

Good Riddance to Step 1 Numeric Grading... But Now What?

Benjamin N. Schneider, MD

(Fam Med. 2021;53(9):751-3.)

doi: 10.22454/FamMed.2021.262563

On January 26, 2022, United States Medical Licensing Examination (USMLE) Step 1 Examination (Step 1) scores will be reported as pass or fail, as the test was initially designed. This decision was made thoughtfully and with broad input from stakeholder organizations as part of the Invitational Conference on USMLE Scoring (InCUS).¹ Family medicine (FM) educators should celebrate this change. The unintended consequences overemphasizing Step 1 have been well described for both faculty² and students.³ The current use of Step 1 as a filter for graduate medical education (GME) applications is a poor predictor of clinical performance,^{4,5} perpetuates structural inequities by race and gender,^{5,6} and negatively impacts student well-being by shifting attention away from institutional undergraduate medical education (UME) performance and extracurricular activities including service and research.³ Finally, there is reason to suspect there will be positive implications of this change concerning FM specialty choice. Chen et al describe how students choosing to specialize in primary care are often assumed to have lower examination scores and students with high Step 1 scores are commonly encouraged to apply to more competitive specialties.

The InCUS group did not recommend USMLE grade changes in isolation. They also recommended a full review of the UME-to-GME transition, leading to the creation of the UME-to-GME Education Review Committee (UGRC). The UGRC recently released their initial report with 43 preliminary recommendations⁷ including issues of advising, competency assessment, information available about applicants, interviews and visiting rotations, equity, intern preparedness, oversight, and transitioning from student to resident. I strongly encourage family medicine educators to review these recommendations. They acknowledge

that the current system is failing applicants, programs, and the public good. Although we can and should debate the details, these recommendations seek to revitalize and improve the entire UME-GME transition process. Additionally, the recommendations around competency assessment and communication form a bridge to the Accreditation Council for Graduate Medical Education (ACGME) Milestones and allow for a more seamless learning framework and opportunities for better evaluation strategies in the future.

During the COVID-19 pandemic, the USMLE Step 2 Clinical Skills Examination (CS) was suspended, then retired.⁸ While the pass rate for US graduates has been at or above 95%, the pass rate for international graduates (IMGs) has been around 75%.⁹ For programs with a large IMG applicant pool, this is yet another stressor. The Education Commission for Foreign Medical Graduates has responded by creating multiple pathways to allow IMGs to demonstrate the necessary skills to enter residency. There are now six possible pathways for the 2022 match.¹⁰ IMGs play an essential role in our workforce, and programs experience unique challenges when trying to review these applicants as they make up approximately 60% of all FM applicants.¹¹

In parallel with these changes, medical schools are shifting to pass/fail grading.¹² This shift is being driven by the heterogeneity of systems and imprecision of meaning,¹³ an increased focus on competency-based standards across the medical education continuum, and evidence that pass/no-pass grading can improve student well-being¹⁴ without impacting performance.¹⁵ Some are concerned that these changes will swing the pendulum too far

From the Department of Family Medicine, Oregon Health & Science University, Portland, OR.

away from medical knowledge, but the requisite knowledge has not changed. Step 1 will still provide, “the ability of medical licensing authorities to use the exam for its primary purpose of medical licensure eligibility.”¹¹

The confluence of these changes likely leaves program directors (PDs) feeling less informed while managing more applications than ever. The mean number of applications per applicant has skyrocketed in the past 20 years in a vicious cycle described as application fever.¹⁶ Despite informational campaigns by medical schools and organizations,¹⁷ the number of applications per applicant continues to climb.¹¹ In 2020, FM programs averaged 1,147 applications,¹¹ At 10 minutes per file, it would take 8 days working around the clock to perform even a cursory review. The current system overwhelms PDs and forces them to seek ways to filter out applicants. The National Residency Matching Program (NRMP) PD Survey reports that only 30% of FM applicants receive an in-depth review!¹⁸ Not only are filters like Step 1 score problematic for equity, well-being, and future performance, our residencies are all going after the same applicants. In 2016, 7% of FM applicants received 50% of all interview offers, and 23% of those who interviewed comprised 50% of all interviews.¹⁹ In the short term, focus will likely shift to the USMLE Step 2 Clinical Knowledge Exam. Eighty-eight percent of internal medicine and orthopedic PDs surveyed reported that the Step 1 grade change will increase emphasis on Clinical Knowledge Exam grades.²⁰ It appears we need filters, but what are the factors we want to filter in or out of our programs?

