

Should Medical Grades be Standardized? Improving the Resident Selection Process

ERIKA SHIMAHARA, MA

Each fall directors and administrators of medical residency programs comb through hundreds of applications to find the best candidates. As the 2008 National Residency Match Program survey of program directors reported, grades acquired in medical school were among the top 5 most frequently cited criteria for selecting candidates for interviews.¹

However, relying on grades can be problematic given that there is often little to no overlap in what those grades mean across schools. For the 2009 academic year, the Stanford University School of Medicine Department of Otolaryngology received over 250 applications from medical students who hailed from 96 (74% of Association of American Medical Colleges [AAMC]) US medical schools. This showed that schools used a wide variety of grading schemes. Complicating matters, the letter representation of a grade across schools varied; the letter “C” could stand for average at schools that use an A/B/C letter grade system, but it could also stand for “Credit” (University of Southern California and University of Arizona), “Commendable” (Ohio State), or “Completed” (Wake Forest), and the letter “H” was not necessarily the highest grade at a school that awarded Honors (Harvard awards “High Honors,” or “HH”).

To collect application materials from individual candidates, Stanford uses the Electronic Residency Application System, the web-based program of the AAMC, used by the majority of residency programs. The Electronic Residency Application System facilitates candidates’ ability to apply to multiple programs. However, for reasons that are not clear, grading keys—the legends that explain the grading system at a particular school—often fail to be included with the grade transcripts, and their absence makes it cumbersome to determine (1) which grade at a particular institution was the highest (eg, “A” or “A+” or “Honors” or “High Honors”) and (2) the relative value of the grades given (when a “P” is on a transcript, is the institution’s grading system Pass/Fail or Honors/Pass/Fail?) Of the applications our department received for the 2009 academic year, grading keys were absent from 41% (39) of US

institutions. In many cases, it is possible to obtain an approximation of the type of grading scale used from information included in the Medical School Performance Evaluation (MSPE), or by reviewing an applicant’s entire transcript to see what other grades were earned. But many schools have distinct grading structures for clerkships and basic science courses, and the MSPE is not submitted until November 1 of each year, which is later than the application deadline for many schools. The lack of clear and consistent explanations about grading schemes suggests that residency application reviewers may guess what grades symbolize, use other selection factors in the assessments of candidates, or spend significant time trying to decipher what a given grade represents.

Grades are only one of the selection criteria considered by residency programs. The most cited criterion is the United States Medical Licensing Examination score (USMLE). The lack of clarity in grades may result in an overreliance on USMLE scores as a measure of a candidate’s academic strength because these are the only objective criteria contained in the applications. Given the uncertain predictability of USMLE scores in physician performance, its use as the major determinant in selecting residents may be misguided. The other top selection criteria are hardly better at foretelling a candidate’s doctoring potential.² The MSPE (second-most cited criterion) lays out the student’s performance in coursework, and despite guidelines instituted by the AAMC in 2002 that it be evaluative and include comparative data, it continues to resemble a recommendation letter advocating for a student, rather than an analysis of the student’s performance, particularly in lower-performing students.³ The personal statement (third-most cited criterion) can be exaggerated and ghostwritten to portray a desirable image, and letters of recommendation (fourth-most cited criterion) are generally one-sided and often indistinguishable from one another.⁴ In contrast to these other factors, institutional grading schemes should enable us to assess student performance with relatively simple analysis.

The TABLE shows examples of the types of grading systems used by the 96 US institutions submitting transcripts to the Stanford Department of Otolaryngology on behalf of applicants for the academic year 2009.

The variety in grading schemes notwithstanding, it is worth noting that many of the grades representing performance that is average (eg, “C”) or worse (eg, “D” or “F”) never appear on the transcript because, in actuality, medical schools simply do not award mediocre to poor grades. In the same vein, no student gets an “F” at a

Erika Shimahara, MA, is an Education Specialist in the Department of Otolaryngology at the Stanford University School of Medicine.

