HOD ACTION: Recommendations in Council on Medical Education Report 4 adopted and the remainder of the report filed.

REPORT 04 OF THE COUNCIL ON MEDICAL EDUCATION (A-23)
Decreasing Bias in Assessments of Medical Student Clinical Clerkship Performance (Resolution 309-A-22, Resolve 2)
(Reference Committee C)

EXECUTIVE SUMMARY

Per a directive from the House of Delegates (HOD), the American Medical Association (AMA) has been asked to study and report back on the impact of two-interval clinical clerkship grading systems on residency application outcomes, clinical performance during residency, and bias.

This report defines two-interval grading (binary pass/fail with no other hierarchical ranking) and notes existing policy regarding pass/fail in non-clinical curricula. This report offers the theoretical background for the importance of pass/fail grading within competency-based medical education and formative assessment. It also highlights the competitive medical education system and the ongoing demand for summative assessment and ranking, particularly due to applicant selection challenges impacting both learners and program directors.

Due to a need for additional future research combining the multiple factors indicated by the HOD’s directive, this report instead summarizes research on each relevant topic individually, including significant variability and bias within clinical clerkship grading; existing recommendations toward improving reliability in this area; background on how grading system data is collected; proportions of two-interval pass/fail grading systems across medical schools; and current overall research on residency application outcomes, longitudinal performance tracking, and bias issues. This report emphasizes the diverse factors and potential unintended consequences that may arise when hierarchy is eliminated in one area of medical education and ranking decisions are shifted to other areas.

This report proposes reaffirmation of current AMA policy and offers new recommendations that continue to encourage work in support of the Coalition for Physician Accountability’s Undergraduate Medical Education-Graduate Medical Education Review Committee “Recommendations for Comprehensive Improvement of the UME-GME Transition”; encourage and support UME institutions’ investment in a) developing more valid, reliable, and unbiased summative assessments for clinical clerkships, including development of assessors’ awareness regarding structural inequities in education and wider society, and b) providing standardized and meaningful competency data to program directors; encourage institutions to publish information related to clinical clerkship grading systems and residency match rates, with subset data for learners from varied groups, including those that have been historically underrepresented in medicine or may be affected by bias; and encourage UME institutions to include grading system methodology with grades shared with residency programs.
HOD ACTION: Recommendations in Council on Medical Education Report 4 adopted and the remainder of the report filed.

REPORT OF THE COUNCIL ON MEDICAL EDUCATION

CME Report 04-A-23

Subject: Decreasing Bias in Assessments of Medical Student Clinical Clerkship Performance (Resolution 309-A-22, Resolve 2)

Presented by: John P. Williams, MD, Chair

Referred to: Reference Committee C

Resolution 309-A-22, “Decreasing Bias in Evaluations of Medical Student Performance,” was introduced by the Medical Student Section at the 2022 Annual Meeting of the American Medical Association (AMA). While Resolve 1 was adopted into AMA Policy D-295.307, Resolve 2 was referred for study. The referred clause asked that our AMA:

1. Study the impact of two-interval clinical clerkship grading systems on residency application outcomes and clinical performance during residency.

Testimony emphasized the current difficulty in accessing data needed to inform such a study and work underway via the AMA ChangeMedEd initiative toward longitudinal tracking. Testimony also highlighted challenges faced by program directors, the delicate balance of wanting more data versus ensuring unbiased data, and equity concerns regarding current grading models and diverse learners. Reference Committee C and the House of Delegates (HOD) felt that these concerns warranted further study. This report is in response to this referral.

BACKGROUND

Clinical Clerkships and Two-Interval Grading

In clinical clerkships, medical students are immersed in learning experiences involving direct patient care and application of clinical sciences.¹ This comprises both core and elective rotations, beginning in the third year of medical school, and with significant variability between clerkship experiences based on seasonal infectious disease cycles, electives chosen, and other considerations.

Two-interval grading refers to grading structures with only two options, either pass or fail, though these grades may also be accompanied by narrative information. Two-interval pass/fail grading is distinct from generalized pass/fail grading insofar as some pass/fail grading structures offer opportunities for grading with honors and other hierarchies, such as “high pass,” as opposed to the binary pass/fail. While AMA Policy H-295.866, “Supporting Two-Interval Grading Systems for Medical Education,” encourages “the establishment of a two-interval grading system in medical colleges and universities in the United States for the non-clinical curriculum,” current policy does not address clinical curriculum.

