



## **AAIM Response to Recommendations of the Invitational Conference on USMLE Scoring (InCUS)**

Approved by the AAIM Board of Directors on July 25, 2019

***InCUS Recommendation #1 - Reduce the adverse impact of the current overemphasis on USMLE performance in residency screening and selection through consideration of changes such as pass/fail scoring.***

### ***AAIM Response:***

The Alliance for Academic Internal Medicine (AAIM) is the only organization that represents educators across the continuum of internal medicine education in the United States, including Clerkship Directors, Residency Program Directors, Fellowship Directors, Chairs of Medicine, and Administrators of Internal Medicine.

The Alliance's member groups concur that numerical score reporting of USMLE Step 1 is not an optimal tool to adequately assess medical students' readiness for training in internal medicine. We acknowledge that the emphasis on achieving high scores on this exam may be counterproductive to student learning priorities and contribute to mental health concerns such as depression, anxiety, and suicidal ideations. As the ability to earn interviews in a particular specialty or program may be limited by a score on this exam rather than any observed performance with patients or health care teams, it is not surprising that focus on a single multiple-choice exam may drive down engagement in other important aspects of medical education.

Although there are many stakeholders in this conversation, the Alliance submits that society and individual patients are ***our most important stakeholders***. Our approach to this issue should ensure the physicians we train are humanistic individuals who are skilled in all six core competencies and practice life-long learning. However, we acknowledge the challenges faced by program directors in screening the thousands of applications received by each program every year given the lack of alternative objective performance metrics other than the numeric USMLE score. Recruiting high caliber candidates is a priority for all. The annual match season burden takes a toll on residency program directors and may lead to a decline in their well-being and a possible factor in their already established high turnover rates.

Despite the areas of consensus stated above, the Alliance recognizes that all the views of our member organization may not be reconciled into a unified position. Below we provide the perspectives of our groups that are most impacted by this important conversation.

### ***Student Clerkship Directors***

Student mental health concerns such as depression, anxiety, and suicide, have been at the forefront of many discussions, and this is certainly an appropriate concern and area of focus. However, we believe it is crucial to hold a patient-centered approach that balances the altruistic tenets fundamental to our profession with the health of our learners. With the focus on the patient, the primary adverse impact on learners is disengagement from the medical school curriculum not deemed directly relevant to success on the USMLE Step 1. The current environment promotes learners to focus on the short-term acquisition of data for a single exam in deference to the development of clinical skills, clinical reasoning, professional identity, communication and interpersonal skills, and many other components vital to the development of competent physicians. We strongly believe this status quo is untenable, and stakeholders across the continuum of medical education will need to take deliberate and thoughtful action in transforming the educational environment to promote behaviors that lead to meaningful learning and development across all domains so that patients will receive the best care from medical school graduates.

Despite significant diversity in clerkship directors across the country, the majority believes a change to the USMLE Step 1 score reporting needs to be part of this transformation to an improved educational environment. There are pros and cons to either a pass/fail or categorical score reporting. Data suggests there is evidence that categorical scoring may be a valid approach (1-5) but may not yield the desired re-focus on other domains of learning. Regardless of how score reporting is modified, the primary issue is the over-reliance on this score for residency program and even fellowship program screening and selection. There is strong agreement that additional assessments of learners across multiple domains outside medical knowledge are needed as a counter-balance to legitimately de-emphasize the Step 1 score. The majority of clerkship directors are in favor of providing more detailed information about learner performance on the clerkship, and there is a desire to develop more standardized assessment language as well as reporting formats (e.g. standardized department letters) to promote greater transparency and comparability in learner ability across medical schools. There is a critical need for increased direct observations of learners by experienced clinician educators across the continuum of medical education with the recognition that these observations provide vital information about learners in a dynamic educational environment beyond what is captured on a multiple-choice exam (6).

### ***Residency and Fellowship Program Directors***

Program Directors would value more objective data on applicants' medical knowledge beyond the single USMLE Step 1 report. Numeric scores are a marker for progress in medical knowledge, and acquisition of medical knowledge is one of the competencies essential to success as a resident and as a practicing physician. Measures of medical knowledge are also important to program directors because they may predict board certification pass rate which is an ACGME criterion for program accreditation.

