

## **Response to Survey Regarding the Summary Report and Preliminary Recommendations from the Invitational Conference on USMLE Scoring (InCUS)**

***Submitted on July 26, 2019***

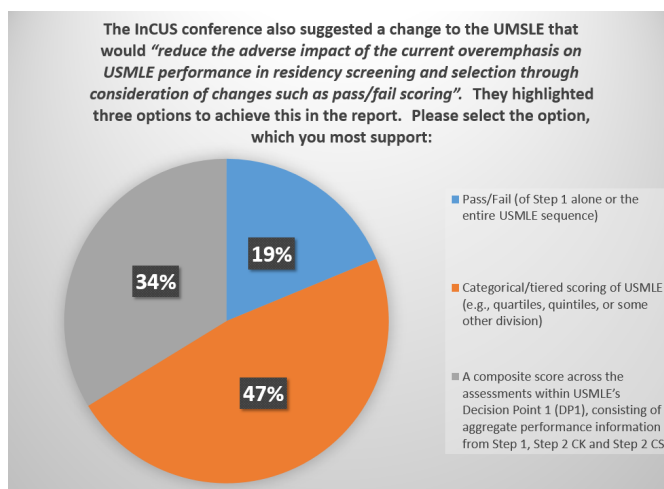
The APCCMPD represents 94% of all ACGME accredited fellowships in Pulmonary, Critical Care and Pulmonary/Critical Care Medicine. The APCCMPD was established in 1984. We have an active and vibrant membership, with a mission to provide leadership and support for Pulmonary, Critical Care Medicine and Pulmonary/Critical Care Medicine training program directors and their programs. As program directors for fellowship programs, we have an extra layer of assessment (residency performance) to shelter us from the USMLE as the main discriminator of uniform performance. However, the role of the USMLE certification system has implications for all levels of GME training.

In an effort to represent the viewpoints of all of the PCCM program directors, we conducted a survey to assess our members' opinions regarding the four recommendations of the InCUS Conference. We surveyed 217 APCCMPD Member program directors between July 8 and 21; we received 82 responses, for a response rate of 38%.

Our members generally agreed with three of the recommendations:

- 96% agreed with accelerating research on the correlation of USMLE performance to measures of residency performance and clinical practice
- 88% agreed with minimizing racial demographic differences that exist in USMLE performance
- 94% agreed with convening a cross-organizational panel to create solutions for the assessment and transition challenges from UME to GME, targeting an approved proposal, including scope/timelines by end of calendar year 2019

One recommendation brought a variety of viewpoints and responses (and many comments, which we have attached in the appendix to this letter). They were distributed as shown below:



There is not widespread agreement on this recommendation, a point that is likely not lost at this point on the organizers of InCUS.

As one of the largest and most active subspecialty program directors' associations in the United States, the APCCMPD would be happy to send a representative to any further discussions/deliberations regarding operationalizing the InCUS recommendations in the future.

# Results of APCCMPD Member Survey Regarding Invitational Conference on USMLE Scoring (InCUS) Report

## Survey Response:

Survey sent to 217 APCCMPD Member Pulmonary, PCCM, and CCM Program Directors

Survey period: July 8 - July 21, 2019

Response Rate: 37.8% (82)

## Survey Intro:

Dear Program Directors:

We are asking for your feedback in order to draft a response to the recent Invitational Conference on USMLE Scoring (InCUS). This survey is short (4 questions) and extremely important to the future structure of student and resident assessment in the United States.

Recently the USMLE convened a conference to discuss the future of USMLE testing. This conference was in response to growing calls for a change in the system, which rewards student focus on this high-stakes test at the expense of other, less testable medical training activities; and the fact that the test is being used by many programs as a discriminator for residency (and fellowship) selection- a function that test was not designed to perform.

The InCUS Conference took place March 11-12, 2019. In short, the conference had four recommendations. They are:

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.
2. Accelerate research on the correlation of USMLE performance to measures of residency performance and clinical practice.
3. Minimize racial demographic differences that exist in USMLE performance.
4. Convene a cross-organizational panel to create solutions for the assessment and transition challenges from UME to GME, targeting an approved proposal, including scope/timelines by end of calendar year 2019.

The last three recommendations have consensus of most medical educators. The first is less widely agreed upon.

The full summary report is a very easy read and can be found here: [https://www.usmle.org/pdfs/incus/incus\\_summary\\_report.pdf](https://www.usmle.org/pdfs/incus/incus_summary_report.pdf).

