Medical School Grade Inflation; An Expanding Problem?

James Jackson Gregory[1], SUSANA DIAZ COTO[1], Marcus Philip Coe
[1]Orthopaedic Surgery

INTRODUCTION:

As of January 2022, the United State Medical Licensing Examination (USMLE) Step 1 exam transitioned to Pass-Fail in an effort to encourage a more wholistic review applicants for residency. While well intentioned, the objective data available for programs to review has decreased, now limited to USMLE Step 2 scores, and core clinical clerkship grades. With hundreds of students vying for a single residency position, the selection process is a challenging endeavor, and understanding differences between institutions is paramount. We sought to investigate if the percentage of students awarded the highest grade differed between different grading schemas, if skew has occurred over time, and to what extent institutions follow the Association of American Medical Colleges (AAMC) recommendations for the Medical Student Performance Evaluation (MSPE).
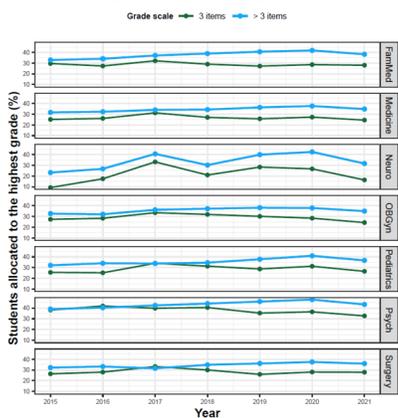
METHODS:

A retrospective review of applications submitted to a single orthopaedic residency program was conducted between 2015-2021. Core Clerkships were defined as Surgery, Family Medicine, Pediatrics, OBGyn, Psychology, Neurology, and Medicine. De-identified MSPE's were used to obtain grading schemas as well as distributions of grades awarded. Grading schemas were categorized as having 3 categories (honors/pass/fail) or greater than 3. Institutions with binary (pass-fail) were excluded. Mixed effects logistic regression models were fitted to determine the odds-ratio of receiving the highest grade, and to assess if skew occurred over the study window.

RESULTS:

Over the 7 year study window, MSPE's from 141 of 155 United States Allopathic Medical Schools were available for review, totaling 994 clerkship records. 6 institutions (4.25%) did not report distributions, 5 institutions (3.5%) used a Pass-Fail core clerkship grading schema, and 1 institution (0.7%) reported quartiles and were excluded from analysis. 91.48% of institutions followed AAMC recommendations in reporting graphical comparison data. In total 760 clerkships/schools/years were included in the analysis. In all years, the proportion of students awarded the highest grade were smaller in schools with 3 categories compared to >3 categories, and odds of being awarded the highest grade was 7% greater in schools with >3 categories. Skew was present, with the odds of receiving the highest grade increasing by 5% per year. Odds of receiving the highest grade were greater for Psychology, Family-medicine, OBGyn, and Pediatrics compared to surgery. Odds of Neurology honors were lower than surgery.

DISCUSSION AND CONCLUSION:

Institutions using 3-category grading schema had statistically lower odds of awarding honors, in contrast to mathematical theory of finer delineation with increased categories. There was statistically significant skew that occurred over the 7 year study window, with an increase in the highest grade of 5% by year. Most institutions follow AAMC MSPE recommendations, although more explicit verbiage pertaining to MSPE comparative data may improve application review. External motivations may contribute to increased awarding of highest grade as medical schools benefit from a high match rate with recruitment of new students. Future studies should seek to investigate the phenomena of skew historically, and the change in grading schemas following the COVID-19 pandemic and change of USMLE Step-1 moving to Pass/Fail.

Proportion of students awarded with highest Grade	Overall	Schools using 3 grade categories	Schools using > 3 grade categories
Overall	35.3	28.6	36.4
Clerkships			
Surgery	33.7	28.4	34.5
Fam. Medicine	36.4	28.8	37.7
Pediatrics	34.8	28.7	35.8
OBGyn	34.8	29.0	35.7
Psych	42.7	37.9	43.4
Neuro	30.9	19.0	32.6
Medicine	33.4	26.6	34.5
Years			
2015	31.0	25.9	32.0
2016	32.2	27.7	33.2
2017	36.0	33.9	36.3
2018	35.6	30.1	36.3
2019	38.0	28.8	39.2
2020	39.5	29.7	40.8
2021	35.0	25.8	36.5

Table 1. Proportion of students allocated to the highest grades expressed in %

	Univariate		Multivariate	
	OR [95% CI]	p-value	OR [95% CI]	p-value
Grade Scale				
3 items	Ref.		Ref.	
> 3 items	1.14 [1.10, 1.18]	< 0.001	1.07 [1.03, 1.11]	<0.001
Clerkships				
Surgery	Ref.		Ref.	
Fam. Medicine	1.14 [1.12, 1.16]	< 0.001	1.15 [1.03, 1.17]	< 0.001
Pediatrics	1.06 [1.04, 1.07]	< 0.001	1.06 [1.04, 1.07]	< 0.001
OBGyn	1.05 [1.03, 1.07]	< 0.001	1.05 [1.03, 1.07]	< 0.001
Psych	1.50 [1.47, 1.53]	< 0.001	1.50 [1.47, 1.53]	< 0.001
Neuro	0.85 [0.83, 0.86]	< 0.001	0.85 [0.83, 0.87]	< 0.001
Medicine	0.99 [0.97, 1.00]	0.149	0.99 [0.97, 1.00]	0.151
Years				
Years	1.05 [1.05, 1.05]	<0.001	1.05 [1.05, 1.05]	< 0.001

Table 2: Unadjusted and adjusted OR from the mixed effect logistic regression models.