5-2007

Clinical Experience and Examination Performance: Is There a Correlation?

Gary L. Beck
University of Nebraska Medical Center

Mihaela Teodora Matache
University of Nebraska at Omaha, dvelcsov@unomaha.edu

Carrie Riha
University of Nebraska Medical Center

Katherine Kerber
University of Nebraska Medical Center

Frederick A. McCurdy
Texas Tech University

CLINICAL EXPERIENCE AND EXAMINATION PERFORMANCE: IS THERE A CORRELATION?

Gary L. Beck, B.S.[1]; Mihaela T. Matache, Ph.D.[2]; Carrie Riha, B.A. [1]; Katherine Kerber, B.S. [1]; Fredrick A. McCurdy, M.D., Ph.D., M.B.A.[3]

[1]Department of Pediatrics, University of Nebraska Medical Center, Omaha, NE

[2]Department of Mathematics, University of Nebraska at Omaha, Omaha, NE

[3]Department of Pediatrics, Texas Tech University Health Sciences Center at Amarillo, Amarillo, TX

Key Words: medical student, examination performance, patient exposure, logbooks

Word Counts:

Abstract: 224

Manuscript: 2,376

Corresponding Author: Gary L. Beck
Department of Pediatrics
University of Nebraska Medical Center
982184 Nebraska Medical Center
Omaha, NE 68198-2184
Email: gbeck@unmc.edu
Office (402) 559-7351
Fax (402) 559-5137

Author Contributions:

Gary L. Beck, B.S. is a co-principal investigator for this project, collecting logbooks, completing statistical analyses, and writing the manuscript.

Mihaela T. Matache, Ph.D. provided oversight of the statistical testing being conducted, making recommendations of how best to analyze the data. She co-wrote the Methods and Results sections as well as edited the entire manuscript.

Carrie Riha, B.A. organized the logbook data as well as the examination data, preparing the information for data analysis. She offered suggestions for approaches to the data analysis. Carrie also reviewed and edited the manuscript.

Katherine Kerber, B.S. was responsible for coding all of the patient logbooks and entering the information into a database. She provided invaluable editing for the manuscript.

Fredrick A. McCurdy, M.D., Ph.D., M.B.A. is a co-principal investigator for this project, collecting logbooks, completing statistical analyses, and writing the manuscript.

Acknowledgements: None

Sources of Funding: None

Competing Interests: None

Ethical Review: The University of Nebraska Institutional Review Board approved this study as exempt under 45 CFR 46:101b, category 4. The IRB number is 140-04-EX.

OVERVIEW BOX

What is already known on this subject:

- Logbook data is used in clinical medical education.
- Little has been reported on the correlation between patient encounters and knowledge-based examination performance.

What this study adds: This study correlates performance on a pediatric clerkship multiple choice examination and patient encounter numbers related to exam topics. Our findings demonstrate increasing patient encounters does not improve exam performance.

Suggestions for further research:

- Study whether student's roles in patient encounters improves the student's knowledge acquisition.
- Develop evaluations for experiential knowledge acquisition during clinical courses to better assess medical student performance.

ABSTRACT

Background: The Liaison Committee on Medical Education (LCME) requires "There must be comparable educational experiences and equivalent methods of evaluation across all alternative instructional sites within a given discipline." The LCME had made an accreditation requirement that students encounter similar numbers of patients with similar diagnoses. However, previous empiric studies have not shown a correlation between numbers of patients seen by students and performance on a multiple-choice examination. **Purpose**: Does students' exposure to patients with specific diagnoses predict performance on multiple-choice examination questions pertaining to those diagnoses? **Methods**: UNMC Pediatrics has collected patient logbooks from clerks since 1994 that contain patient demographic information and the students' role in patient care. During the seventh week of an 8-week course, students took an examination intended to help them prepare for their final examination. Logbooks and pre-examination questions were coded using standard ICD-9 codes. Data were analyzed using Minitab statistical software to determine dependence between patient encounters and test scores. **Participants**: Convenience sample of students completing the clerkship from 1997 through 2000. **Results**: From our analysis, performance on a multiple-choice examination is independent of numbers of patients seen. **Conclusions**: Our data suggest knowledge-based examination performance cannot be predicted by the volume of patients seen. Therefore, too much emphasis on examination performance in clinical courses should be carefully

weighed against clinical performance to determine successful completion of clerkships.

INTRODUCTION

Third-year medical student clerkships in the United States are expected to meet two essential goals: provide an adequate quantity and quality of clinical exposure to students and increase students' knowledge of the broader aspects of medicine. To satisfy these requirements, more medical schools are sending increasing numbers of students to community sites to complete the clinical components of their training due to reduced numbers of hospitalized patients as well as to emphasize managed care models.

Based on requirements by the Liaison Committee on Medical Education (LCME), the accrediting authority for medical education in the United States and Canada, clerkships with more than one site must provide equivalent experiences. Even though it is difficult to assess equivalency, having students maintain logbooks has been shown to be one way that is reasonably accurate and consistent (1-3). In fact, other studies have shown students tend to under-report patient encounters (4). In a previous study we were unable to show there was a relationship between student exposure to patients and overall multiple-choice examination performance (5), which is considered the objective benchmark for successfully completing a clerkship.

Students who completed their third-year pediatric clerkship at the university and in community-based practices do report significant differences in their overall experiences (5-7). They also report that community-based sites provide a richer experience and the students logged a greater volume of patients. However, after completing a standardized multiple-choice examination and a

structured oral examination, no discernable differences between students could be determined based on training location (5).

