

Medical Student Education

USMLE Step 1 and 2 Scores Correlate With Family Medicine Clinical and Examination Scores

Thomas Myles, MD; Rosa Galvez-Myles, MD

Background and Objectives: We sought to validate the family medicine shelf examination by determining whether correlations exist between the US Medical Licensing Examination (USMLE) family medicine final examination (FMF) scores and the USMLE Step 1 or Step 2 scores. We also evaluated for correlations between the family medicine clinical evaluation scores (CES), final clerkship grades, and all of these examinations. **Methods:** The above scores (first attempts only) of 258 third-year medical students at Texas Tech University at Amarillo from July 1994 to June 2001 were obtained. Linear regression models were made between scores. Low CES and examination or family medicine clerkship failures were statistically compared. **Results:** The average scores were USMLE Step 1: 203.7, USMLE Step 2: 203.8, FMF: 83.3, and CES: 90.1. Positive linear correlations were seen between the Step 1 scores and both the FMF scores and CES. Positive linear correlations between the Step 2 scores and both the FMF scores and the CES were seen. Students failing the USMLE Step 1 were more likely to fail the FMF as well as the family medicine clerkship. Students with a CES less than the 10th percentile were more likely to fail the FMF as well as the family medicine clerkship. **Conclusions:** USMLE Step 1 and Step 2 scores correlate linearly with both the FMF and CES of the third-year family medicine clerkship. Students failing their USMLE Step 1 examination or having a low CES are at risk for failure of the FMF as well as the family medicine clerkship.

(Fam Med 2003;35(7):510-3.)

Core rotations in family medicine have become an integral part of the third year of medical school in the United States. As with other third-year rotations, final grades for family medicine rotations are a composite of an objective examination as well as a clinical evaluation. In most medical schools, clinical evaluations are based on the subjective assessments of residents and faculty.

Predictions as to which students may have difficulty with a clerkship are important, since early assistance to a student at risk for failure has been shown to be beneficial.¹ Campos-Outcalt et al studied 482 students and noted a correlation between the scores on the National Board of Medical Examiners (NBME) Part 1 examination and clinical evaluation scores on family medicine rotations.² These clinical evaluation scores also

correlated with the NBME Part 2 examinations. In addition, scores on NBME Parts 1 and 2 correlated with the authors' departmentally generated family medicine final examination (FMF).²

Investigators in other specialties investigated correlations between scores on NBME subject examinations (given at the end of the rotation) and scores on the NBME Part 2, US Medical Licensing Examination (USMLE) Step 1, and USMLE Step 2 test.³⁻¹⁰ However, a few studies have evaluated the relationships between examination scores and subjective clinical evaluations, and these studies yield varying results. Benenson et al, studying 280 students,¹¹ reported that no relationship existed between the subjective and objective grades on a surgery clerkship, while Jacobsen et al, studying 289 students,¹² found such a relationship did exist. Neither Stenchever et al (111 students)¹³ nor Miller et al (248 students)¹⁴ identified a relationship between the subjective and objective grades in an OB-GYN clerkship, while Metheny (82 students)¹⁵ reported a weak correlation. A report by Marrienfeld and Reid (188 students)¹⁶ noted a strong correlation.

Because of the uncertain relationship between clinical performance and objective examination scores, the validity of the NBME (now the USMLE) family medicine subject examination has been questioned.^{17,18} As a result, we sought to validate the USMLE family medicine subject examination. Specifically, our objective was to determine whether correlations exist between family medicine subject examination scores and scores on either the USMLE Step 1 or Step 2 and if there are correlations between any of these three examinations and clinical evaluation scores on a family medicine clerkship. Finally, we evaluated the students who either failed the USMLE Step 1 or had a clinical evaluation score below the 10th percentile with respect to failure of the family medicine subject examination scores on the family medicine clerkship.

