The Journal of Emergency Medicine, Vol. 50, No. 2, pp. 302–307, 2016
Copyright © 2016 Elsevier Inc.
Printed in the USA. All rights reserved
0736-4679/$ - see front matter

ELSEVIER

http://dx.doi.org/10.1016/j.jemermed.2015.09.010

Education

TIMING OF EMERGENCY MEDICINE STUDENT EVALUATION DOES NOT AFFECT SCORING

Katherine M. Hiller, MD, MPH, Anna Waterbrook, MD, and Kristina Waters, BS

Department of Emergency Medicine, University of Arizona, Tuscon, Arizona
Reprint Address: Katherine M. Hiller, MD, MPH, Department of Emergency Medicine, University of Arizona, 1501 North Campbell Avenue, Tucson, AZ 85724

☐ **Abstract—Background: Evaluation of medical students rotating through the emergency department (ED) is an important formative and summative assessment method. Intuitively, delaying evaluation should affect the reliability of this assessment method, however, the effect of evaluation timing on scoring is unknown. Objective: A quality-improvement project evaluating the timing of end-of-shift ED evaluations at the University of Arizona was performed to determine whether delay in evaluation affected the score. Methods: End-of-shift ED evaluations completed on behalf of fourth-year medical students from July 2012 to March 2013 were reviewed. Forty-seven students were evaluated 547 times by 46 residents and attendings. Evaluation scores were means of anchored Likert scales (1–5) for the domains of energy/interest, fund of knowledge, judgment/problem-solving ability, clinical skills, personal effectiveness, and systems-based practice. Date of shift, date of evaluation, and score were collected. Linear regression was performed to determine whether timing of the evaluation had an effect on evaluation score. Results: Data were complete for 477 of 547 evaluations (87.2%). Mean evaluation score was 4.1 (range 2.3–5, standard deviation 0.62). Evaluations took a mean of 8.5 days (median 4 days, range 0–59 days, standard deviation 9.77 days) to complete. Delay in evaluation had no significant effect on score ($p = 0.983$). Conclusions: The evaluation score was not affected by timing of the evaluation. Variance in scores was similar for both immediate and delayed evaluations. Considerable amounts of time and energy are expended tracking down delayed evaluations. This activity does not impact a student's final grade.** © **2016 Elsevier Inc.**

☐ **Keywords—medical student; emergency medicine; clerkship; evaluation; assessment; timing; recall bias; education; missing at random**

INTRODUCTION

End-of-shift evaluation of medical students is a commonly used summative and formative feedback mechanism used in the emergency department (ED) (1,2). Workflow around evaluation varies and end-of-shift evaluations can be delayed due to clinical demands or prioritization of other academic duties. Although there are no published studies on the effect of the timing of evaluation on evaluation scores themselves, delaying evaluation of a student introduces the possibility of recall bias. Recall bias is a well-recognized source of error in epidemiologic studies ranging from dietary habits to obstetric histories to medication use (3). In such studies, the time interval from exposure to evaluation has a significant effect on recall bias (4,5). These types of studies, however, all utilize self-evaluation. Little is known about how recall bias affects the evaluation of someone else.

Certainly, a delay in evaluation affects a student's formative evaluation, especially in a rotation with a shorter duration, such as emergency medicine (EM). Additionally, intuitively, it would seem that a delay of end-of-shift evaluation would adversely affect a student's summative evaluation, however, there are no published

RECEIVED: 7 August 2015;
ACCEPTED: 4 September 2015

data to support this assumption. Determining the effect, if any, of delay of evaluation of a student has practical implications to clerkship directors and coordinators as well. Clerkship coordinators and directors spend more time and energy collecting evaluations that are delayed. Whether the effort to collect such delayed evaluations is important to the student's overall grade has yet to be determined.

