compared to allopathic students (adjust OR 1.454, 95% CI 0.515–4.106).<sup>23</sup> This sentiment is supported by our findings, as only a third of osteopathic trainees agreed with the change, and more osteopathic students felt they would be disadvantaged compared to their allopathic counterparts.

The primary limitation of this study is a low response rate, which is estimated using the largest possible denominator. The actual number of trainees that received the surveys is unknown. Thus, the respondents may not represent all medical students and GME trainees. In addition, other than pilot testing for clarity and consistency, the survey has limited validity evidence, and respondents may have interpreted questions differently than intended. In particular, demographic categories may not have included all possible responses. As this was an exploratory project, the  $P$  level of .05 was not corrected for multiple comparisons, and some differences may not be true. Next research steps may include qualitative approaches to examine concerns expressed by medical and osteopathic students regarding the upcoming reporting change, as well as mitigating strategies to ensure some students, such as osteopathic or international medical students, are not disadvantaged by the change.

## Conclusions

Trainee perspectives regarding pass/fail USMLE Step 1 score reporting are mixed. Overall, a third of trainees agreed with the score reporting change, with more students than residents/fellows supporting the change. UiM trainees were more likely to support the change than non-UiM trainees. In contrast, DO and IMG trainees were less in favor of the change relative to MD trainees.

---

**References**

- National Resident Matching Program. Results of the 2020 NRMP Program Director Survey. <https://mk0nrmp3oyqui6wqfm.kinstacdn.com/wp-content/uploads/2020/08/2020-PD-Survey.pdf>. Accessed July 8, 2021.
- National Resident Matching Program. Results of the 2016 NRMP Program Director Survey. <https://www.nrmp.org/wp-content/uploads/2016/09/NRMP-2016-Program-Director-Survey.pdf>. Accessed July 8, 2021.
- National Resident Matching Program. Results of the 2010 NRMP Program Director Survey. <https://mk0nrmp3oyqui6wqfm.kinstacdn.com/wp-content/uploads/2013/08/programresultsbyspecialty2010v3.pdf>. Accessed July 8, 2021.
- Grubbs V. Diversity, equity, and inclusion that matter. *N Engl J Med*. 2020;383(4):e25. doi:10.1056/NEJMp2022639
- Youmans QR, Essien UR, Capers Q. A test of diversity—what USMLE pass/fail scoring means for medicine. *N Engl J Med*. 2020;382(25):2393–2395. doi:10.1056/NEJMp2004356
- 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. doi:10.1097/ACM.0000000000002366
- United States Medical Licensing Examination. Summary report and preliminary recommendations from the Invitational Conference on USMLE Scoring (InCUS) 2019. [https://www.usmle.org/pdfs/incus/incus\\_summary\\_report.pdf](https://www.usmle.org/pdfs/incus/incus_summary_report.pdf). Accessed July 8, 2021.
- Makhoul AT, Pontell ME, Ganesh Kumar N, Drolet BC. Objective measures needed—program directors' perspectives on a pass/fail USMLE Step 1. *N Engl J Med*. 2020;382(25):2389–2392. doi:10.1056/NEJMp2006148
- Accreditation Council for Graduate Medical Education. Sponsor search. <https://apps.acgme.org/ads/Public/Sponsors/Search>. Accessed July 8, 2021.
- Association of American Medical Colleges. AAMC Medical School Members. <https://members.aamc.org/eweb/DynamicPage.aspx?webcode=AAMCOrgSearchResult&orgtype=Medical%20School>. Accessed July 8, 2021.
- American Association of Colleges of Osteopathic Medicine. U.S. Colleges of Osteopathic Medicine. <https://www.aacom.org/become-a-doctor/u-s-colleges-of-osteopathic-medicine>. Accessed July 8, 2021.
- Rickards G, Magee C, Artino AR Jr. You can't fix by analysis what you've spoiled by design: developing survey instruments and collecting validity evidence. *J Grad Med Educ*. 2012;4(4):407–410. doi:10.4300/JGME-D-12-00239.1
- Association of American Medical Colleges. Underrepresented in Medicine Definition. <https://www.aamc.org/what-we-do/equity-diversity-inclusion/underrepresented-in-medicine>. Accessed July 8, 2021.
- Altman DG, Bland JM. How to obtain the confidence interval from a P value. *BMJ*. 2011;343:d2090. doi:10.1136/bmj.d2304
- Association of American Medical Colleges. Enrollment Up at U.S. Medical Schools. <https://www.aamc.org/news-insights/press-releases/enrollment-us-medical-schools>. Accessed July 8, 2021.
- American Osteopathic Association. Osteopathic Medical Profession Surges above 150,000 Physicians and Students. <https://osteopathic.org/2020/04/27/osteopathic-medical-profession-surges-above-150000-physicians-and-students/>. Accessed July 8, 2021.
- Association of American Medical Colleges. Total Enrollment by U.S. Medical School and Race/Ethnicity (Alone or In Combination), 2020-2021. <https://www.aamc.org/media/9621/download>. Accessed July 8, 2021.
- American Association of Colleges of Osteopathic Medicine. Osteopathic Medical College Diversity in Enrollment. <https://www.aacom.org/become-a-doctor/diversity-in-ome/diversity-data/diversity-in-enrollment>. Accessed July 8, 2021.
- 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(9):1318–1321. doi:10.1097/ACM.0000000000003490
- Al-Akchar M, Salih M, Fanari Z. USMLE step 1 pass/fail: the impact on international medical graduates. *Avicenna J Med*. 2021;11(1):40–41. doi:10.4103/ajm.ajm\_154\_20
- Boulet JR, Pinsky WW. Reporting a pass/fail outcome for USMLE Step 1: consequences and challenges for international medical graduates. *Acad Med*. 2020;95(9):1322–1324. doi:10.1097/ACM.0000000000003534
- 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(9):1308–1311. doi:10.1097/ACM.0000000000003495
- Ehrlich H, Sutherland M, McKenney M, Elkbuli A. Implications of the United States Medical Licensing Examination Step 1 examination transition to pass/fail on medical students education and future career opportunities [published online ahead of print



December 19, 2020]. *Am Surg*. doi:10.1177/  
0003134820973382



**Nishant Ganesh Kumar, MD**, is a Resident, Section of Plastic Surgery, Department of Surgery, University of Michigan Health System; **Matthew E. Pontell, MD**, is a Resident, Department of Plastic Surgery, Vanderbilt University Medical Center; **Alan T. Makhoul, BA**, is a Medical Student, School of Medicine, Vanderbilt University; and **Brian C. Drolet, MD**, is Associate Professor and Program Director, Department of Plastic Surgery,

Department of Biomedical Informatics, Center for Biomedical Ethics and Society, Vanderbilt University Medical Center

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

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

**Corresponding author:** Brian C. Drolet, MD, Vanderbilt University Medical Center, [brian.c.drolet@vumc.org](mailto:brian.c.drolet@vumc.org)

Received December 23, 2020; revisions received April 2, 2021, and June 7, 2021; accepted June 11, 2021.