¹ Competency-Based Medical Education and the “Growth Mindset”

© 2023 American Medical Association. All rights reserved.
The current rationale for two-interval grading centers around learner trust and growth within the move toward competency-based medical education, or CBME (see also AMA policy D-295.317). Specifically, for medical education to focus on outcomes via a developmental approach, vulnerability for learners must be acknowledged and institutional culture must demonstrate trustworthiness, as learner gaps and needs may only be addressed if acknowledged rather than hidden due to performance pressure. Thus, two-interval pass/fail frees the learner from striving for a specific performative grade, allowing more transparency around gaps. This redirects focus to effectively meeting required competencies (passing) after careful consideration of areas for improvement, rather than concealing difficulties to rank higher. Equity between learners is complex and not inherently achieved by grading system changes alone, as discussed in later sections. Biases related to race, gender, disability, or other factors exist in a wider societal structure, and interventions require a multi-pronged approach. However, even highly rigorous and non-biased assessments would drive undesired behaviors (concealment versus transparency toward growth) if graded or ranked. Nonetheless, larger medical education and societal structures currently create a demand for ranking, as discussed below.

Applicant Selection Challenges

A significant concern regarding possible elimination of tiered rankings in clerkship grades involves the increasing number of residency applications and growing challenges for programs when selecting from an overwhelming number of candidates. The United States Medical Licensing Examination® (USMLE®) Step 1 examination’s shift to pass/fail in January 2022 sparked concerns in this regard from residency program directors: a study of internal medicine program directors found that, in the absence of graded Step 1 examination scores, program personnel would be increasingly likely to weight such variables as ranked clerkship grades, Step 2 exam scores, personal knowledge of the applicant, and audition electives; respondents also expressed the belief that osteopathic applicants may potentially be further disadvantaged. Data regarding actual impact is unknown because not enough time has passed. Without an overhaul of the application process and infrastructure supportive of the time necessary for holistic review of applicants or transition away from competition-based processes (i.e., randomization via lottery), eliminating rankings in certain areas may indeed pose challenges. However, clerkship grades are an unreliable measure for evaluating residency applicants and challenged by inconsistencies and bias, as further described in the next section.

Unreliability and Variability in Clinical Clerkship Grades

Despite perceptions of their importance in selecting program applicants, clinical clerkship grades are generally found to be inconsistent and unreliable. In one study, most students believed that clerkship grades were unfair and that being liked by specific supervisors most influenced grading; further data confirms the detachment of clerkship grades from useful assessment criteria. One study noted that most medical schools used a four-tier system of fail, pass, high pass, or honors, but all defined these words subjectively and inconsistently, even within the same programs; this variability across schools and even within programs poses a challenge to accurate stratification of applicants. U.S. News & World Report Top 20 medical schools were also more likely to disproportionately assign the highest clerkship grade to a higher percentage of students than other medical schools, even though these schools were also less likely to implement grade comparison at all. Clerkship grades often suggest the “illusion of objectivity,” despite no standard approach to assigning grades or rank, flawed data not based on actual observations, high stress for students, and time-based grading paradigms that promote inequities.

Equity and Diversity Concerns Within Medical School Assessment
Beyond concerns of general unreliability, equity and diversity concerns also arise within clinical clerkship assessment. One 2018 study (which defined “underrepresented in medicine” narrowly as students from the racial or ethnic groups Black, Latina/o/x, Native American, and Alaska Native) demonstrated differences in clerkship director ratings that consistently favored non-underrepresented students, and while these differences were small, they created an amplification cascade later in the educational experience, compounding challenges already faced by these students due to structural racism. Another 2019 study demonstrated that, even after accounting for confounding variables, grades were more likely to favor white students above both underrepresented and non-underrepresented students of color. Even prior to grading itself, the training environment and overall social environment already hinders students from marginalized racial/ethnic groups, depleting cognitive resources and interfering with learning, such that even with more “objective” grading standards, societal bias already creates an inequitable environment for learning. Finally, while research that addresses the specific topic of clinical clerkship assessment for other marginalized identities/experiences is limited, learners are subjected to systemic biases in many realms, such as LGBTQ issues, socioeconomic status, and disability.

**DISCUSSION**

Course grades perform two purported functions: giving students a summative evaluation of their course performance and providing a standardized means of communicating student performance to third parties. Grades should be distinguished from formative assessments, which are focused on improving student learning. As a summative evaluation, grades should be based on valid and reliable data and contain sufficient information to be useful to students and third parties, with attention to the ways larger systemic bias and inequitable assignment of merit influences even otherwise reliable data. Current data demonstrated above indicates significant reliability concerns in current grading systems.

Little data exists to demonstrate the impact of two-interval clinical clerkship grading on residency application outcomes and clinical performance during residency, and even less data that includes analysis by race, gender, socioeconomic class, disability, or other relevant demographics. This report seeks to split the question into its various components, provide background on how some data is collected and reported, offer currently available research, and offer suggestions on how this data might be gathered in the future.

**Current Data and Challenges Regarding Pass/Fail in Clinical Clerkships**

Much current research suggests that two-interval pass/fail grading systems improve learner well-being in the preclinical years, and academic performance remains similar, with an increased opportunity for a reduction of stress and less competitive learning environment. Proponents of CBME also generally advocate to reframe two-interval pass/fail as two-interval “only pass/not yet pass” and to utilize criterion-referenced assessment such that learners will pass in time. Support for CBME is inherently linked to removing hierarchical grading structures in all aspects of medical education.