They are asking for a call for input from both individuals and stakeholder societies. We are one of those societies. We are gathering as much input from our members as possible to include in the response from the APCCMPD.

This is a complicated topic that will have far-reaching consequences for medical education and may change dramatically the landscape of medical trainees in the United States, so it is worth putting a few minutes into reading the full summary report (I read it in about 15 minutes). There is a LOT more that goes into this topic than is found in the report, and it does not present opinions on each side of the issue. Anyone interested can email me at [gbosslet@iu.edu](mailto:gbosslet@iu.edu) and I can share resources on the topic.

Please click here for a VERY BRIEF 4-question survey that we will use to inform our response to the InCUS committee, and we will share the response once it is submitted.

Sincerely,

Gabriel Bosslet, MD  
President-Elect, APCCMPD

## 1. Please indicate whether you agree or disagree with the following recommendations of InCUS:

### 1a. Accelerate research on the correlation of USMLE performance to measures of residency performance and clinical practice

Response	N
Agree	79 (96.3%)
Disagree	3 (3.7%)

**Comments from respondents that Agreed:**

- In my day we had Internal Medicine boards to determine whether or not we had learned a lot. We also during Internship had to do a Hx and PExam in front of some attending (I am sure I was at best average since I get nervous). That was it. It think I did O.K. Now I have published a lot, got a Ph.D. from Columbia (after all my fellowships) in Epi in my spare time, etc. Passed my Public Health boards, etc. I am sure I was not that impressive as Intern or Resident and I did not talk a lot.
- This is arguably the most objective information in a candidate's application. While I'm not sure how much correlation there will be would like to know what is there?

**Comments from respondents that Disagreed:**

*No comments recorded.*

**1b. Minimize racial demographic differences that exist in UMSLE performance**

Response	N
Agree	72 (87.8%)
Disagree	10 (12.2%)

**Comments from respondents that Agreed:**

- Some differences may be due to differences in educational opportunities that have resulted in lower performance. Others may be due to bias in the exam. We should minimize bias in the exam, but not try to use the exam to adjust for differences in educational attainment even if they are due to societal inequities.
- Not sure what this means.
- I'm not sure I understand how to go about doing this.
- This is tricky; if the racial demographic differences represent important outcomes, then we need to measure those differences. If the differences are artifacts of their background that do not bear on future performance, we should eliminate them.

**Comments from respondents that Disagreed:**

- I don't understand this statement. Knowledge, a race blind concept, is required to develop clinical competence (the application of that knowledge to the care of patients). Some assessment of knowledge is necessary.
- I think the broader goal is to avoid bias and to minimize racial differences in achievement and competence. Minimizing racial differences in USMLE performance is only desirable if the USMLE scores inaccurately reflect differences in achievement and competence.

**1c. Convene a cross-organizational panel to create solutions for the assessment and transition challenges from UME to GME, targeting**

**an approved proposal, including scope/timelines by end of calendar year 2019**

Response	N
Agree	77 (93.9%)
Disagree	5 (6.1%)

**Comments from respondents that Agreed:**

- Not sure what UME means but sounds good.
- But that sounds like an ambitious deadline

**Comments from respondents that Disagreed:**

*No comments recorded.*

- 2. The InCUS conference also suggested a change to the UMSLE that would “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”. They highlighted three options to achieve this in the report. Please select the option, which you most support:**

<b>Response</b>	<b>N</b>
Pass/Fail (of Step 1 alone or the entire USMLE sequence)	15 (18.75%)
Categorical/tiered scoring of USMLE (e.g., quartiles, quintiles, or some other division)	38 (47.5%)
A composite score across the assessments within USMLE’s Decision Point 1 (DP1), consisting of aggregate performance information from Step 1, Step 2 CK and Step 2 CS	27 (33.75%)

**Comments from respondents “Pass/Fail (of Step 1 alone or the entire USMLE sequence)”:**

- I think that a pass/fail for step 1 is appropriate with possibly a quartile range followed by a score for step 2 as currently done.
- Pass/Fail across the entire sequence
- Pass/fail would be acceptable to me.
- If a change was made to pass/fail, there must be corresponding attempts to allow GME programs to better differentiate applicants and efforts made to reduce the exorbitant number of schools students are applying to, and applications that must be reviewed by GME programs.
- This will require more work from UME deans to ensure that they are accurately representing the individual student's skill set.
- I think these scores are often given too much weight in decisions for fellowship acceptance.