Methods

Family Medicine Rotation

Our family medicine rotation is 6 weeks long. It consists of both outpatient and inpatient responsibilities (including some overnight call). At the outpatient clinic, students are scheduled to see several patients, whom they must present to a faculty member. The final examination (USMLE family medicine subject examination) is given on the final Friday of the rotation.

The clinical evaluation scores consist of the faculty and resident ratings of the student's performance in each of eight categories. At least eight separate evaluations are required for each student. The eight categories graded include fund of knowledge, history and interviewing skills, patient exam skills, oral/written presentations, problem solving/judgment, staff/peer interactions, patient interactions, and professional behavior. Within each category, there are six grade choices: superior, very good, good, satisfactory, marginal, and failing. Faculty and residents are instructed during their orientation as to the manner in which to evaluate students, and additional conferences about evaluation criteria are provided. The eight component scores are combined into a composite score, and only the composite evaluation scores were used in this study.

Students must obtain a combined score on the clinical evaluation and final examination of 75 or above for a student to successfully complete the rotation. Students averaging less than 75 must repeat some or all of the rotation (at the discretion of the clerkship director) including retaking the final examination. Students with an exam score below 70 must repeat the entire rotation and retake the examination.

All students are required to take the USMLE Step 1 at the conclusion of their second year of medical school; students are required to pass this examination to graduate from medical school. All students also must take the USMLE Step 2 exam, but passing this examination is not a requirement for graduation.

Data Collection

After approval of the Institutional Review Board, we evaluated the examination scores of the 258 medical students assigned to the family medicine clerkship at Texas Tech University at Amarillo from July 1994 through June 2001. The USMLE Step 1, Step 2, FMF (USMLE subject examination), and the composite family medicine clinical evaluation scores (CES) were all obtained from the registrar's office. Only the first attempt at each examination was included in our analysis.

Data Analysis

We performed linear regressions between the USMLE Step 1 and 2 scores and both the FMF and CES. The family medicine subject examination scores were also compared with the CES. We determined the 10th percentile scores for the CES. Chi square analysis, odds ratios (OR), Fischer exact tests, and Student's *t* tests were used (where appropriate) to compare USMLE Step 1 examination, family medicine subject examination, and CES. The risks for failure of the family medicine clerkship or family medicine subject examination were determined for the students who either failed the USMLE Step 1 or had a clinical evaluation score less than the 10th percentile. Significance was set at $P < .05$. All statistical analysis was done with SPSS 9.0 (SPSS Inc, Chicago).

Results

Relationships Between Examination Scores

The students' average scores for the three examinations and the clinical evaluations are shown in Table 1. A linear relationship was seen between the family medicine subject examination scores and the USMLE Step 1 scores ($r^2 = .175$, $P < .001$) and Step 2 scores ($r^2 = .276$, $P < .001$). Only a weak linear relationship was seen, however, between CES and USMLE Step 1 ($r^2 = .03$, $P < .006$) and Step 2 scores ($r^2 = .089$, $P < .001$). A comparison between CES and family medicine subject examination scores also revealed a weak linear relationship ($r^2 = .022$, $P < .049$). All correlations were positive.

Students Who Failed Examinations

Eighteen students failed the USMLE Step 1, 11 students failed the USMLE Step 2, and 29 students failed the family medicine subject examination (score < 75). Nine students failed the family medicine clerkship due to an exam score < 70 . One student failed the clerkship due to a combined score of less than 75 and dropped out of school prior to taking the USMLE Step 2. Therefore, 10 students failed the family medicine clerkship. None of these students who failed the clerkship had a clinical evaluation score less than 75.

Students who failed the USMLE Step 1 had lower family medicine subject examination scores than those who passed (77.6 versus 83.8, $P < .001$), and these

Table 1
Descriptive Statistics

	Score	SD	10th Percentile	90th Percentile
USMLE Step 1 (n=258)	203.7	19.1	182.9	228
USMLE Step 2 (n=257)	205.0	20.7	178	234
Family medicine final examination (n=258)	83.3	7.9	73	94
Subjective clinical scores (n=258)	90.1	2.6	87	92

USMLE—US Medical Licensing Examination
SD—standard deviation

students were more likely to fail the family medicine clerkship (OR: 6.8, 1.5–30.1). These students were not, however, at increased risk to have a clinical evaluation score less than the 10th percentile (OR: 2.5, .6–9.3).