Corresponding author: Erika Shimahara, MA, Stanford University School of Medicine, 801 Welch Road, Stanford, CA 94305, 650.725.6500, eshimahara@stanford.edu

Received November 24, 2009; revision received February 23, 2010; accepted March 29, 2010.

DOI: 10.4300/JGME-D-09-00098.1

TABLE	EXAMPLES OF GRADING SYSTEMS OF INSTITUTIONS SUBMITTING TRANSCRIPTS TO THE STANFORD UNIVERSITY DEPARTMENT OF OTOLARYNGOLOGY RESIDENCY	
Examples of Grades	Institution(s)	
Honors (H), Pass (P), Fail (F) grading system ^a		
H, P, F	University of Vermont, UCSF, UCSD, State University of New York at Stonybrook, Texas Tech University	
H, COM (commendable), S (satisfactory), U (unsatisfactory)	Case Western University	
H, E (excellent), S, N (no pass)	University of Minnesota	
H, H-, P, F	University of Iowa	
H, NH (near honors), S, M (marginal), F	Oregon Health Sciences University School of Medicine	
H, HP (honors pass), P/\ ("pass after remediation"), P* ("pass repeated course")	Northwestern University	
H, HS, S+, S, U, P ("advanced standing")	State University of New York at Buffalo	
H, HP, P, MP (marginal pass), F	Baylor College of Medicine	
H, HS (honors satisfactory), S	Drexel University College of Medicine	
HH (High honors), H, S, U	Harvard Medical School	
C (completed), H, P, LP (low pass)	Wake Forest University School of Medicine	
CR (credit), H, NC (no credit/unsatisfactory), Inc (incomplete/unsatisfactory)	University of Hawaii	
H, HP, P, U for clerkship grades, except in the courses called "Surgery B" and "Surgery C," which are graded either pass or fail	Yale School of Medicine	
A/B/C letter grading system ^b		
A, AB, B	University of Wisconsin	
A, B, C, D, F	Uniformed Services University	
A, B, C, D, Honors, P, F	Rosalind Franklin University	
A, B, P	LSU Health Sciences Center-Shreveport	
A, B, S	Emory University School of Medicine	
A, B+, B, C	UT Southwestern Medical Center	
A+, A, A-, B+, B, B-, C+	University of Virginia School of Medicine	
A+, B, C, D, F	Medical College of Georgia	
A, B+, B, C+, C, D, E	College of Medicine-University of Florida	
Pass/fail (P/F), ^c grade point system, ^d or other ^e		
HON (honors), EXCEL (excellent), GOOD, MAR (marginal competence)	Thomas Jefferson University	
HN (honors), LC (letter of commendation), S+ (satisfactory with honors), S (satisfactory)	University of Missouri-Columbia School of Medicine	
O (outstanding), Adv (advanced), PR (proficient), S	University of Illinois at Chicago	
S, P, F, C	University of Arizona	
SU (superior), HS (high satisfactory), SA (satisfactory)	University of Kansas	

Abbreviations: UCSF, University of California at San Francisco; UCSD, University of California at San Diego; LSU, Louisiana State University; UT, University of Texas.

^a Sixty-one schools (64%) used either an H/P/F grading system or a variant.

^b Sixteen institutions (17%) used a variant of the A/B/C grading system in which an "A" or "A+" was the highest grade.

^c Five schools (5%) including the University of Alabama, University of California at Los Angeles, University of Chicago, University of Michigan, and Stanford University used a P/F system.

^d Three institutions (3%)—Medical University of South Carolina College of Medicine, University of Medicine & Dentistry of New Jersey/RW Johnson Medical School, and University of Medicine & Dentistry of New Jersey/New Jersey Medical School—used a grade point system in which the highest grade was represented by "4."

^e Eleven schools (11%) used either combinations of the aforementioned systems or systems that were undecipherable.