Applicants want better data too. There is no realistic way for an applicant to determine which of the 700+ FM GME programs they would be a good fit for, and the lack of transparency around what programs are looking for in applicants further increases pressures to overapply. Applicants realize that if they don't look like a student in the top 10%, the data suggest they should apply broadly to get enough interviews to match. Applicants and advisors have a variety of tools at their disposal including those produced by the American Academy of Family Physicians (AAFP),²¹ the NRMP,^{18,22} the Association of American Medical Colleges,^{23,24,17} the American Medical Association (AMA),²⁵ University Collaboration,²⁶ and third-party platforms.²⁷⁻²⁹ However, these data often focus on easily-measured factors of questionable significance, such as “number of job experiences.” Despite these resources, a

lack of transparency around selection criteria remains. The AAFP Residency Directory and AMA's FREIDA come closest to allowing students to search and filter programs, but even these are limited to searching by geography, program size, community served, and program type. Are these the factors we want our training programs defined by? The fit of applicants to programs can improve if we increase transparency about what we are looking for and add more meaningful programmatic data to the AAFP Residency Directory.

It will take multiple interventions in parallel to get us out of this mess. We must decrease the applications per applicant, clarify what we care about in applicants, be transparent in the mission and outcomes of our programs, and help build a system that allows for greater bidirectional transparency and sorting. We should not be satisfied with a system that screens out 70% of applicants. Except for application caps, it is unlikely that any single intervention will get us back to the ratios we saw in the early 2000s, but there are other options to consider. Staged applications (early acceptance), preference signaling, and interview caps could all reduce the load on programs but do little to help applicants know which programs may be a good fit for them.

Step 1 scores poorly predict GME success and have a hard time measuring the knowledge, skills, and attitudes that do predict success. For over a decade we've talked about the need for collaboration between UME and GME so interns are ready to hit the ground running.³⁰ Recent graduates have higher USMLE scores than any prior generation of physicians. With less focus on Step 1 content, this is an opportunity to acknowledge and focus on the skills across all the ACGME Core Competencies allowing applicants to thrive as residents.

Finally, the FM community should begin treating Step 1 as a pass/no-pass exam with the class of '23. The pandemic's impact on this cohort has been incredible. As a student affairs dean, I've heard countless stories of how COVID-19 personally impacted students and their families during the dedicated study period. As students tried to sit for the exam, the disruption continued. Nearly half the students at my institution had their exams canceled because of testing site closures. Some were notified the day of their exam and many had multiple cancellations.

Three-digit Step 1 scores will soon be history, and many are nervous about their ability to differentiate applicants without these data.

These scores were never validated for this purpose and the unintended consequences of the Step 1 climate and application fever are ultimately bad for programs, applicants, and the public good. We should work within our specialty and across academic medicine to break the vicious cycle of overapplication. We should be transparent about what we are looking for in applicants and what we strive for in our graduates. Now is the time to define what we care about, and how we can assess and communicate those data. In coordination with each of the above suggestions, we should study how these interventions meet the needs of our programs, applicants, and their future patients.

CORRESPONDENCE: Address correspondence to Dr Benjamin N. Schneider, Robertson Collaborative Life Sciences Building, 2730 S Moody Ave, Portland, OR 97201. 503-346-4749. Schneibe@ohsu.edu.