Students with CES less than the 10th percentile had lower family medicine subject examination scores (76.3 versus 83.9, $P<.005$). The OR for these students to fail the family medicine clerkship was 7.2 (1.3–41.3), while the OR for them to fail the family medicine subject examination was 5.6 (1.4–22.4).

The OR for failure of the family medicine subject examination for students who either failed the USMLE Step 1 or received a CES in the lowest 10th percentile (n=29) was 3.5 (1.3–9.6). In the case of these students, the OR for failure of the family medicine clerkship was 4.0 (1.0–16.8). The two students who received both a CES less than the 10th percentile and failed the USMLE Step 1 examination failed both the family medicine clerkship (OR: 24.3 [12.3–47.8]) and subject examination (OR: 7.2 [5.1–10.2]).

Discussion

This is the first study (to the best of our knowledge) to compare family medicine performances (subjective and objective) to USMLE Step 1 and 2 scores. Our study indicates that a modest, but statistically significant, positive correlation exists between scores on both USMLE Step 1 and Step 2 and scores on the USMLE family medicine subject examination. These correlations are the most important finding of our study, since they demonstrate the validity of the USMLE family medicine subject examination as a means of assessing a student's knowledge of family medicine topics. Certainly if the USMLE family medicine exam did not at all correlate with any other test of medical knowledge, it would be invalid. Our findings are consistent with studies that

report correlations between other third-year NBME subject examinations and licensure examinations.⁵⁻¹⁰

The second important finding in this study is the potential ability to identify the “at-risk” student. Students at high risk for failing either the family medicine clerkship or the family medicine subject examination have either failed the USMLE Step 1 or received a clinical evaluation score less than the 10th percentile. Students satisfying both of these criteria appear to be at extremely high risk for clerkship failure. The potential application of a “midterm” test to further delineate these at-risk students could also be considered, since it has been shown to also have some predictive potential.¹⁹ The potential benefit of early intervention to assist the at-risk student has been described.¹

A potential weakness of the study is that it includes data from only one medical school. However, this could also serve as an advantage due to increased uniformity in teaching and evaluation techniques. An additional limitation lies with the subjective nature of the clinical evaluations. The lack of objectivity could provide a source of error. Though our evaluation forms cover eight categories, only the composite scores were used.

The weak correlation observed between the subjective and objective scores suggests that these entities measure different factors. However, since students with poor clinical evaluations did have lower final examination scores, an alternative explanation could be that the observed relationships may be weak due to occasional grade inflation. This potential error source suggests a need for either greater objectivity by evaluators or a need to reteach evaluators methods of student evaluation. Studies suggest the existence of grade inflation (particularly for the aggressive learner).^{11,13,16,20} It is hoped that the large number of evaluations required (eight) would average out the grades of each student (with respect to the easy/hard graders).

An additional potential source for error is the effect of rotation sequence. This could affect the correlations negatively, since students clerking with the department later in the year could score higher on the final examination. Such effects have been shown in other disciplines.^{7,21,22} We did not break down this information for this paper due to the lack of sufficient power in several categories. Subjective grades have not been shown to vary by semester of clerkship.^{21,22}

Conclusions

Our study validates the USMLE family medicine subject examination through our observed correlations of that examination to the USMLE Step 1 and 2 examinations. The enhanced ability to predict the at-risk family medicine students by means of USMLE Step 1 failure or low clinical evaluation scores is highly important.

Corresponding Author: Address correspondence to Dr Myles, St Louis University, Department of Obstetrics and Gynecology, Saint Mary's Health Center, 6420 Clayton Road, St Louis, MO 63117. mylesth@slucare1.sluh.edu.

