Perceived impact of USMLE Step 1 pass/fail scoring change on neurosurgery: program director survey

*Sakibul Huq, BS, Adham M. Khalafallah, MD, David Botros, BS, Adrián E. Jimenez, BS, Shrvika Lam, BS, Judy Huang, MD, and Debraj Mukherjee, MD, MPH

Department of Neurosurgery, Johns Hopkins University School of Medicine, Baltimore, Maryland

Neurosurgery is a small field with an increasingly competitive match in which the number of applicants outweighs the number of available positions.¹ In recent years, the average quantity of applications from each student has risen dramatically, making it a resource-intensive process for program directors (PDs) and associate program directors (APDs) to evaluate applications.²,³ Programs are subsequently tasked with narrowing down a large applicant pool to a small subset of qualified, mutually interested candidates to interview for 1–4 residency slots. In the setting of this increasingly competitive and taxing process, as well as a general shift toward pass/fail grading systems in preclinical education,⁴ there has been a premium placed on objective, standardized measures of performance to compare students from different institutions.⁵–⁷ Within neurosurgery, one of the most important and scrutinized of these metrics has been the United States Medical Licensing Examination (USMLE) Step 1 score.⁴ In prior published surveys, nearly half of responding neurosurgery programs have indicated that they have a minimum cutoff USMLE Step 1 score as a screening mechanism; perhaps unsurprisingly, the USMLE Step 1 score has been described as one of the most important parts of neurosurgery residency applications⁶ and the strongest predictor of match outcomes in the field.⁸ Use of the USMLE Step 1 score for resident selection has been the subject of some debate, given its questionable predictive value on meaningful performance and career outcomes.⁹,¹⁰ However, this practice has been justified by the need for a pragmatic filtering mechanism to narrow a burdensome number of applications, the exam’s possible predictive value on resident attrition and academic career trajectories,⁹ and its correlation with American Board of Neurological Surgery (ABNS) written board exam scores.¹¹ Nevertheless, the neurosurgical literature on factors predicting resident success is minimal,¹⁰ particularly compared with other specialties.¹²–¹⁴

In recent years, increasing attention has been brought to the notion that the USMLE Step 1 was originally designed as an evaluation system for medical licensure rather than a comparative achievement test well suited for secondary purposes (i.e., residency selection).¹⁵ Various stakeholders have expressed concerns about unintended consequences resulting from the increased importance of the exam. At one institution, greater than 70% of students reported anxiety about a mismatch between their institution’s preclinical curriculum and content that could appear on USMLE Step 1.¹⁵ This may explain student disengagement with institutional curricula in favor of “parallel curricula” driven in part by for-profit board preparation resources that, taken collectively, frequently cost students upwards of $1000.¹⁵–¹⁷ One study in the radiation oncology literature noted that screening protocols based on USMLE Step 1 score may promote reductions in resident diversity by disproportionately targeting applicants who are women, underrepresented in medicine, and/or older than 30 years.¹⁸ Others have criticized a lack of holistic review of residency applications and negative impact on student well-being due to the USMLE “Step 1 climate.”¹⁵,¹⁷,¹⁹

On February 12, 2020, following a national conversation about USMLE score reporting and the broader transition from undergraduate to graduate medical education,¹⁴,¹⁵,²⁰ the National Board of Medical Examiners (NBME) and Federation of State Medical Boards (FSMB) announced a series of changes to the USMLE sequence (https://www.usmle.org/inCus/#decision). This included changing USMLE Step 1 outcomes from a three-digit numeric score to a pass/fail scoring system, taking effect no earlier than January 1, 2022.
We found anecdotally that this change was an important topic of conversation among neurosurgeons and medical students contemplating how it would impact the field and how to prepare accordingly. Recognizing these concerns, we undertook a survey of all neurosurgery PDs and APDs in the United States to provide baseline perceptions regarding the value of USMLE Step 1 in neurosurgery resident selection, offer insight into the potential consequences of the scoring change on the field, and present recommendations for programs and applicants moving forward.

