

# Family Medicine Clerkship Assessment Analysis: Application and Implications

Anna Virani, MD | Sarah E. Stumbar, MD, MPH | Rodolfo Bonnin, PhD

PRIMER. 2021;5:32.

Published: 9/27/2021 | DOI: 10.22454/PRIMER.2021.915080

## Abstract

**Introduction:** Clerkship assessment structures should consist of a systematic process that includes information from exam and assignment data to legitimize student grades and achievement. Analyzing student performance across assessments, rather than on a single assignment, provides a more accurate picture to identify academically at-risk students. This paper presents the development and implications of a structured approach to assessment analysis for the Family Medicine Clerkship at Florida International University Herbert Wertheim College of Medicine.

**Methods:** The assessment analysis included a table presenting the distribution of all assessment performance results for 166 clerkship students from April 2018 to June 2019. A correlation table showed linear relationships between performance on all graded activities. We conducted a Pearson analysis ( $r$ ), coefficient of determination ( $r^2$ ), multiple regression analysis, and reliability of performance analysis.

**Results:** Performance on one assessment—the core skills quiz—yielded a statistically significant correlation ( $r=.409$ ,  $r^2=.16$ ,  $P<.001$ ) with the final clerkship grade. The reliability of performance analyses showed low performers ( $<-1.7$  SD), had both a low mean quiz score (59.6) and final grade (83). Top performers ( $>-1.7$  SD) had both a high mean quiz score (88.5) and final grade (99.6). This was confirmed by multiple regression analysis.

**Conclusion:** The assessment analysis revealed a strong linear relationship between the core skills quiz and final grade; this relationship did not exist for other assignments. In response to the assessment analysis, the clerkship adjusted the grading weight of its assignments to reflect their utility in differentiating academic performance and implemented faculty development regarding grading for multiple assignments.

## Introduction

Identifying low-performing students early in medical school can help facilitate early implementation of additional academic support services. As such, assessments that effectively measure medical knowledge related to a learning experience are critical.<sup>1,2</sup> Low-stake assessments like quizzes suffer from a lack of student motivation and may not be a quality measure.<sup>3</sup> However, as academic performance among medical students is relatively consistent and predictive,<sup>4</sup> low-stake assessments are tools for early identification of academically at-risk students; little scholarship exists on this.<sup>5,6</sup>

Medical schools in the United States are highly selective, thereby creating a homogeneous population of academically successful medical students<sup>7,8</sup> who should perform in a consistent manner from assessment to assessment.<sup>9,10</sup> The content of clerkship assessments measures knowledge related to a set of learning objectives (eg, family medicine clerkship curriculum),<sup>11</sup> which makes quizzes and other assessments integrated as related measures where performance should be consistent.

Multiple assessments within a clerkship measure knowledge about one domain. A strong positive relationship among assessments supports convergent validity as knowledge measures,<sup>11</sup> while statistical correlations between assessment outcomes predict learning.<sup>12</sup> Therefore, assessment quality should be viewed as a holistic system.<sup>13</sup> Analyzing student performance across assessments provides a more accurate picture that identifies at-risk students and predicts later performance. This paper presents the assessment analysis approach used to evaluate the relationships between Florida International University Herbert Wertheim College of Medicine (FIU HWCOCOM)'s Family Medicine Clerkship's graded assignments and final grade.

## Methods

---

We conducted an assessment analysis for 166 students who completed the required 8-week, third-year family medicine clerkship between April 2018 and June 2019. The FIU HWCOCOM Family Medicine Clerkship includes multiple graded assignments that are due throughout the clerkship (Table 1).

The statistical foundation of the assessment analysis model is based on linear associations between the outcomes of the assessments. Students who perform well on one assessment should perform well on others. Even with individual variability, performance should remain relatively stable up and down the performance scale. We conducted a Pearson correlation ( $r$ ) and coefficient of determination ( $r^2$ ), resulting in a table that provides a statistical interpretation of the relationships.

In addition to the correlation table, we created a line chart to confirm linear relationships. This chart serves as a visual feedback mechanism to identify characteristics of the outcomes including spreads in scores, group outliers, and relationships with other assessments. Finally, we performed a multiple regression analysis. We obtained institutional review board exemption for this study.

## Results

---

The analysis revealed that only the core skills quiz had a statistically significant relationship with student performance across all other assessments and with the final grade ( $r=.409$ ,  $r^2=.167$ ,  $P<.0001$ ; Table 2). The results are supported in Figure 1, which shows the performance relationships across the clerkship assessments.