Data around usage of pass/fail grading systems in clinical clerkships is collected by the Liaison Committee on Medical Education (LCME) for allopathic schools and by the American Association of Colleges of Osteopathic Medicine (AACOM) for osteopathic schools, but few analyses of impact exist.

The LCME’s files indicated the following data for each portion of the curriculum:
### LCME Part II Totals: Type of Grading System Used (2019-2020)

<table>
<thead>
<tr>
<th>Grading system</th>
<th>Required clinical clerkships</th>
<th>Fourth-year selectives/sub-internships</th>
<th>Electives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass-fail</td>
<td>11</td>
<td>32</td>
<td>84</td>
</tr>
<tr>
<td>Honors-pass-fail</td>
<td>26</td>
<td>27</td>
<td>21</td>
</tr>
<tr>
<td>Honors-high pass-pass-fail</td>
<td>85</td>
<td>68</td>
<td>57</td>
</tr>
<tr>
<td>Numerical grade</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Letter grade</td>
<td>24</td>
<td>19</td>
<td>10</td>
</tr>
<tr>
<td>Other</td>
<td>13</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

### LCME Part II Totals: Type of Grading System Used (2020-2021)

<table>
<thead>
<tr>
<th>Grading system</th>
<th>Required clinical clerkships</th>
<th>Fourth-year selectives/sub-internships</th>
<th>Electives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass-fail</td>
<td>24</td>
<td>37</td>
<td>92</td>
</tr>
<tr>
<td>Honors-pass-fail</td>
<td>25</td>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>Honors-high pass-pass-fail</td>
<td>81</td>
<td>72</td>
<td>54</td>
</tr>
<tr>
<td>Numerical grade</td>
<td>7</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Letter grade</td>
<td>20</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>10</td>
<td>12</td>
</tr>
</tbody>
</table>

### LCME Part II Totals: Type of Grading System Used (2021-2022)

<table>
<thead>
<tr>
<th>Grading system</th>
<th>Required clinical clerkships</th>
<th>Fourth-year selectives/sub-internships</th>
<th>Electives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass-fail</td>
<td>20</td>
<td>37</td>
<td>90</td>
</tr>
<tr>
<td>Honors-pass-fail</td>
<td>26</td>
<td>27</td>
<td>18</td>
</tr>
<tr>
<td>Honors-high pass-pass-fail</td>
<td>82</td>
<td>73</td>
<td>55</td>
</tr>
<tr>
<td>Numerical grade</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Letter grade</td>
<td>19</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

As seen above, within required clinical clerkships, two-interval pass/fail accounted for only about seven percent of grading systems in 2019-2020 and 14 percent in 2020-2021, with a slight decline in 2021-2022 to 20 schools out of 155, or about 13 percent. In fourth-year medical selective rotations, two-interval pass/fail grading systems accounted for about 21 percent in 2019-2020, 22 percent in 2020-2021, and 23 percent in 2021-2022. Elective clerkships were more likely to be two-interval pass/fail than other clerkships, as this accounted for about 47 percent of grading systems in both 2019-2020 and 2020-2021, and about 49 percent in 2021-2022.

The most recent AACOM data available showed that 28 schools used pass/fail to grade required clinical clerkships, while 21 schools used pass/fail for elective/selective grading. However, this data reflects multi-interval pass/fail variants including honors and does not indicate which, if any, use two-interval grading. Looking closer, a 2020 study of transcripts indicated that osteopathic medical schools’ grading system distribution in clinical years was 59.5 percent honors, 29.7.
percent letter grade, and 10.8 percent other systems. Only one of the 37 osteopathic medical schools participating in this study used two-interval pass/fail systems without tiered indicators such as “high pass” in the clinical years. This study demonstrated the variability between grading systems, both within and between allopathic and osteopathic schools, and the rarity of two-interval pass/fail in clerkship years.

Given limited implementation of two-interval pass/fail, research on the impact of this grading mechanism is even more limited. In 2021, faculty from one institution responded to the elimination of tiered clerkship grades with optimism for well-being and the learning environment, as well as hesitations, such as lack of readiness for hierarchies in later educational structures and concerns about the residency selection process. Students in a different 2021 qualitative study shared that implementation of two-interval pass/fail core clerkship grading, in combination with enhanced formative feedback, resulted in benefits to intrinsic motivation, increased ability to seek feedback and improvements, lowered stress, and perceived mitigation of equity concerns. However, this perceived mitigation was not confirmed with outcomes-based data, nor are these perceptions disaggregated by respondent demographics. In another study from 2022, transitioning to two-interval clinical clerkship grades with enhanced feedback was related to moderate to large improvements in students’ perceptions of grading and the learning environment, toward that of "mastery-oriented learning" rather than performative behavior. Simultaneously, deeper learner concerns around bias in evaluators and inequitable narrative summaries remained.