A decision to report USMLE as pass/fail will likely shift the "high stakes" focus to another test such as USMLE Step 2 or to new specialty-specific tests adding additional test preparation time, stress, and financial burden to our medical students. Alternative objective strategies for reporting of medical knowledge measures should be considered, including release of Step 2 CS domains, shelf exam scores, or the development of composite scores.

There are other potential unintended negative consequences of a decision to report Step 1 as pass/fail only. Programs, unable to adequately assess the thousands of applications received each year may

preferentially recruit from institutions that have greater name recognition, increase their reliance on personal connections, or require audition rotations – all of which will disproportionately harm students from disadvantaged backgrounds and schools. Programs might also begin to require secondary applications to try to limit the pool of applicants to those who are truly interested in them. Some may be driven to withdraw from the “all in” match to minimize constraints on recruiting resources.

A national standardized exam of medical knowledge is especially critical in the assessment of International Medical Graduates (IMG) who play a vital role in many residency programs, especially community-based programs. Metrics of performance in international medical school are often unhelpful in assessment of adequate residency preparation. Any change to USMLE scoring must address the needs of both US and IMG students, be considered in parallel with the addition of other objective metrics of performance, be phased in over a significant time frame to minimize disruption, and ensure maintenance of the high standards our patients expect of our trainees.

***InCUS Recommendation #2 - Accelerate research on the correlation of USMLE performance to measures of residency performance and clinical practice.***

***AAIM Response:***

There is already a substantial body of literature exploring the correlation between USMLE scores and a trainee’s performance on subsequent measures of medical knowledge such as In Training Exam scores or performance on board certification exams. There is scant data, however, that explores the relationship of USMLE scores to a trainee’s overall performance during residency and even less on outcomes in clinical practice.

Furthermore, there are concerns that the content of USMLE Step 1 may not be relevant to clinical practice and may require a complete reassessment of its structure or even consideration for consolidating it with Step 2.

However, the Alliance believes that further research on performance outcomes related to USMLE scores in isolation would not be useful. All stakeholders would be better served by exploring the relationship of alternative metrics beyond medical knowledge alone that correlate student performance to subsequent residency success and ultimately clinical practice. In particular, research on ways in which searchable/filterable assessment data in all competency areas could be integrated into reports on student achievement would be valuable along with study of the potential utility of approaches such as composite summative scores from USMLE data.

***InCUS Recommendation #3 - Minimize racial demographic differences that exist in USMLE performance.***

***AAIM Response:***

The Alliance supports the proposal to assess for racial bias in the USMLE tests; correlation to socioeconomic factors in general is equally important. There is also a need for more data about accessible test preparation resources for under-represented minorities and equity in resource allocation in medical school.

***InCUS Recommendation #4 - Convene a cross-organizational panel to create solutions for the assessment and transition challenges from UME to GME.***

***AAIM Response:***

The Alliance believes that a comprehensive plan to improve the UME/GME transition is imperative. The proposed UME-GME consortium should consider innovative approaches that might include reframing expectations on how schools should assess and report student achievement as well as a transformation of the entire match process. While some of these efforts (such as changes to the match process) may be ready for implementation in two to three years, transformative changes to assessment and reporting are likely to require pilots and rigorous outcome research over a minimum of three to five years.

- *Consideration identified by InCUS: Reducing the number of applications perceived by residency applicants as necessary to obtain a position*

Tools such as the “Residency Explorer” being piloted by AAMC/NBME and others are a potentially useful first step. It is likely that more robust information about residency programs than is currently available will be required in order to make the approach effective. The loss of an objective measure like a USMLE score could lead to further application inflation since students may no longer have the “guidance” that a numerical score offers regarding their competitiveness for certain residency programs or specialties.

The Alliance encourages exploration of creative mechanisms to decrease application inflation, such as alternative match strategies including “early decision,” multiple smaller match cycles, and a greater ability for applicants to indicate their preferences in ERAS (such as geography, type or size of program, etc.). More aggressive counseling of students by their medical schools may be required as well.