Pass/Fail school. (Pass/Fail systems arguably provide the least information about a student's academic performance because they also do not indicate the margin by which students are passing.) Grade inflation is the obvious explanation for the absence of low marks,⁵ which raises another question: should the *distribution* of grades across institutions also be standardized? That is, should not the percentage of each grade awarded be the same for all institutions? If our goal is to be able to compare the academic performance of students from different schools, and in some cases of the same school—especially Pass/Fail schools, the answers are obvious.

The application data from the Stanford program resonates with survey results reported by Takayama et al⁶ covering 121 AAMC/Liaison Committee on Medical Education teaching hospitals, in which 76 schools used a variant of the Honors/Pass/Fail system, 22 used a letter grade system, and the remainder used such systems as Outstanding/Advanced/Proficient, or Pass/Fail, or no interpretable system at all. The concluding recommendation of that article was that the AAMC should standardize grading.

Surely, standardizing grading is not a cure-all that by itself will enable us to choose the best candidates for residency. First, individual medical schools will have different standards by which they award grades, just as they have different standards for admitting students. Second, residency programs will continue to receive applications from non-US graduates whose transcripts will reflect their own institutions' standards. Third, we should continue to use a variety of criteria to make decisions about residency candidates to allow us a fuller and fairer picture of the student. Finally, and more broadly, the issue of standardized grading raises other issues that cannot be separated from questions surrounding the very purpose of assessment in medical school. For example, can grades accurately measure the type of learning needed to become a superb doctor? Can standardized grades be compatible with a learning environment that fosters collegiality, cooperation, and learning for the sake of learning? These fundamental questions, addressing both theoretical and operational

implications, should be part and parcel of any discussion of standardizing grades.

If we are to believe Cooke et al⁷ who say that “Rigorous assessment has the potential to inspire learning, influence values, reinforce competence, and reassure the public,” the concept of grades has an essential purpose in medical education. Moreover, if we are to assume that the goal of assessment is, as Epstein⁸ argues, “to optimize the capabilities of all learners and practitioners by providing motivation and direction for future learning, to protect the public by identifying incompetent physicians, and to provide a basis for choosing applicants for advanced training,” then medical school grading, as it exists now, needs standardization. Without it we continue with our conventional, misguided practices. We should instead aspire to being held accountable for using grades to select physicians-in-training. Standardizing grades will make them more meaningful to the student, residency programs, and ultimately the public who is eventually under the care of the candidate who we chose, at least in part, on the basis of their grades.

References

- 1 National Resident Matching Program, Data Release and Research Committee. Results of the 2008 NRMP Program Director Survey. Available at: www.nrmp.org/data/programresultsbyspecialty.pdf. Accessed January 28, 2009.
- 2 Lee AG, Golnik KC, Oetting TA, et al. Re-engineering the resident applicant selection process in ophthalmology: a literature review and recommendations for improvement. *Surv Ophthalmol*. 2008;53(2):164–176.
- 3 Shea JA, O'Grady E, Morrison G, Wagner BR, Morris JB. Medical student performance evaluations in 2005: an improvement over the former dean's letter? *Acad Med*. 2008;83(3):284–291.
- 4 Messner A, Shimahara E. Letters of recommendation to an otolaryngology/head and neck surgery residency program: their function and the role of gender. *Laryngoscope*. 2008;118(8):1335–1344.
- 5 Cacamese SM, Elnicki M, Speer AJ. Grade inflation and the internal medicine subinternship: a national survey of clerkship directors. *Teach Learn Med*. 2007;19(4):343–346.
- 6 Takayama H, Grinsell R, Brock D, Foy, H, Pellegrini C, Horvath K. Is it appropriate to use core clerkship grades in the selection of residents? *Curr Surg*. 2006;63(6):391–396.
- 7 Cooke M, Irby DM, Sullivan W, Ludmerer KM. American medical education 100 years after the Flexner report. *N Engl J Med*. 2006;355(13):1339–1344.
- 8 Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356(4):387–96.