Current Clinical Clerkship Recommendations for Eliminating Grading Bias

Grappling with known equity issues, the Alliance for Academic Internal Medicine’s 2021 report, “Aiming for Equity in Clerkship Grading: Recommendations for Reducing the Effects of Structural and Individual Bias” indicated the scarcity of evidence-based resources for eliminating bias in clinical clerkship grading. Using a socioecological model, the authors suggest several possible interventions for further implementation and study, including but not limited to faculty development, non-normative competency-based grading, and refraining from standardized cut-off scores to designate honors in grading, though recommendations do not explicitly suggest removal of honors within grading.

Also regarding systemic bias concerns in grading, the Coalition for Physician Accountability’s Undergraduate Medical-Education-Graduate Medical Education Review Committee recommended the following in 2021: “To eliminate systemic biases in grading, medical schools must perform initial and annual exploratory reviews of clinical clerkship grading, including patterns of grade distribution based on race, ethnicity, gender identity/expression, sexual identity/orientation, religion, visa status, ability, and location (e.g., satellite or clinical site location), and perform regular faculty development to mitigate bias. Programs across the UME-GME continuum should explore the impact of bias on student and resident evaluations, match results, attrition, and selection to honor societies.”

In 2022, Russo et al. demonstrated the bias present within clinical clerkship grades and suggested that two-interval pass/fail grading as one component may mitigate the impact of bias, though it will not eliminate bias itself. “Shifting to a competence-based assessment model will give the learner multiple opportunities over time to demonstrate their mastery of skills and knowledge, thereby reducing the power of a single biased assessment.”

Due to the complexities of bias within clinical clerkship grading systems, the need for innovation is clear, but additional evidence is required to understand whether two-interval pass/fail grading effectively addresses these challenges.
Current Data and Challenges Regarding Pass/Fail and Residency Application Outcomes

When considering how to understand the impact two-interval pass/fail in clinical clerkships may have on residency application outcomes, especially regarding bias and equity, one must first consider what data is needed, and how this data is currently collected.

Match results from applications to residency programs are reported in aggregate by both the National Resident Matching Program (NRMP) and by medical schools. While it might be possible to determine some correlation between the schools that use two-interval pass/fail in clinical clerkships and their aggregate Match results, all other confounding factors would need to be considered, including other aspects of the school and all other determining factors considered in applications, both on larger-scale and individual learner levels. When also considering learner diversity and any potential impacts of bias, information would need to be disaggregated into multiple categories, such as race, ethnicity, disability, gender identity, sexual orientation, socioeconomic status, and more. Some of this information is currently collected in aggregate ways, such as through the Association of American Medical Colleges’ (AAMC) Medical School Graduation Questionnaire, but not all aspects of bias are addressed; these results are not tied to specific application outcomes or individuals due to privacy concerns. Further insights on two-interval pass/fail grading systems’ impact on bias in residency application outcomes would require the limited number of schools with two-interval pass/fail in clinical clerkship to study this specifically, comparing archival data before two-interval grading with current data, and with a student population large enough to ensure confidentiality for participants. This data would then need to be published. Multiple schools would need to achieve this to provide sufficient numbers to allow for comparison between institutions, and between allopathic versus osteopathic programs.

Outside of medical schools, in a related field, a 2019 study found that for Doctor of Pharmacy students within advanced pharmacy practice experiences, there was little statistical difference in residency match rates between applicants with two-interval pass/fail grades and tiered grades to assess clinical experiences. However, pharmacy education exists in a different context than medical education, and extrapolations cannot necessarily be made.

As discussed in earlier sections, it is well-known that bias is a concern in residency application outcomes. A 2019 study found no statistically significant differences in residency application outcomes in one institution when pre-clinical grades are pass/fail, but no such research currently exists for clinical clerkships. Current research merely indicates that clinical clerkship grades overall are not useful for ranking residency applications. A 2021 study suggested that receiving honors in clinical clerkship grading contributed to matching into the applicant’s top five programs in OB/GYN where honors were available, but that minority and male students were less likely to receive honors, suggesting further need for research into grading disparities.

Residency programs must currently create a rank list of applicants for admission, and in numerous specialties and for many residency programs, the number of qualified applicants to be evaluated greatly exceeds the number of positions available. Medical school clerkship grades are among several factors used by residency programs to determine the ranking of applicants. Though these grades are currently unreliable, as discussed above, conversion to two-interval pass/fail grading systems for clerkships without other interventions will require residency programs to weigh other data points more heavily when reviewing applications, such as recommendation letters or perceived medical school reputation. It is uncertain if these alternative factors are more valid or subject to less bias than clerkship grades, and the impacts on diverse student groups are still uncertain. While further knowledge is gathered, medical schools can invest in improving their grading systems to
decrease bias, provide transparency to residency programs regarding their grading system methodologies, and invest in methods of providing more useful information to residency programs.