Methods

Survey Construction

A 23-question survey was built in Qualtrics (https://jhmi.co1.qualtrics.com/jfe/form/SV_eP1TzW3pEyA71HL) (Appendix A). In an effort to increase response rate, the survey was anonymous but contained the option to self-identify. There were 5 demographic questions, followed by 18 multiple choice, ranking, and free response questions addressing baseline USMLE Step 1 information about each program (2 questions), global perceptions about the change in scoring (4 questions), perceived predictive ability of USMLE Step 1 on neurosurgery resident performance (2 questions), perceptions of who would benefit or suffer from the change (2 questions), and impact on and advice for future neurosurgery applicants (8 questions). Where appropriate, 5-point Likert scores were used, with a neutral option intentionally kept in order to capture uncertainty. The survey was reviewed and piloted by the study team, which included one neurosurgery residency PD (J.H.).

Survey Dissemination

A roster of all neurosurgery PDs in the United States was created based on a directory from the American Association of Neurological Surgeons (https://www.aans.org/en/Trainees/Residency-Directory) and cross-referenced with the American Medical Association’s Fellowship and Residency Electronic Interactive Database (https://freida.ama-assn.org/Freida/#/). Where applicable, APDs were identified through manual review of program websites. The final list contained 119 PDs and 26 APDs. Personalized emails were sent to each PD and APD by the study team (A.M.K. and S.H.) on behalf of the two senior authors (D.M. and J.H.). A follow-up email was sent to all survey respondents 1 week later.

Data Collection and Statistical Analysis

Survey responses were captured in Qualtrics and then exported into a database in Microsoft Excel version 16.33 (Microsoft Corp.). Comparisons to answers between PDs and APDs were compared using Fisher’s exact test. Answer distributions between PDs and APDs were statistically insignificant for all questions except for one regarding “impact on my program” described more fully in Results, and thus PD and APD responses were subsequently pooled for descriptive analysis.

Likert scores were represented visually using histograms. Questions with categorical responses were represented with bar graphs. Free responses underwent minor copyediting for capitalization and punctuation but were otherwise reported verbatim.

Statistical analysis was performed using Stata (StataCorp LLC), and figures were created using GraphPad Prism version 8.3.1 (GraphPad).

Results

Survey Respondent Characteristics

A total of 60 complete survey responses were received after the initial email wave, and 15 additional responses were received following a second email 1 week later. In total, 59 PDs (50% response rate) and 16 APDs (62% response rate) participated, representing a 52% overall response rate (Table 1). The most common region represented was the Southeast (35%), followed by the Northeast (24%). All program sizes were represented, with the most common being 2 residents per year (47%). Responding programs were affiliated with a variety of medical schools, including traditionally high-matching programs.

Table 1. Characteristics of survey respondents

| Variable                                      | No. of Respondents (%)
|-----------------------------------------------|------------------------
| Survey respondents                            |                        |
| PD                                            | 59 (79)                |
| APD                                           | 16 (21)                |
| Regions represented                           |                        |
| Northwest                                     | 3 (4)                  |
| Southwest                                     | 10 (13)                |
| Northeast                                     | 18 (24)                |
| Southeast                                     | 26 (35)                |
| Midwest                                       | 17 (23)                |
| Other                                         | 1 (1)                  |
| Residents/yr                                  |                        |
| 1                                             | 16 (21)                |
| 2                                             | 35 (47)                |
| 3                                             | 19 (25)                |
| 4                                             | 5 (7)                  |
| Total applicants from home institution (last 3 yrs) |                |
| ≤3                                            | 34 (45)                |
| 4–9                                           | 23 (31)                |
| 10–15                                         | 8 (11)                 |
| No affiliated medical school                   | 10 (13)                |
| Routinely screen by Step 1?                   |                        |
| Always                                        | 58 (77)                |
| Sometimes                                     | 12 (16)                |
| Never                                         | 5 (7)                  |
| Average Step 1 score of incoming interns last 2 yrs |                |
| <230                                          | 1 (1)                  |
| 230–239                                       | 5 (7)                  |
| 240–249                                       | 41 (55)                |
| 250–259                                       | 26 (35)                |
| ≥260                                          | 2 (3)                  |

TABLE 1. Characteristics of survey respondents

Unauthenticated | Downloaded 09/21/23 03:25 AM UTC
and low-matching ones (≥ 1% and < 1% of the graduating class per year matching into neurosurgery, respectively).!