To account for individual factors related to student performance variability, the stratified averages are based on the clerkship grade. Each data point represents the mean score of 16 students arranged from top to bottom; the bottom six groups each include 17 students. The performance outcomes for the core skills quiz and final grade show a consistent or stable pattern of performance with only one set of means with crossover. Additionally, there is a distinct visual separation between the bottom group of students and the next group up, allowing for easy identification of the group of students most at-risk for poor performance. Furthermore, for the eight students who were the lowest performers, scoring below -1.7 SD from the cohort mean, there was an average core skills quiz score of 59.6, compared to the cohort mean of 78.5 (excluding the bottom eight scores). The mean final grade was 83, compared to the cohort mean of 92.6 (excluding the bottom eight scores). Students scoring above -1.7 SD had a mean core skills quiz score of 78.5 and mean final grade of 92.6.

We conducted a multiple regression analysis to examine the relationship between the final clerkship grade and

various potential predictors such as the core skills quiz, Clinical Assessment of Student Performance (CASP), End of Clerkship Assessment (EOCA), and National Board of Medical Examiners (NBME) subject exam (Table 3).

The multiple regression model with all four predictors produced  $R^2=.659$ ,  $F(4, 161)=77.62$ ,  $P<.0001$ . The scores of the core skills quiz, CASP, EOCA, and NBME subject exam had significant positive regression weights, indicating students with higher scores on these scales are expected to have higher final grades, after controlling for other variables in the model (Table 4).

Our results show, for each additional 1-point score in the core skills quiz, the average final grade is expected to increase by 0.044, assuming that NBME subject exam, CASP, and EOCA final scores remain constant. The coefficients from the output of this model can be used to create the following estimated regression equation: Final Clerkship Grade= $19.261+0.044*\text{Core Skills Quiz}+ 0.272*\text{CASP Final Score}+0.312*\text{EOCA Final Score}+0.175*\text{NBME subject exam}$ , and predict the final clerkship grade for a student, based on the scores of the core skills quiz, CASP, EOCA and NBME subject exam.

## Conclusions

---

The assessment analysis revealed a strong linear relationship between the core skills quiz and the final clerkship grade; this did not exist for other assignments. The core skills quiz is an objective measure of medical knowledge, whereas other assignments worth less than 10% of the final grade are subjective, and all students tend to perform well on them.

The regression analysis further confirmed the predictive nature of the core skills quiz to the final grade along with its relationship to other assessments within the clerkship.<sup>11</sup> Even though the core skills quiz is weighted at 4% of the overall grade, our results indicate that it contributes meaningfully to the assessment of students' achievement of the clerkship learning objectives.

Through the assessment analysis, it became clear that, while subjective assignments are important to students' learning, they do not assist in predicting lower-performing students warranting early intervention by faculty. However, the core skills quiz can serve as an evidence-based predictor to later performance. This system could be used to identify academically at-risk students early in the clerkship, thereby allowing faculty to identify poor performance, intervene, and monitor the student's performance. Students can be informed of the core skills quiz's predictive nature and have time to refocus their studying.<sup>14</sup>

This study had limitations. While the core skills quiz assesses medical knowledge, it is not a direct predictor of clinical skills, which are evaluated through the CASP and EOCA. However, the core skills quiz had a strong statistical association with these clinical performance assessments. Unsurprisingly, students with high levels of medical knowledge are also high performers in the clinical setting. Another limitation was this included an examination of the relationships of assessments in one clerkship at one medical school, and these specific findings are not generalizable to clerkships at other medical schools. However, our methodology, as outlined above, could easily be implemented for other clerkships and different institutions.

As a result of the assessment analysis, the clerkship adjusted the grading weight of its assignments to more accurately reflect their utility in differentiating academic performance. Faculty development was implemented regarding grading for the subjective assignments, with additional adjustments made to the assessment rubrics to assist in stratifying performance.

The model used in this study provides a user-friendly approach to identifying and predicting student performance. Observing the statistical associations between assessments can also serve as an additional feedback mechanism to enhance assessment quality. A measurable understanding of how assessments and

the final clerkship grades are statistically associated with one another can be used to further develop curricular plans.

## Tables and Figures

**Table 1: Family Medicine Clerkship Graded Assignments**

Assignments	Percent of Final Grade	Due Date (Week of Clerkship)	Description
Core skills quiz	4	1	Thirteen multiple choice questions that cover topics such as antibiotics, asthma, urology, dermatology, and gynecology. Supplements a hands-on skills simulation activity such as use of metered dose inhaler, urine dipstick analysis, prostate and cervical exam.
Patient note	3	2	A full written history and physical; followed by several reflective questions regarding students' performance in the encounter.
Student teaching	10	2-6	Interactive 15-minute student teaching presentation on an assigned core family medicine topic, such as hypertension, GERD, constipation.
Narrative medicine assignment	4	6	A faculty-led small group and written narrative essay, which ask students to reflect on a significant patient encounter.
Clinical assessment of student performance (CASP)	25	8	Preceptor's clinical evaluation of student
End of clerkship assessment (EOCA)*	17	8	Three-station OSCE, requiring students to evaluate a standardized patient, write a patient note, and present the case to a faculty member.
NBME Subject Exam	30	8	National board of medical examiners family medicine subject exam

Abbreviations: GERD, gastrointestinal reflux disorder; NBME, National Board of Medical Examiners; OSCE, objective structured clinical examination.