**Current Data and Challenges Regarding Longitudinal Tracking into Residency**

Additional challenges arise when seeking data on how two-interval pass/fail grades in clinical clerkship and bias may impact residency performance outcomes. For longitudinal tracking into residency, current data sources include feedback from program directors to school deans, either sent by the school or coordinated by the AAMC Resident Readiness Survey. Information published by the AAMC does not track comparatively across schools, and even comparative school data would need to account for confounding factors, not merely each school’s clinical clerkship grading system. As with application outcome challenges, residency performance outcome challenges also include the need to collect and disaggregate demographic information for learners without violating learner privacy.

There is currently no pre-existing research to draw from on the direct impact of two-interval pass/fail clinical clerkship grading systems on residency performance outcomes, with or without the consideration of equity and bias. One 2019 study that begins to approach the topic is a meta-analysis of program directors’ perceptions of residency performance among residents from schools using two-interval pass/fail versus tiered clerkship grading, which found no significant difference in perceptions of overall performance between these groups. However, perceptions of performance do not inherently translate to actual actions taken nor actual criterion-referenced performance and carry the additional limitation of reflecting only on those who were already admitted into residency.

Some progress has been made on overall development of longitudinal tracking, though not related to these topics specifically. For instance, the AMA Accelerating Change in Medical Education Consortium created a personalized graduate profile for 32 medical schools, addressing three core questions of workforce, clinical exposure, and quality of care. This may serve as “a proof of concept” for further research into the topics of this report. The Accreditation Council for Graduate Medical Education (ACGME) also collects milestone data by specialty, but this data is not currently compared with data on pass/fail grading systems in clinical clerkships. There is also evidence to suggest that racial and ethnic biases may impact milestone levels. For instance, a 2022 study in pediatric programs found race and gender disparities in assessments of trainees in residency programs.

**RELEVANT AMA POLICY**

The AMA has extensive policy related to grading systems and mitigating bias in medical education. Some examples are as follows:

- **D-200.985**, “Strategies for Enhancing Diversity in the Physician Workforce,” recommends that residency/fellowship programs use holistic assessments of applicants that take into account the diversity of preparation and the variety of talents that applicants bring to their education.
- **D-310.945**, “Mitigating Demographic and Socioeconomic Inequities in the Residency and Fellowship Selection Process,” encourages medical schools, medical honor societies, and residency/fellowship programs to work toward ethical, equitable, and transparent recruiting processes, which are made available to all applicants.
• D-295.988, “Clinical Skills Assessment During Medical School,” works with appropriate stakeholders to assure the processes for assessing clinical skills are evidence-based and most efficiently use the time and financial resources of those being assessed.

• D-295.317, “Competency Based Medical Education Across the Continuum of Education and Practice,” continues to study and identify challenges and opportunities and critical stakeholders in achieving a competency-based curriculum across the medical education continuum and other health professions that provides significant value to those participating in these curricula and their patients.

• D-295.318, “Competency-Based Portfolio Assessment of Medical Students,” develops pilot projects to study the impact of competency-based frameworks on student graduation, the residency match process, and off-cycle entry into residency programs.

• D-295.963, “Continued Support for Diversity in Medical Education,” works with appropriate stakeholders to commission and enact the recommendations of a forward-looking, cross-continuum, external study of 21st century medical education focused on reimagining the future of health equity and racial justice in medical education.

• D-295.307, “Decreasing Bias in Evaluations of Medical Student Performance,” works with appropriate stakeholders to promote efforts to evaluate methods for decreasing the impact of bias in medical student performance evaluation as well as reducing the impact of bias in the review of disciplinary actions.

• D-295.983, “Fostering Professionalism During Medical School and Residency Training,” continues to study the clinical training environment to identify the best methods and practices used by medical schools and residency programs to foster the development of professionalism.

• H-350.979, “Increase the Representation of Minority and Economically Disadvantaged Populations in the Medical Profession,” supports increasing the representation of minorities in the physician population.

• D-295.322, “Increasing Demographically Diverse Representation in Liaison Committee on Medical Education Accredited Medical Schools,” studies medical school implementation of the Liaison Committee on Medical Education (LCME) Standard IS-16 and share the results with appropriate accreditation organizations and all state medical associations for action on demographic diversity.

• H-295.866, “Supporting Two-Interval Grading Systems for Medical Education,” works with stakeholders to encourage the establishment of a two-interval grading system in medical colleges and universities in the United States for the non-clinical curriculum.

These policies are listed in full detail in Appendix A.