Baseline USMLE Step 1 Information

A majority of programs (77%) indicated that they always screen applications by USMLE Step 1 score. The average USMLE Step 1 score of incoming interns at the respondents’ programs over the past 2 years was most commonly between 240 and 249, followed by 250 and 259 (Table 1).

Global Perceptions of USMLE Step 1 Scoring Change

A large majority of PDs and APDs strongly or somewhat disagreed with the scoring change (79%) (Fig. 1A) and also strongly or somewhat disagreed that it would have a positive effect on neurosurgery (72%) (Fig. 1B). A majority strongly or somewhat disagreed that it would have a positive effect on their program (67%) (Fig. 1C) and on the field of medicine as a whole (68%) (Fig. 1D).

Perceived Impact on Respondents’ Trainees

A majority (64%) of respondents strongly or somewhat agreed that USMLE Step 1 is a strong predictor of resident success in their programs (Fig. 2A). Perceptions of the scoring change’s impact on residents’ ABNS written exam scores were mixed: 35% thought scores would de-
crease, 12% thought residents would pass later in training, 24% thought there would be no change, and none thought that residents would achieve higher scores or pass earlier in training. Many were uncertain (29%) (Fig. 2B).

Perceived “Winners and Losers” of USMLE Step 1 Scoring Change

PDs and APDs were prompted to select all applicant types they thought would benefit or suffer from the scoring change. Responses were mixed, although common themes emerged. A majority (57%) thought that students from prestigious schools would benefit, while students from less prestigious schools would suffer (60%) (Fig. 3). International medical graduates (IMGs; 29%) and residency PDs and APDs (52%) were also thought to largely suffer from the change. Survey respondents had the option to provide free-text responses to this question. A total of 8 respondents specifically typed out that “no one” would benefit from the change, while 1 respondent said “no one” would suffer. Other free-text answers largely reflected the

FIG. 2. Perceived impact of USMLE Step 1 on resident success and board exams. A: Agreement or disagreement that USMLE Step 1 is a strong predictor of resident success in the respondent’s individual program. B: Perceived impact that the scoring change will have on the respondent’s residents’ ABNS written exam scores.

FIG. 3. Perceived applicant types that may benefit or suffer from the USMLE Step 1 scoring change.
responses to the predefined options, echoing that students with less strong academic backgrounds, pedigrees, and resources might suffer relative to their counterparts.

Impact on and Advice for Applicants

A majority of PD and APD respondents (73%) thought that the USMLE Step 1 scoring change would increase the number of applicants to neurosurgery (Fig. 4A). There was some uncertainty regarding the potential impact on medical student research productivity; 49% said that research output would definitely or probably increase, 27% said it would definitely or probably not increase, and 24% were uncertain (Fig. 4B).

Nearly half of participating PDs and APDs indicated that Step 2 Clinical Knowledge (CK) would become “the new Step 1” for their program (47%), while many remained uncertain (39%) (Fig. 4C). Fewer suggested that Step 2 CK would not become the new Step 1 (15%).

A majority of PDs and APDs recommended that, if students were given the option to take USMLE Step 1 as scored or pass/fail in the future, they should take it for a score (81%) (Fig. 4D). A majority (67%) also reported plans to advise future applicants to take USMLE Step 2 CK before submitting their Electronic Residency Application Service (ERAS) application (Fig. 4E).

Free Responses

The last question of the survey was an optional free-response question: “How will this change affect your review of students’ applications?” A total of 39 survey respondents (52%) left a free response. There were 4 favorable responses (10%), 17 mixed/neutral responses (44%), and 18 unfavorable responses (46%). Full responses are included in Appendix B.

Discussion

Key Results and Interpretation

Our survey response rate was comparable to those of similar studies in the neurosurgical literature. Based
on the speed of survey completion, as well as the number of lengthy responses for the optional free-text question, we infer that this scoring change is an important topic at the forefront of many PDs’ and APDs’ minds. Program demographics represented in our survey, including size, location, affiliated institution, and average USMLE Step 1 score grossly approximate the overall landscape of neurosurgery residency programs in the United States. While there were notable exceptions, we found that PDs and APDs generally expressed dissatisfaction with the scoring change. They most frequently thought that it would have a negative effect on neurosurgery, their own programs, and medicine as a whole. Key themes emerged in free-response answers, including the notion that USMLE Step 1 is a reflection of work ethic and fund of knowledge, the predictive ability (or lack thereof) of USMLE Step 1 scores in foreshadowing resident success, concerns about removing an objective data point from the match process, worries about the downstream impact on resident ABNS board scores, and more.