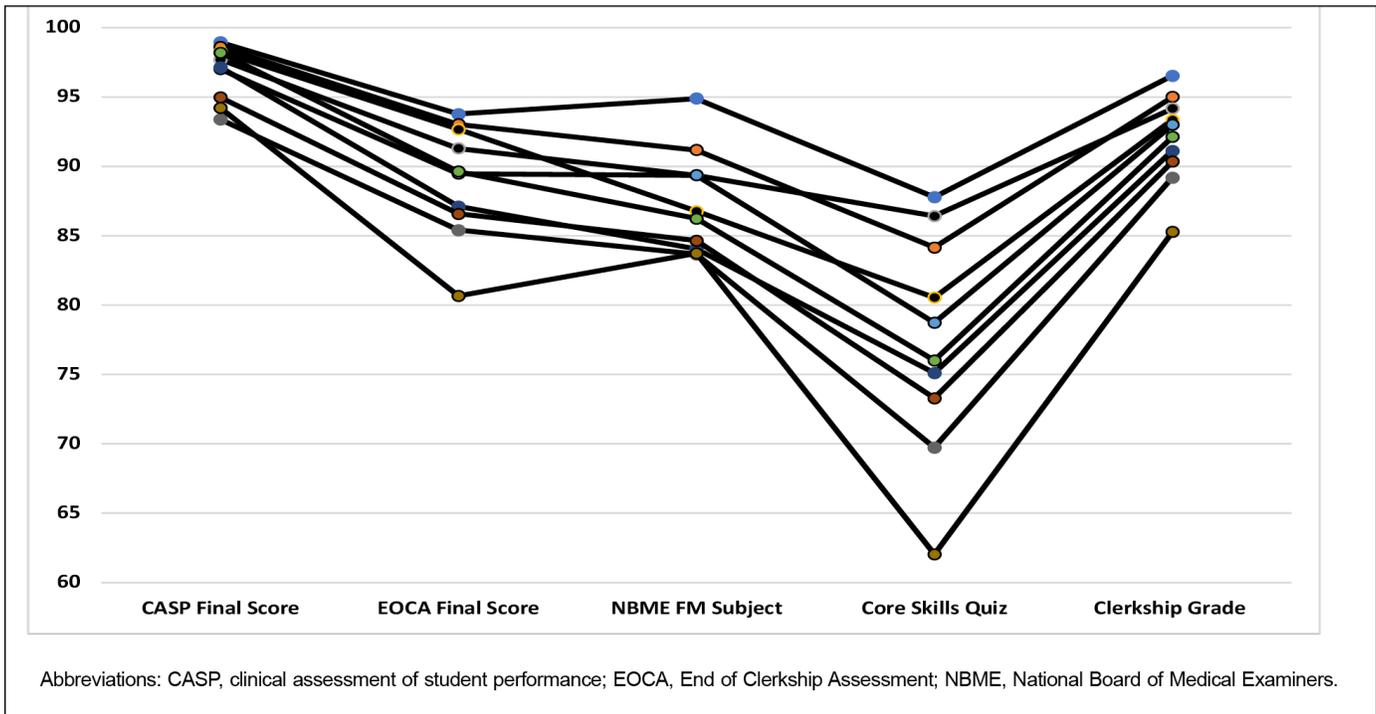
**Table 2: Correlation Table**

Family Medicine Clerkship	Core Skills Quiz (Pearson $r$   $r^2$ )		Clinical Assessment of Student Performance (Pearson $r$   $r^2$ )		End-of-Clerkship Assessment (Pearson $r$   $r^2$ )		NBME Family Medicine Subject (Pearson $r$   $r^2$ )		Clerkship Score (Pearson $r$   $r^2$ )
Core skills quiz	1								
Clinical assessment of student performance	0.16994	0.0289	1						
	$P=.0286$								
End-of-clerkship assessment*	0.23052	0.0531	0.09625	0.0093	1				
	$P=.0026$		$P=.2174$						
NBME family medicine subject	0.18003	0.0324	0.13633	0.0186	0.25809	0.0666	1		
	$P=.0195$		$P=.0799$		$P=.0007$				
Clerkship score	0.40902	0.1673	0.40597	0.1648	0.64863	0.4207	0.48937	0.2395	1
	$P<.0001$		$P<.0001$		$P<.0001$		$P<.0001$		

Abbreviation: NBME, National Board of Medical Examiners.