SUMMARY AND RECOMMENDATIONS

Fair and equitable assessment in medical school improves career opportunities for medical students and benefits the public which deserves a more diverse physician workforce. Grades are one form of summative assessment of student performance, and summative assessment should provide third parties with important information about learner competencies and readiness. Current research demonstrates that despite the weighting of clinical clerkship grades in residency applicant selection, these grades are currently inconsistent, unreliable, and biased. Thus, medical schools should invest in developing valid, reliable, unbiased, and informative assessments for clerkships. Two-interval pass/fail clinical clerkship grading systems are rare in allopathic and osteopathic schools alike, and understanding their impacts on residency application outcomes and clinical performance during residency, especially from an equity lens, will require significant effort by researchers and medical education stakeholders. Efforts toward longitudinal tracking in general are
also still in the early stages. However, both AMA policy and pre-existing research do support
overall well-being and learning environment improvements related to two-interval pass/fail grading
systems in the pre-clinical years. Not all schools have implemented this grading structure, and
continued encouragement to do so is warranted.

Learners, including learners experiencing systemic oppression in one or many domains, are not a
monolith, and the need for nuance is paramount as these issues are addressed. Inequity in clinical
clerkship assessment may be one symptom of the wider culture of systemic bias as well as a
reflection of the current learning environment of competition within medical education. The
“bottleneck” within the popularity of certain specialties over others also amplifies the competitive
environment. Without a greater shift within medical education’s values, or without tending to the
entire landscape of medical education, modifying one component piece may send varying intended
and unintended ripple effects outwards to the other components of learner assessment—potentially
shifting pressure and bias from one area to another, and having unknown and heterogeneous effects
on a variety of learners. It is difficult to assess only one piece of the overall system to reflect an
understanding of overall equity in assessment, and even more challenging to correct only one piece
of a much wider puzzle. Despite these challenges, further gathering of data and the exploration of
innovations across the continuum of medical education is beneficial, with an emphasis on attention
to the needs of unique populations, especially those that are underrepresented in medicine or
experience bias. An evidence base for best practices and interventions can and should be gathered.
Strategies must focus on the wider whole, including evaluating the benefits and challenges of
moving to a competency-based system with equity at the forefront, rather than a time-based and
competitive system.

The Council on Medical Education therefore recommends that the following recommendations be
adopted in lieu of Resolution 309-A-22, Resolve 2, and the remainder of this report be filed:

That our American Medical Association (AMA):

1. Continue to encourage work in support of the Coalition for Physician Accountability’s
   Undergraduate Medical Education-Graduate Medical Education Review Committee
   “Recommendations for Comprehensive Improvement of the UME-GME Transition.”
   (Directive to Take Action)

2. Encourage and support UME institutions’ investment in a) developing more valid, reliable,
   and unbiased summative assessments for clinical clerkships, including development of
   assessors’ awareness regarding structural inequities in education and wider society, and b)
   providing standardized and meaningful competency data to program directors. (New HOD
   Policy)

3. Encourage institutions to publish information related to clinical clerkship grading systems
   and residency match rates, with subset data for learners from varied groups, including
   those that have been historically underrepresented in medicine or may be affected by bias.
   (New HOD Policy)

4. Encourage UME institutions to include grading system methodology with grades shared
   with residency programs. (New HOD Policy)

5. Reaffirm the following policies:

   • D-295.307, “Decreasing Bias in Evaluations of Medical Student Performance”
H-295.866, “Supporting Two-Interval Grading Systems for Medical Education”

D-295.317, “Competency Based Medical Education Across the Continuum of Education and Practice”

D-295.318, “Competency-Based Portfolio Assessment of Medical Students”

Fiscal note: $1,000

APPEndix A: RELEVANT AMA POLICY

Strategies for Enhancing Diversity in the Physician Workforce D-200.985
1. Our AMA, independently and in collaboration with other groups such as the Association of American Medical Colleges (AAMC), will actively work and advocate for funding at the federal and state levels and in the private sector to support the following: (a) Pipeline programs to prepare and motivate members of underrepresented groups to enter medical school; (b) Diversity or minority affairs offices at medical schools; (c) Financial aid programs for students from groups that are underrepresented in medicine; and (d) Financial support programs to recruit and develop faculty members from underrepresented groups.

2. Our AMA will work to obtain full restoration and protection of federal Title VII funding, and similar state funding programs, for the Centers of Excellence Program, Health Careers Opportunity Program, Area Health Education Centers, and other programs that support physician training, recruitment, and retention in geographically-underserved areas.

3. Our AMA will take a leadership role in efforts to enhance diversity in the physician workforce, including engaging in broad-based efforts that involve partners within and beyond the medical profession and medical education community.

4. Our AMA will encourage the Liaison Committee on Medical Education to assure that medical schools demonstrate compliance with its requirements for a diverse student body and faculty.

5. Our AMA will develop an internal education program for its members on the issues and possibilities involved in creating a diverse physician population.

6. Our AMA will provide on-line educational materials for its membership that address diversity issues in patient care including, but not limited to, culture, religion, race and ethnicity.

7. Our AMA will create and support programs that introduce elementary through high school students, especially those from groups that are underrepresented in medicine (URM), to healthcare careers.