A commonly cited consequence of the scoring change was the potential effect on groups that may benefit (students from prestigious schools), and those that may suffer (residency PDs and APDs, students from less prestigious schools, and IMGs). Responses regarding the latter 2 groups are intuitive, as PDs may be more likely to rely on medical school and mentor reputations in the absence of known “equalizers.” While components of the challenging match process for IMGs have been characterized, further study will be needed to determine the impact on and way forward for these groups adversely affected by pass/fail scoring. We suspect that residency PDs and APDs thought that they themselves would suffer due to the loss of a frequently used data point (and screening mechanism); they will now likely have to allocate more time and resources in order to meaningfully review a large number of applications. Despite these common themes, we note a great deal of heterogeneity in survey responses, which may reflect uncertainty in a time of change.

Similarly, there was uncertainty about whether the change would increase applicants’ research output. We speculate, based on prevailing trends in survey responses, that many programs may place additional weight on research experience, including a heightened expectation for undergraduate to graduate medical education. These efforts, as they relate to neurosurgery, would be aided by further research investigating pre-residency determinants of eventual residency and career performance.

Prior Literature on USMLE Step 1 in Neurosurgery

A common thread that has emerged in the literature and in our data involves the potential relationship between USMLE Step 1 and ABNS written exam scores. ABNS exam performance is considered an essential metric of residency success. It is also a factor for accreditation by the Accreditation Council for Graduate Medical Education (ACGME). Critics of the scored USMLE exam argue that previous test-taking ability is known to predict future test-taking ability but not other meaningful outcomes, including clinical performance. However, our data suggest that neurosurgery PDs and APDs do feel that USMLE Step 1 provides some predictive value on residency success, including on the ABNS written exam. Their perceptions align with one prior study reporting a direct correlation between USMLE Step 1 scores and ABNS written exam scores of graduating residents (Spearman correlation coefficient 0.7; coefficient of determination 0.429). In the same study, the median USMLE Step 1 score and ABNS scores were 249 and 505, respectively, and the ranges were 184–271 and 341–651 (all of which exceed the passing threshold), respectively. Our survey data imply that many PDs are in fact concerned about test-taking ability, perhaps in an effort to avoid issues with passing the ABNS written board exam (and consequent ACGME penalties).

Prior Literature on USMLE Step 1 in Other Specialties

These considerations regarding emphasis on USMLE Step 1 and potential correlations with subspecialty board exam scores are not unique to neurosurgery. In orthopedic surgery—another specialty frequently employing a USMLE Step 1 score cutoff—it has been shown that residents with higher subspecialty board scores had higher USMLE Step 1 scores. Similar patterns have been observed in surgery, obstetrics and gynecology, emergency medicine, and internal medicine. However, there are also bodies of literature that rebut these claims and instead highlight the predictive value of USMLE Step 2 CK, rather than USMLE Step 1, on later board exam success.

Anticipated Emphasis on USMLE Step 2 CK

We specifically inquired about the future role of USMLE Step 2 CK based on personal communications with colleagues who anticipated that this exam might fill the screening gap created by the loss of USMLE Step 1 scores. Our data did indeed reflect that many PDs and APDs view USMLE Step 2 CK as the new USMLE Step 1. This shift was anticipated by NBME and FSMB. While some may believe that no single standardized exam should receive such considerable weight, it has been suggested that if programs did want to see standardized exam data, USMLE Step 2 CK might be more suitable than USMLE Step 1 due to its more clinical nature.

USMLE Step 2 CK has traditionally played a smaller role in the neurosurgery application process relative to USMLE Step 1. Unsurprisingly, there is a paucity of literature describing its predictive value in neurosurgery. One institution found that USMLE Step 2 CK scores did not correlate with ABNS scores (while Step 1 scores did!), although USMLE Step 2 CK scores did correlate with attempts needed to score above 300 on the ABNS exam.