\*End of clerkship assessment is an objective structured clinical examination requiring students to evaluate a standardized patient, write a patient note, and present the case to a faculty member.

**Figure 1: Performance Stability Across Assessments Given During Family Medicine Clerkship**



**Table 3: The Descriptive Statistics**

Descriptive Statistics					
Variable	N	Mean	SD	Minimum	Maximum
Core skills quiz	168	77.56	14.30	31	100
CASP Final Score	166	96.91	3.45	85	100
EOCA Final Score	168	89.10	5.30	73	100
NBME FM Subject	168	87.39	5.27	75	100
Final Clerkship Grade	168	92.10	3.27	75	99

Abbreviations: CASP, clinical assessment of student performance; EOCA, End of Clerkship Assessment; NBME, National Board of Medical Examiners.

**Table 4: The Multiple Regression Model**

	Beta	P Value	CI
Intercept	19.261	.000	9.386 - 29.136
Core Skills Quiz	0.044	<.0001	0.022 - 0.066
CASP Final Score	0.272	<.0001	0.184 - 0.360
EOCA Final Score	0.312	<.0001	0.252 - 0.371
NBME FM Subject	0.175	<.0001	0.115 - 0.235
Observations	166		
R <sup>2</sup> /adjusted R <sup>2</sup>	0.659/0.650		
F-test	77.62		

Abbreviations: CI, confidence interval; CASP, clinical assessment of student performance; EOCA, End of Clerkship Assessment; NBME, National Board of Medical Examiners.

## Corresponding Author

Anna Virani, MD

Herbert Wertheim College of Medicine, Florida International University, 11200 SW 8th Street, AHC2-585, Miami, FL 33199.

[avirani@fiu.edu](mailto:avirani@fiu.edu)

## Author Affiliations

Anna Virani, MD - Herbert Wertheim College of Medicine, Florida International University, Miami, FL

Sarah E. Stumbar, MD, MPH - Herbert Wertheim College of Medicine, Florida International University, Miami, FL

Rodolfo Bonnin, PhD - Herbert Wertheim College of Medicine, Florida International University, Miami, FL

## References

1. Downing SM. The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Adv Health Sci Educ Theory Pract.* 2005;10(2):133-143. doi:10.1007/s10459-004-4019-5
2. Vanderbilt AA, Feldman M, Wood IK. Assessment in undergraduate medical education: a review of course exams. *Med Educ Online.* 2013;18(1):1-5. doi:10.3402/meo.v18i0.20438
3. Cole JS, Bergin DA, Whittaker TA. Predicting student achievement for low stakes tests with effort and task value. *Contemp Educ Psychol.* 2008;33(4):609-624. doi:10.1016/j.cedpsych.2007.10.002
4. Griffin B, Bayl-Smith P, Hu W. Predicting patterns of change and stability in student performance across a medical degree. *Med Educ.* 2018;52(4):438-446. doi:10.1111/medu.13508
5. Milton O, Pollio HR, Eison JA. *Making Sense of College Grades.* San Francisco: Jossey-Bass Publishers; 1986.
6. Franke M. Final exam weighting as part of course design. *Teach Learn Inq.* 2018;6(1):91-103. doi:10.20343/teachlearningqu.6.1.9
7. Baron T, Grossman RI, Abramson SB, et al. Signatures of medical student applicants and academic success. *PLoS One.* 2020;15(1):e0227108. doi:10.1371/journal.pone.0227108
8. Scott JN, Markert RJ, Dunn MM. Critical thinking: change during medical school and relationship to performance in clinical clerkships. *Med Educ.* 1998;32(1):14-18. doi:10.1046/j.1365-2923.1998.00701.x
9. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach.* 2011;33(6):478-485. doi:10.3109/0142159X.2011.565828
10. van der Vleuten CP, Schuwirth LW, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach.* 2012;34(3):205-214. doi:10.3109/0142159X.2012.652239
11. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ.* 2003;37(9):830-837. doi:10.1046/j.1365-2923.2003.01594.x
12. Goldsmith TE, Jorgenson PJ, Acton WH. Assessing structural knowledge. *J Educ Psychol.* 1991;83(1):88-96. doi:10.1037/0022-0663.83.1.88
13. Crowe A, Dirks C, Wenderoth MP. Biology in bloom: implementing Bloom's Taxonomy to enhance student learning in biology. *CBE Life Sci Educ.* 2008;7(4):368-381. doi:10.1187/cbe.08-05-0024
14. Alharbi Z, Cornford J, Dolder L, De La Iglesia B. Using data mining techniques to predict students at risk of poor performance. *SAI Computing Conference (SAI).* 2016: 523-531, doi:10.1109/SAI.2016.7556030

Copyright © 2021 by the Society of Teachers of Family Medicine