The American Journal of Surgery (2015) 209, 760-764

The American Journal of Surgery®

Surgical Education

The sum is greater than its parts: clinical evaluations and grade inflation in the surgery clerkship

Robert E. S. Bowen, M.D., M.P.H., Wendy J. Grant, M.D., Kimberly D. Schenarts, Ph.D.*

Department of Surgery, University of Nebraska Medical Center, 983280 Nebraska Medical Center, Omaha, NE 68198-3280, USA

KEYWORDS:
Grade inflation;
Surgery clerkship;
Clinical evaluations

Abstract

BACKGROUND: This study examines grading component distributions to determine whether alterations in clinical grade determination reduce skew and improve predictive capability of the clinical evaluation.

METHODS: Rotation evaluations, examination scores, and final grades were collected for third-year medical students over a 2-year period. Conditional logistic regression and ordinary least squares regression models were run using SAS 9.3.

RESULTS: Conditional logistic regression demonstrated significant association between global clinical score and final grade and between average clinical evaluation score and final grade. Inclusion of shelf score into either model demonstrated increase in overall final grade.

CONCLUSIONS: Regressions using global and average clinical evaluation score indicate that average score is a better fit for a norm-based grading system. Arguably, the Shelf measures clinical knowledge more objectively than clinical evaluation, but both were significant. Clinical evaluation is prone to inflation because of its subjective nature; conceivably, inflation leads to the decreased correlation with shelf score.

The purpose of this study was to examine surgery clerkship grading and determine the relationship between clinical evaluation and final grade in a clerkship with norm-referenced grading. Grade inflation in medical school clerkship and subinternship evaluations has been a concern for many years. Repeated surveys of internal medicine program directors have demonstrated that a significant proportion (18% in 2004, 38% in 2009) have admitted to passing students who should have failed.[1,2] More than 50% of students in medicine subinternships in 2009 in the United States received the highest grade possible,[2] and more than 60% of students in psychiatry clerkships received the highest grade possible,[3] with one institution boasting a 76% honors rate for medicine subinternships.[2] One of the factors driving grade inflation is the clinical evaluation. Clinical evaluations tend to be an inherently subjective measure of student performance, utilizing observation and interaction with the student to inform evaluation. With work-hour restrictions altering how residents and

The authors declare no conflicts of interest.

* Corresponding author. Tel.: +1-402-559-5905; fax: +1-402-559-3356.

E-mail address: kim.schenarts@unmc.edu

Manuscript received May 29, 2014; revised manuscript September 30, 2014

0002-9610/$ - see front matter © 2015 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amjsurg.2014.10.023

Component	Percentage
NBME Subject Examination Grade	45%
Clinical Evaluation	40%
Oral Examination	15%

Figure 1 Grading rubric used to determine final grade for the years included.

attending physician workflow occurs,[4,5] the time that can be spent in direct supervision of students is often limited, and knowledge deficiencies can be gilded by hardworking or cheerful attitudes.

Some authors refer to grade inflation as grade compression, arguing that the grades themselves have not lost their value and that students are still able to receive low grades, but B's are becoming A's and C's are becoming B's.[6]

Common causes for grade inflation mentioned include unhappy, upset students, little formal education in student evaluation, subjective nature of clinical evaluations, difficulty forming bonds with medical students, and a desire to help students acquire the best residency possible. Faculty often found the evaluation forms to be vague and confusing, with some schools using numerical scales, with descriptions reserved for best and worst scores.[1–3,6–12]

The most frequent cause mentioned, however, is a culture of entitlement prevalent among millennial medical students.[2] Evaluators wish to avoid litigious and angry medical students and the hassle required to either alter the evaluation or assert its legitimacy.[6] The desire to avoid confrontation and direct criticism appears to be particularly strong—one study out of the University of Michigan demonstrated that negative feedback dropped significantly when evaluators met with students' face-to-face.[13]

Patients and Methods

The study population included a total of 250 third-year medical students, 124 from the first year, and 126 from the second year. This included students remediating the clerkship. Rotation evaluations, National Board of Medical Examiners (NBME) Subject Examination (Shelf examination) scores, oral examination scores, and final grades were collected for third-year medical students rotating though the surgery clerkship over a 2-year period, yielding a total of 1,048 observations. The grades are weighted as shown in Fig. 1, with the largest percentages derived from the NBME Subject Examination and the clinical evaluations. Students rotate through four 2-week periods consisting of 2 general surgery rotations (including subspecialties) and 2 surgical specialty rotations (like orthopedics or ophthalmology). These rotations were weighted equally; each 2-week rotation comprised 25% of the clinical evaluation score. Overall, therefore, general surgery rotations contributed 50% of the clinical grade, and surgical specialties contributed 50% of the clinical grade. As seen in Figure 2, students are currently evaluated using a questionnaire that includes 6 questions regarding clinical skills and 4 questions regarding professionalism. An average evaluation score was derived

Evaluation Questions		Needs improvement	Satisfactory	Above Average	Superior
Does the student develop a plan for self-directed learning to include preparation for lectures, clinic, wards, ward rounds and the OR?					
Rate the student concerning:	Teachability and Initiative				
	Reliability and Responsibility				
Does the student Formulate a differential diagnosis by synthesizing information from the history any physical examination and diagnostic material and develop a management plan using the principles of evidence-based medicine?					
Does the student incorporate considerations of cost, efficacy, and ethics involved into recommendations for procedures and treatments for patients?					
Does the student participate in preoperative management of patients including: consideration of patient's developmental stage, preoperative evaluation, nutritional support, wound healing, coagulation disorders, fluid balance, and considering potential postoperative complications?					
Does the student recognize emergent surgical problems and develop a plan for appropriate triage, initial management and referral?					
Does the student engage in professional behavior, including communication skills, honesty and integrity, respect, and maintenance of personal health?					
Rate the student concerning:	Communication Skills				
	Honesty and Integrity				
	Respect				
	Personal Health and Demeanor				
Overall Student Performance					

Figure 2 Evaluation form used by the surgery clerkship for the years included.