Future Directions for Applying to Neurosurgery Residency

The change in USMLE Step 1 scoring is ultimately one piece of a broader initiative to overhaul the transition from undergraduate to graduate medical education. These efforts, as they relate to neurosurgery, would be aided by further research investigating pre-residency determinants of eventual residency and career performance.

Based on our data and prior literature, we anticipate that emphasis on interviews, clerkship grades, and subinternship performance will remain constant or increase, while the existing emphasis on USMLE Step 1 may shift toward
USMLE Step 2 CK. Programs may continue to consider other experiences and accomplishments, leadership attributes, community engagement, diversity, and more.6,7 We speculate that some programs may also use this change as an opportunity to experiment with new assessment methods, such as emotional intelligence testing,35,36 dexterity challenges,37 virtual/augmented reality,11 personality assessments,38 tests of endurance and resilience,9,39 specialty-specific knowledge exams, prior life experiences, and more.

Impact for Programs and Applicants
Despite the uncertainty surrounding this change, we believe that our data provide an important starting point for programs and applicants moving forward. We hope that programs will benefit from and discuss responses to our survey. While we urge caution in extrapolating our data beyond their original intention of assessing PD and APD perceptions and advice, we do note several tangible findings. If future students pursuing neurosurgery are given the option to take USMLE Step 1 as scored or pass/fail, most PDs and APDs surveyed recommend taking the exam for a score. Similarly, many PDs and APDs note that USMLE Step 2 CK will become the new Step 1, and they consequently plan to advise applicants to take USMLE Step 2 CK before submitting ERAS. There is uncertainty regarding the impact of USMLE Step 1 scoring changes on research and publication productivity among applicants, although we conjecture that applicant publications may increase over time, heightening the research “arms race”40 that has been previously reported. Aspiring residents would do well to stay abreast of conversations within neurosurgery regarding the impact of the scoring change and, as always, should continue to consult with trusted mentors throughout the application process. Actionable recommendations resulting from this survey include advising future applicants to take USMLE Step 1 for a score if given the option, to take USMLE Step 2 CK prior to submitting ERAS, and increasing meaningful research productivity prior to applying to neurosurgery residency.

Limitations
The USMLE Step 1 scoring change is part of a broader paradigm shift in medical education, a full discussion of which is beyond the scope of this article. The scoring change is recent, with many parts in flux; neurosurgery PD and APD opinions may evolve over time and eventually differ from those noted in our study. Components of our study may become obsolete depending on future announcements by the USMLE and FSMB. Our survey had an incomplete response rate and may be subject to response bias, limiting its generalizability to all PDs and APDs. Our data may or may not fully reflect the diversity and distribution of all neurosurgeons’ opinions. Our analyses are descriptive in nature, and we advise caution in applying their general findings to individual programs. Given the anonymous nature of many survey responses, we did not perform subset analyses based on program characteristics. Despite these limitations, our data contain clear trends and themes that we believe will be actionable for both programs and applicants.

Conclusions
The change in USMLE Step 1 scoring from a three-digit score to pass/fail was made in good faith with intentions of improving the medical education environment and the transition from undergraduate to graduate medical education. Strongly held opinions about this change exist among different stakeholder groups, including within neurosurgery. We found that neurosurgery PDs and APDs were frequently dissatisfied with the change, uneasy about its potential downstream effects on resident success, and concerned about potentially negative consequences on the resident selection process. Many PDs and APDs indicated that USMLE Step 2 CK would fill the existing role of USMLE Step 1, and they plan to advise applicants to take USMLE Step 2 CK before submitting ERAS. There were notable exceptions to these trends, reflecting differences in opinion and a general sense of uncertainty about the impact of this change on the field. Our data provide a baseline “pulse” of neurosurgery PDs and APDs taken shortly after the announcement of the scoring change; its eventual impact on neurosurgery remains to be seen and warrants further investigation.

Acknowledgments
We thank all participating Program Directors and Associate Program Directors for their time and thoughtful responses to our survey. We additionally thank the Neuro-Oncology Surgical Outcomes Laboratory for helpful discussion and edits.

References
review and applicability to neurosurgical training. World Neurosurg. 2018;110:475–484.e10.