References

- Invitational Conference on USMLE Scoring. United States Medical Licensing Examination. Accessed May 17, 2021. <https://www.usmle.org/incus/>
- Andolsek KM. One small step for Step 1. *Acad Med.* 2019;94(3):309-313. doi:10.1097/ACM.0000000000002560
- Chen DR, Priest KC, Batten JN, Fragoso LE, Reinfeld BI, Laitman BM. Student perspectives on the "Step 1 Climate" in preclinical medical education. *Acad Med.* 2019;94(3):302-304. doi:10.1097/ACM.0000000000002565
- McGaghie WC, Cohen ER, Wayne DB. Are United States Medical Licensing Exam Step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? *Acad Med.* 2011;86(1):48-52. doi:10.1097/ACM.0b013e3181ffacdb
- 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. doi:10.1002/hsr.2.161
- 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
- Mejicano G. Initial Summary Report and Preliminary Recommendations of the Undergraduate Medical Education to Graduate Medical Education Review Committee (UGRC). Coalition for Physician Accountability. Accessed July 1, 2021. <https://physicianaccountability.org/wp-content/uploads/2021/04/UGRC-Initial-Summary-Report-and-Preliminary-Recommendations-1.pdf>
- Announcements. United States Medical Licensing Examination. Accessed May 17, 2021. <https://www.usmle.org/announcements/?ContentId=309>
- Performance Data. United States Medical Licensing Examination. Accessed May 18, 2021. https://www.usmle.org/performance-data/default.aspx#2019_step-2-cs
- Step 2 CS Archives. Educational Commission for Foreign Medical Graduates - News. Accessed May 18, 2021. <https://www.ecfmg.org/news/category/step-2-cs/>
- ERAS Statistics. Association of American Medical Colleges. Accessed May 18, 2021. <https://www.aamc.org/data-reports/interactive-data/eras-statistics-data>
- Number of Medical Schools Using Selected Grading Systems. Association of American Medical Colleges. Accessed May 18, 2021. <https://www.aamc.org/data-reports/curriculum-reports/interactive-data/grading-systems-use-us-medical-schools>
- Alexander EK, Osman NY, Walling JL, Mitchell VG. Variation and imprecision of clerkship grading in U.S. medical schools. *Acad Med.* 2012;87(8):1070-1076. doi:10.1097/ACM.0b013e31825d0a2a
- Wasson LT, Cusmano A, Meli L, et al. Association between learning environment interventions and medical student well-being: a systematic review. *JAMA.* 2016;316(21):2237-2252. doi:10.1001/jama.2016.17573
- McDuff SGR, McDuff D, Farace JA, Kelly CJ, Savoia MC, Mandel J. Evaluating a grading change at UCSD school of medicine: pass/fail grading is associated with decreased performance on preclinical exams but unchanged performance on USMLE step 1 scores. *BMC Med Educ.* 2014;14(1):127. doi:10.1186/1472-6920-14-127
- Carmody JB, Rosman IS, Carlson JC. Application fever: reviewing the causes, costs, and cures for residency application inflation. *Cureus.* 2021;13(3):e13804. doi:10.7759/cureus.13804
- Apply Smart: Data to Consider When Applying to Residency. Association of American Medical Colleges. Accessed May 17, 2021. <https://students-residents.aamc.org/apply-smart-residency/apply-smart-data-consider-when-applying-residency>
- Results of the 2018 NRMP Program Director Survey. National Resident Matching Program. June 2018. Accessed May 16, 2021. <https://www.nrmp.org/wp-content/uploads/2018/07/NRMP-2018-Program-Director-Survey-for-WWW.pdf>
- Lee AH, Young P, Liao R, Yi PH, Reh D, Best SR. I dream of Gini: quantifying inequality in otolaryngology residency interviews. *Laryngoscope.* 2019;129(3):627-633. doi:10.1002/lary.27521
- Choudhary A, Makhoul AT, Ganesh Kumar N, Drolet BC. Impact of Pass/Fail USMLE Step 1 Scoring on the Internal Medicine Residency Application Process: a Program Director Survey. *J Gen Intern Med.* 2020. doi:10.1007/s11606-020-05984-y
- Residency Directory Index. American Academy of Family Physicians. Accessed May 17, 2021. <https://www.aafp.org/medical-education/directory/residency/search>
- Charting Outcomes in the Match: Senior Students of U.S. MD Medical Schools Characteristics of U.S. MD Seniors Who Matched to Their Preferred Specialty in the 2020 Main Residency Match 2nd Edition. Washington, DC: National Resident Matching Program; 2020. Accessed May 16, 2021. https://mk0nrmp3oyqui6wqfm.kinstacdn.com/wp-content/uploads/2020/07/Charting-Outcomes-in-the-Match-2020_MD-Senior_final.pdf
- Residency Explorer Tool: Home. Accessed May 17, 2021. <https://www.residencyexplorer.org/>
- Careers in Medicine. Association of American Medical Colleges. Accessed May 17, 2021. <https://www.aamc.org/cim/>
- FREIDATM AMA Residency & Fellowship Programs Database. Accessed May 17, 2021. <https://freida.ama-assn.org/>
- Texas STAR: Medical School Student Affairs - UT Southwestern, Dallas, TX. Accessed May 17, 2021. <https://www.utsouthwestern.edu/education/medical-school/about-the-school/student-affairs/texas-star.html>
- Welcome to Residency Navigator | Doximity. Accessed May 17, 2021. https://www.doximity.com/residency/?_remember_me_attempted=yes
- Family Medicine. Reddit feed. Accessed May 17, 2021. <https://www.reddit.com/r/FamilyMedicine/>
- Family Medicine | Student Doctor Network. Accessed May 17, 2021. <https://forums.studentdoctor.net/forums/family-medicine.37/>
- Hall K, Schneider B, Abercrombie S, et al. Hitting the ground running: medical student preparedness for residency training. *Ann Fam Med.* 2011;9(4):375. doi:10.1370/afm.1285