- *Consideration identified by InCUS: Improving Residency Program Directors’ ability to more holistically evaluate candidates*

Holistic assessment of candidates requires an ability to weigh student achievement in all six core competencies in addition to extracurricular, demographic and social factors that are not easily identified in ERAS. At present, there is only one of the six core competencies for which we have an objective tool that permits comparison across the spectrum of medical schools both in the United States and internationally. Validated metrics that provide insight into the other core competencies are essential to the goal of “holistic assessments.” These metrics will need to be available and searchable within ERAS in addition to other applicant characteristics mentioned above. The vast number of applications received by many residency programs preclude an individualized review of every applicant.

- *Consideration identified by InCUS: Improving the trust of school-based assessments for residency screening and selection*

Lack of trust among residency program directors has grown in parallel to the loss of transparency on the part of medical schools. This has occurred in part due to the move toward pass/fail reporting of student performance in the preclinical years and to “grade inflation” on clinical rotations. MSPEs and chair letters from some medical schools may provide scant core competency-related information on a student’s readiness for residency. Time spent in remediation or failures on shelf exams are not always disclosed. As a consequence, there may be crucial measures of a student’s performance that are not readily accessible or identifiable by program directors. A critical piece of the solution is more reporting

of direct observations of students by well-trained and trusted faculty. However, financial challenges within medical schools and work demands of faculty by health systems make allocation of faculty resources to these activities extremely difficult.

Measures that should be explored to improve assessment and reporting include core competency-based standards for MSPEs and structured requirements for chair letters (such as the “standardized letters of evaluation” used by emergency medicine). Content for these “SLOEs” could include the six core competencies as well as procedural skills, documentation skills (H&Ps, discharge and transfer summaries, medication reconciliation). Comparative reviews of medical schools by accrediting organizations such as LCME on their assessment practices and compliance of their reporting with recommended standards could be helpful to minimize variation across medical schools.

Many programs receive a large number of applications from international graduates who eventually make an important contribution to our country’s health care system, particularly in primary care and underserved communities. A differing approach to USMLE reporting on these applicants may be needed. Aggressive efforts by ECFMG to partner with international medical schools will also be important to allow more effective reviews of applicants from these settings regardless of whether the pass/fail decision applies to them or not.

#### ***Additional comments on solutions to the UME/GME transition challenges***

- Core Entrustable Professional Activities for Entering Residency frameworks (CEPAERs) have been developed in internal medicine to ensure that students are prepared for residency training. Such frameworks could conceivably be used to develop more holistic evaluation metrics and become the basis for a “Standardized Letter of Evaluation.” In addition to summative measures, they may be able to provide formative reports on student progress as well.
- Development of specialty-specific competency-based learning portfolios could provide balanced views of a student’s achievements. To be effective in the match process, these portfolios must include filterable and searchable metrics that allow a program to screen candidates.
- The rigorous competency-based rather than time-based “Education in Pediatrics Across the Continuum” (EPAC) model developed by some pediatric residencies in collaboration with medical schools may have potential application in other specialties. Pilots in efforts like this should be supported. For example, students interested in internal medicine could opt into the program at their school where in addition to standard assessments (shelf exam, clinical grades, etc.) they participate in a program of competency-based evaluations for the affiliated internal medicine programs to which they apply. (The AMA’s “Reimagining Residency Initiative” has recently awarded several grants to institutions that aim to undertake similar initiatives.)
- Expansion of educational opportunities that offer conditional acceptance into a university’s (or its affiliates’) GME program when accepted into their UME program may be desirable. Coupled with objective assessment reports (such as SLOEs or competency-based learning portfolios) this could be a highly effective way to reduce application volume for both students and programs in participating medical school partnerships.
- Medical school partnerships with residency programs, including community settings, create rotation opportunities for students. Students from schools who participate in these partnerships could have the option to participate in a “first match” with those programs. These “first match” events would occur earlier than the traditional match cycle and benefit students who have specific preferences by limiting the number of applications needed. Programs would likewise

benefit by already knowing the skillset of students who have rotated with them and may thus need to review fewer applicants.

- Modified match schedules including early decision opportunities and serial “SOAP” sessions may reduce the burden of application volume for both students and programs.