8. Our AMA will create and support pipeline programs and encourage support services for URM college students that will support them as they move through college, medical school and residency programs.

9. Our AMA will recommend that medical school admissions committees and residency/fellowship programs use holistic assessments of applicants that take into account the diversity of preparation and the variety of talents that applicants bring to their education with the goal of improving health care for all communities.

10. Our AMA will advocate for the tracking and reporting to interested stakeholders of demographic information pertaining to URM status collected from Electronic Residency Application Service (ERAS) applications through the National Resident Matching Program (NRMP).

11. Our AMA will continue the research, advocacy, collaborative partnerships and other work that was initiated by the Commission to End Health Care Disparities.

12. Our AMA opposes legislation that would undermine institutions' ability to properly employ affirmative action to promote a diverse student population.

13. Our AMA will work with the AAMC and other stakeholders to create a question for the AAMC electronic medical school application to identify previous pipeline program (also known as
Mitigating Demographic and Socioeconomic Inequities in the Residency and Fellowship Selection Process D-310.945

Our AMA will: 1. encourage medical schools, medical honor societies, and residency/fellowship programs to work toward ethical, equitable, and transparent recruiting processes, which are made available to all applicants.

2. advocate for residency and fellowship programs to avoid using objective criteria available in the Electronic Residency Application Service (ERAS) application process as the sole determinant for deciding which applicants to offer interviews.

3. advocate to remove membership in medical honor societies as a mandated field of entry on the Electronic Residency Application Service (ERAS)—thereby limiting its use as an automated screening mechanism—and encourage applicants to share this information within other aspects of the ERAS application.

4. advocate for and support innovation in the undergraduate medical education to graduate medical education transition, especially focusing on the efforts of the Accelerating Change in Medical Education initiative, to include pilot efforts to optimize the residency/fellowship application and matching process and encourage the study of the impact of using filters in the Electronic Residency Application Service (ERAS) by program directors on the diversity of entrants into residency.

5. encourage caution among medical schools and residency/fellowship programs when utilizing novel online assessments for sampling personal characteristics for the purpose of admissions or selection and monitor use and validity of these tools.

Clinical Skills Assessment During Medical School D-295.988

1. Our AMA will encourage its representatives to the Liaison Committee on Medical Education (LCME) to ask the LCME to determine and disseminate to medical schools a description of what constitutes appropriate compliance with the accreditation standard that schools should "develop a system of assessment" to assure that students have acquired and can demonstrate core clinical skills.

2. Our AMA will work with the Federation of State Medical Boards, National Board of Medical Examiners, state medical societies, state medical boards, and other key stakeholders to pursue the transition from and replacement for the current United States Medical Licensing Examination (USMLE) Step 2 Clinical Skills (CS) examination and the Comprehensive Osteopathic Medical Licensing Examination (COMLEX) Level 2-Performance Examination (PE) with a requirement to pass a Liaison Committee on Medical Education-accredited or Commission on Osteopathic College Accreditation-accredited medical school-administered, clinical skills examination.

3. Our AMA will work to: (a) ensure rapid yet carefully considered changes to the current examination process to reduce costs, including travel expenses, as well as time away from educational pursuits, through immediate steps by the Federation of State Medical Boards and National Board of Medical Examiners; (b) encourage a significant and expeditious increase in the number of available testing sites; (c) allow international students and graduates to take the same examination at any available testing site; (d) engage in a transparent evaluation of basing this examination within our nation's medical schools, rather than administered by an external organization; and (e) include active participation by faculty leaders and assessment experts from U.S. medical schools, as they work to develop new and improved methods of assessing medical student competence for advancement into residency.

4. Our AMA is committed to assuring that all medical school graduates entering graduate medical education programs have demonstrated competence in clinical skills.
5. Our AMA will continue to work with appropriate stakeholders to assure the processes for assessing clinical skills are evidence-based and most efficiently use the time and financial resources of those being assessed.

6. Our AMA encourages development of a post-examination feedback system for all USMLE test-takers that would: (a) identify areas of satisfactory or better performance; (b) identify areas of suboptimal performance; and (c) give students who fail the exam insight into the areas of unsatisfactory performance on the examination.

7. Our AMA, through the Council on Medical Education, will continue to monitor relevant data and engage with stakeholders as necessary should updates to this policy become necessary.

Competency Based Medical Education Across the Continuum of Education and Practice D-295.317

1. Our AMA Council on Medical Education will continue to study and identify challenges and opportunities and critical stakeholders in achieving a competency-based curriculum across the medical education continuum and other health professions that provides significant value to those participating in these curricula and their patients.

2. Our AMA Council on Medical Education will work to establish a framework of consistent vocabulary and definitions across the continuum of health sciences education that will facilitate competency-based curriculum, andragogy and assessment implementation.

3. Our AMA will continue to explore, with the Accelerating Change in Medical Education initiative and with other stakeholder organizations, the implications of shifting from time-based to competency-based medical education on residents' compensation and lifetime earnings.