REFERENCES

1. Parenti CM. A process for identifying marginal performers among students in a clerkship. *Acad Med* 1993;68:575-7.
2. Campos-Outcalt D, Witzke DB, Fulginiti JV. Correlations of family medicine clerkship evaluations with scores on standard measures of academic achievement. *Fam Med* 1994;26(2):85-8.
3. Spellacy WN, Dockery JL. A comparison of medical student performance on the obstetrics and gynecology National Board Part II Examination and a comparable examination given during the clerkship. *J Reprod Med* 1980;24:76-8.
4. Metheny WP, Holzman GB. Student performance on the NBME Part II subtest and subject examination in obstetrics-gynecology. *J Med Educ* 1988;63:456-62.
5. Armstrong A, Dahl C, Haffner W. Predictors of performance on the National Board of Medical Examiners obstetrics and gynecology subject examination. *Obstet Gynecol* 1998;91:1021-2.
6. Myles TD. US Medical Licensure Examination Step 1 scores and obstetrics-gynecology clerkship final examination. *Obstet Gynecol* 1999;94:1049-51.
7. Ripkey DR, Case SM, Swanson DB. Predicting performances on the NBME surgery subject test and USMLE Step 2: the effects of surgery clerkship timing and length. *Acad Med* 1997;72:S31-S33.
8. Case SM, Ripkey DR, Swanson DB. The effects of psychiatry clerkship timing and length on measures of performance. *Acad Med* 1997;72:S34-S36.
9. Ripkey DR, Case SM, Swanson DB. Identifying students at risk for poor performance on the USMLE Step 2. *Acad Med* 1999;74:S45-S48.
10. Myles TD, Henderson RC. Medical licensure examination scores: relationship to obstetrics and gynecology examination scores. *Obstet Gynecol* 2002;100:955-8.
11. Benenson TF, Stimmel B, Aufses A Jr. Concordance of surgical clerkship performance and National Board of Medical Examiners Part II subtest scores. A validation model. *Surgery* 1981;89:692-6.
12. Jacobson MJ, Sherman L, Perlman I, Lefferts R, Soroff H. Clerkship site and duration: do they influence student performance? *Surgery* 1986;100:306-11.
13. Stenchever MA, O'Toole B, Irby D. Evaluating student performance in an obstetrics and gynecology clerkship. *Am J Obstet Gynecol* 1979;134:235-7.
14. Miller JM, Smith IK, Sosnowski R, Hester L. Evaluation of student performance in an obstetrics and gynecology clerkship. *J Reprod Med* 1982;27:443-6.
15. Metheny WP. Limitations of physician ratings in the assessment of student clinical performance in an obstetrics and gynecology clerkship. *Obstet Gynecol* 1991;78:136-41.
16. Marienfeld RD, Reid JC. Subjective versus objective evaluation of clinical clerks. *N Engl J Med* 1980;302:1036-7.
17. Schneider FD. What is a national board exam in family medicine? *Fam Med* 1996;28(6):441.
18. Beasley JW. Validity of NBME family medicine exam questioned. *Fam Med* 1997;29(1):7.
19. Hemmer PA, Markert RJ, Wood V. Using in-clerkship tests to identify students with insufficient knowledge and assessing the effect of counseling on final examination performance. *Acad Med* 1999;74:73-5.
20. Myles TD, Galvez-Myles R. Relationship between subjective grade and examination score in obstetrics and gynecology clerkship at community and university sites. *J Reprod Med* 2001;46:729-34.
21. Hampton HL, Collins BJ, Perry KG, Meydrech EF, Wisner WL, Morrison JC. Order of rotation in third-year clerkships—influence on academic performance. *J Reprod Med* 1996;41:337-40.
22. Manetta A, Manetta E, Emma D, Keegan KA, Williams JH. Effects of rotation discipline on medical student grades in obstetrics and gynecology throughout the academic year. *Am J Obstet Gynecol* 1993;169:1215-7.