METHODS

In this quality-improvement project, we conducted a retrospective review of all clinical shift evaluations collected at the University of Arizona, Department of Emergency Medicine from July 2012 to March 2013 on fourth-year medical students enrolled in a selective EM/critical care rotation. End-of-shift evaluations were completed electronically and time stamped on completion through New Innovations (Uniontown, OH). Clinical evaluations were scored based on an average of anchored Likert scale scores for the domains of energy/interest, fund of knowledge, judgment/problem-solving ability, clinical skills, personal effectiveness, and systems-based practice. The overall evaluation score had a range of 1 to 5. Students performed a total of 12 shifts each, with one evaluation per shift recorded. Students were typically paired with a senior EM resident (postgraduate year [PGY] 3, PGY 4, or PGY 5) in EM who was responsible for completing their evaluation. Rarely, a student was paired with an attending or a junior (PGY 1 or PGY 2) resident.

Data collected included the student, the evaluator, the date of the shift performed, the date the evaluation was completed, and the evaluation score. Delay in evaluation in days was calculated based on the date the electronic evaluation was completed compared with the shift for which it was intended.

Because evaluation delay and evaluation score are both continuous variables, linear regression was performed to determine the effect, if any, of delay in evaluation on evaluation score. An additional categorical analysis was performed using immediately completed evaluations (0–1 days), evaluations completed within the week (2–7 days) and evaluation completed in more than 1 week (8 + days). Analysis of variance was performed to determine whether a difference within these categories existed.

In order to detect a difference of 0.83 in evaluation score (a difference that would account for a 5% difference in the end-of-rotation score, or the difference between honors, high pass, and pass), using a two-tailed test with a power of 90% and a significance level of 0.05, a power analysis reveals only 24 individual student evaluations would be needed to observe the difference between immediately completed and delayed evaluations. To look for a 1% difference in end-of-rotation score (a difference that might be significant when it comes to ranking students on a program list), a difference in evaluation score of 0.17 was used, yielding a total of 380 evaluations needed for a two-tailed test with a power of 90% and a significance level of 0.05.

In order to determine whether there was an individual student-level difference between early and late evaluations, descriptive statistics were performed on individual student means. In addition, a chi-square analysis of students who had a mean delay of $\geq$8 days compared to those with a mean delay of <8 days was performed.

Because evaluations completed earlier in the academic year could differ in terms of the evaluation score as well as the delay in evaluation, a chi-square analysis of evaluation score and delay in evaluation was performed on data from the first and fourth quartile dates of data collection.

Data were collected in Microsoft Excel 2007. Data analysis was performed using StataMP 11 (College Station, TX). This quality-improvement study was determined by the University of Arizona Institutional Review Board (IRB) to be exempt from formal IRB review.

RESULTS

A total of 47 fourth-year medical students were evaluated 547 times by 46 EM residents and attendings (97.0% completion rate). Evaluators each completed between 1 and 27 evaluations during this time frame (Table 1). Mean number of evaluations per evaluator was 10.4, however, the data were not parametric. Eighty-three percent of evaluations were completed by 24 evaluators and the median number of evaluations completed was 12.5. Complete data were available for 477 of 547 evaluations (87.2%). Missing data were due to evaluations completed on paper rather than electronically (n = 15), no time stamp/unknown date of completion (n = 52), and missing evaluation (n = 3).

The mean evaluation score was 4.1 (range 2.3–5, standard deviation 0.62). Evaluations took a mean of 8.5 days

Table 1. Mean Evaluation Score and Delay in Days of Medical Student Evaluations by Level of Training of Evaluators

Training Level of Evaluator (No. of Evaluations)	Mean Score (SD)	Mean Delay (SD)	*p* Value
PGY 1 (n = 3 [0.6%])	4.00 (0.79)	30.3 (28.5)	<0.01
PGY 2 (n = 35 [7.3%])	4.26 (0.59)	8.3 (7.7)	0.84
PGY 3 (n = 396 [83.0%])	4.19 (0.6)	8.6 (9.8)	Ref
PGY 4 (n = 26 [5.5%])	3.76 (0.66)	8.0 (9.1)	0.73
PGY 5 (n = 12 [2.5%])	4.34 (0.75)	3.3 (2.5)	0.06
Attending (n = 5 [1.0%])	3.94 (0.64)	4.6 (2.5)	0.36

PGY = postgraduate year; SD = standard deviation.

Download English Version:

https://daneshyari.com/en/article/6084894

Download Persian Version:

https://daneshyari.com/article/6084894

Daneshyari.com