Competency-Based Portfolio Assessment of Medical Students D-295.318

1. Our AMA will work with the Association of American Medical Colleges, the American Osteopathic Association and the Accreditation Council for Graduate Medical Education, and other organizations to examine new and emerging approaches to medical student evaluation, including competency-based portfolio assessment.

2. Our AMA will work with the NRMP, ACGME and the 11 schools in the AMA's Accelerating Change in Medical Education consortium to develop pilot projects to study the impact of competency-based frameworks on student graduation, the residency match process and off-cycle entry into residency programs.

Continued Support for Diversity in Medical Education D-295.963

Our AMA will: (1) publicly state and reaffirm its stance on diversity in medical education; (2) request that the Liaison Committee on Medical Education regularly share statistics related to compliance with accreditation standards IS-16 and MS-8 with medical schools and with other stakeholder groups; (3) work with appropriate stakeholders to commission and enact the recommendations of a forward-looking, cross-continuum, external study of 21st century medical education focused on reimagining the future of health equity and racial justice in medical education, improving the diversity of the health workforce, and ameliorating inequitable outcomes among minoritized and marginalized patient populations; and (4) advocate for funding to support the creation and sustainability of Historically Black College and University (HBCU), Hispanic-Serving Institution (HSI), and Tribal College and University (TCU) affiliated medical schools and residency programs, with the goal of achieving a physician workforce that is proportional to the racial, ethnic, and gender composition of the United States population.

Decreasing Bias in Evaluations of Medical Student Performance D-295.307

Our AMA will work with appropriate stakeholders to promote efforts to evaluate methods for decreasing the impact of bias in medical student performance evaluation as well as reducing the impact of bias in the review of disciplinary actions.
Fostering Professionalism During Medical School and Residency Training D-295.983

(1) Our AMA, in consultation with other relevant medical organizations and associations, will work to develop a framework for fostering professionalism during medical school and residency training. This planning effort should include the following elements: (a) Synthesize existing goals and outcomes for professionalism into a practice-based educational framework, such as provided by the AMA's Principles of Medical Ethics. (b) Examine and suggest revisions to the content of the medical curriculum, based on the desired goals and outcomes for teaching professionalism. (c) Identify methods for teaching professionalism and those changes in the educational environment, including the use of role models and mentoring, which would support trainees' acquisition of professionalism. (d) Create means to incorporate ongoing collection of feedback from trainees about factors that support and inhibit their development of professionalism.

(2) Our AMA, along with other interested groups, will continue to study the clinical training environment to identify the best methods and practices used by medical schools and residency programs to fostering the development of professionalism, to include an evaluation of professional behavior, carried out at regular intervals and employing methods shown to be valuable in adding to the information that can be obtained from observational reports. An ideal system would utilize multiple evaluation formats and would build upon educational experiences that are already in place. The results of such evaluations should be used both for timely feedback and appropriate interventions for medical students and resident physicians aimed at improving their performance and for summative decisions about progression in training.

Increase the Representation of Minority and Economically Disadvantaged Populations in the Medical Profession H-350.979

Our AMA supports increasing the representation of minorities in the physician population by: (1) Supporting efforts to increase the applicant pool of qualified minority students by: (a) Encouraging state and local governments to make quality elementary and secondary education opportunities available to all; (b) Urging medical schools to strengthen or initiate programs that offer special premedical and precollegiate experiences to underrepresented minority students; (c) urging medical schools and other health training institutions to develop new and innovative measures to recruit underrepresented minority students, and (d) Supporting legislation that provides targeted financial aid to financially disadvantaged students at both the collegiate and medical school levels. (2) Encouraging all medical schools to reaffirm the goal of increasing representation of underrepresented minorities in their student bodies and faculties. (3) Urging medical school admission committees to consider minority representation as one factor in reaching their decisions. (4) Increasing the supply of minority health professionals. (5) Continuing its efforts to increase the proportion of minorities in medical schools and medical school faculty. (6) Facilitating communication between medical school admission committees and premedical counselors concerning the relative importance of requirements, including grade point average and Medical College Aptitude Test scores. (7) Continuing to urge for state legislation that will provide funds for medical education both directly to medical schools and indirectly through financial support to students. (8) Continuing to provide strong support for federal legislation that provides financial assistance for able students whose financial need is such that otherwise they would be unable to attend medical school.
Increasing Demographically Diverse Representation in Liaison Committee on Medical Education
Accredited Medical Schools D-295.322
Our AMA will continue to study medical school implementation of the Liaison Committee on
Medical Education (LCME) Standard IS-16 and share the results with appropriate accreditation
organizations and all state medical associations for action on demographic diversity.

Supporting Two-Interval Grading Systems for Medical Education H-295.866
Our AMA will work with stakeholders to encourage the establishment of a two-interval grading
system in medical colleges and universities in the United States for the non-clinical curriculum.
REFERENCES