**Comments from respondents “Categorical/tiered scoring of USMLE (e.g., quartiles, quintiles, or some other division)”:**

- Due to the many external forces described in the full report, program directors need a quick way to distinguish among students.
- Choice 2 seems the most reasonable. Admittedly, I'm not intimately familiar with data on correlation between USMLE testing performance and future performance as a physician. If the push to go away from standardized testing is to reduce discomfort by applicants, I'm less receptive to that. If this is part

- of a program to ensure we're testing the right things to optimally predict the best clinicians in the fairest way possible, I'm all for that.
- I don't think USMLE with scores should go away; it is one measure of one aspect of performance. It is the disproportionate emphasis that is potentially problematic.
  - Pass/Fail is not helpful when two candidates are similar otherwise
  - I would strongly advise against pass/fail. It is important to have granular data than that, and I do see at least anecdotally for our fellows that scores correlate with GME performance.
  - The USMLE does provide some information for those at risk of failing boards in the future. I would like to see some categorical or tiered scoring to help me decipher those at risk - mainly so if we consider an applicant with low scores or in the lowest tier that we would be prepared to track their in training exam scores and if they remain low start early intervention plans and early individualized learning plans for board review. So in essence, the score is not the only thing that matters, but identifying low tier performers can help PDs prepare better. So yes the score matters, but mostly for the bottom 20% or so.
  - I am highly in favor of doing something to address the issues with USMLE Step 1. As both a program director and an educator within the foundations phase of our medical school curriculum, I have a fairly good perspective on the role of this test and the impact it is having on medical education. From the standpoint of running a major block in our undergraduate curriculum, I can state that the current overemphasis on Step 1 is starting to ruin this phase of training. The level of stress among our students is very high and it is having an adverse effect on the way they engage (or... in many cases... don't engage) with their main classes. Many are quite content to forego core course resources we provide and focus instead on commercial prep resources or spend time that should be spent learning our course material on doing Step 1 prep, even though Step 1 may be well over a year away. Then there is, of course, the endless complaining about how our curriculum doesn't adequately cover Step 1 material because we don't mention every factoid that shows up in First Aid for the Boards, which the students look at as the bible, even though it is arguably one of the worst resources one could use if they truly wanted to understand material needed to care for patients. I could go on... but the bottom line is that it has become a total mess. From the standpoint of a program director, I need some metrics that allow me to know if an applicant has the skills to succeed in our program. Board scores do provide that to some extent but to me it doesn't make a huge difference if someone has a 240 vs. a 250 on their exam, whereas I'd like to know an applicant had a barely passing score of 198. For that reason, I can see how having quartiles or some other broad categories would be helpful. The problem as most have likely recognized is that we lack other good metrics of how students do. Letters of recommendations are typically useless. Many schools have rampant grade inflation and Dean's letters don't convey much in the way of negative or accurate information that might give one pause about an applicant. There is too much reading between the lines. For that reason, working to identify some other metrics that won't breed ridiculous levels of stress among students would be useful.
  - it should not be pass/fail in my opinion - the scores themselves do matter and help predict future test taking skills

**Comments from respondents "A composite score across the assessments within USMLE's Decision Point 1 (DP1), consisting of aggregate performance information from Step 1, Step 2 CK and Step 2 CS":**

- A composite score may not be helpful as an assessment for fellowship performance any more than Step 3 scores are but provide some measure of medical knowledge. Test scores can not be utilized in isolation to judge fitness for residency or fellowship.
- I do not use absolute scores alone for selection. I typically look to see how many failures they have rather than scores. If have multiple failures, this is of concern about test taking or other skills that may impact passing boards in the future.
- I do not feel strongly that any of the above options would be a positive change. I do feel there is an overemphasis by medical students on their USMLE scores and that this is because of the importance placed on it by residency programs, and that it is likely that the amount of time devoted to preparing for the USMLE should probably be utilized elsewhere if the goal is to train the best physicians. However, I do not know what the solution here is, as I do believe as a GME program you should have the availability to as much information as possible, and to mask scores as "pass/fail" or quartiles is not a positive thing. Instead, the focus should probably be on extensive research into what does produce the highest quality of training, which could end up being the current system though I would suspect not. If programs simply knew what produced the best trainee and correlated with success in residency the most strongly then they would likely not factor USMLE scores into their decision making as strongly. But the decision over what attributes to consider and how to weight them when choosing residents should I think ultimately be left up to the individual programs, perhaps with reporting of the research results to help guide programs decision making.

**Comments from respondents who left this question blank:**

- Keep as is. None of the above
- Don't agree