The purpose of this study was to investigate in more detail if a correlation existed between reported patient encounters and performance on a multiple-choice examination. Since all study participants had completed essentially identical medical education and training within the same environment and physical resources until their third year of training, their education may be considered equivalent. Clerkship settings were apportioned to two tracks: the more traditional university-based experience and the private practice community experience. All of the students had the opportunity to take the multiple-choice examination review during the seventh week of the clerkship. This arrangement provided the opportunity to study the correlation between demonstration of knowledge and patient exposure.

METHODS

Design

All third-year students completed the same course orientation with explicitly stated expectations (e.g., curriculum content, supplemental study materials, online resources, grading policy, and required documentation). Instrumental in this process, supervisory staff at every practice site received a formal orientation to these expectations along with annual updates to any changes in the curriculum. A clerkship coordinator oversaw all administrative tasks, attended all meetings pertaining to curriculum design decisions, and facilitated consistency of data collection across all clerkship training sites.

Students at all sites had the opportunity to take the exam review. The exam review was administered as an actual examination with a time limit of 90 minutes. Once completed, the students returned the scoring sheets and had the opportunity to review the examination with the clerkship director. All examinations were retained at the end of the session to maintain test security.

Sites

Patients were seen in either the university hospital outpatient clinic/inpatient ward setting or in 1 of 9 community practice (CP) sites located in cities from 50 to 475 miles from the medical school campus. In scheduling the clerkship rotations, students had an opportunity to self-select a CP site or the university site. The clerkship coordinator completed the schedule based on students' requests, site availability, and previous academic performance. As long as a student had not repeated a course during the first two years, requests for a community site were granted. Students who chose the community sites for

their clerkship experience were provided with living provisions so they encountered little additional financial hardship relative to students remaining at the university.

Sample

Study participants included third-year students completing their 8-week pediatric clerkship over three years from 1997 to 2000. Each academic year consists of six clerkship groups with approximately 20 students in each rotation. A total of 243 students completed the course over the three year period - 174 at the university and 69 in CP sites. Of these, 154 logbooks were returned, coded and entered into a secure database - 117 from university and 37 from CP rotations.

Students maintained logbooks of their patient encounters. These were returned to the clerkship coordinator on the last day of the course. Patient logs included observed patient's age, primary diagnosis, and the student's role in the encounter. Logbook entries total 20,464 for this time period; university students reported seeing 9,962 patients (an average of 85 patients per student over 8 weeks) and CP students reported 10,502 (an average of 210 per student over 8 weeks).

A co-author rendered each encounter into specific codes using Code-it-Fast software (Ingenix, Salt Lake City, Utah). This software allows the user to enter exact words or phrases to obtain the International Classification of Diseases ICD-9 code, standardized alpha-numeric code numbers for specific diagnoses used for patient billing. Initially, this coder's work was thoroughly

reviewed by one of the authors (FAM) to ensure the accuracy and reliability of the coding process. This software was also used to code test items that pertained to a particular diagnosis for comparison. Students at the university logged 1,090 different ICD-9 codes and the students in the CP sites logged 953 different ICD-9 codes.

Evaluation Tools

During the three years of this study, students took an exam review, a multiple-choice examination (MCE), in the seventh week of the clerkship. Students were given 90 minutes to complete the examination. The MCEs were graded and entered into a database. Each test item pertained to knowledge of a diagnosis that the faculty believed was important. The curriculum objectives had been constructed to emphasize knowledge of each of these diagnostic entities. This allowed one of the co-authors (FAM) to assign a single ICD-9 code to each test item to correlate to the logbooks.

For their final examination, students took the National Board of Medical Examiners (NBME) Subject Examination, a nationally standardized examination consisting of 100 objective multiple-choice questions. Students were allowed 2 hours to complete this examination, which covered a broad range of topics encompassing pediatric medicine. Each of these test questions was not available for coding with the ICD-9 code. Since all of this information is collected as part of the clerkship, we received exempt approval from the UNMC Institutional Review Board to collect and analyze this data.

Validity/Reliability

The MCE has been administered to the students as a means of reviewing for the NBME final examination. Based on a Kuder Richardson Formula 20 test for reliability, this test does not meet minimum standards for reliability (KR-20=0.62). An exam is considered reliable when KR-20≥0.70. Expert validity was obtained by having the clerkship directors of the Council on Medical Student Education in Pediatrics develop and review the examination. All the directors agreed the examination was fair and valid based on the standardized curriculum for pediatric clerkships.

Analyses

The statistical analyses of the data consisted of contingency tables, which test dependence of categorized data, to determine if the examination scores were dependent on the volume of patient encounters. The analyses included a separation of students by type of examination (MCE and NBME), location (university and community), and experience (students at the beginning of the year versus students at the end of the year). Contingency table analyses were further verified using a one-way analysis of variance (ANOVA). Pearson correlation analyses were performed on scores for MCE or NBME scores versus number of patients seen. The MCE questions with specific ICD-9 codes versus number of patients seen with similar diagnoses were similarly analyzed.

RESULTS

This study includes patient logbook data, pre-examination results, NBME examination results, and overall grades from 154 students over the course of academic years 1997 through 2000.

Various statistical analyses were performed on the available sample. Students were arbitrarily grouped based on the numbers of patient encounters logged (<50, 51-100, 101-150, >150). Along with the grouping by patient encounters, we also grouped students by examination scores into five groups (90%, 80%, 70%, 60%, <60%). We initially reviewed descriptive statistics to obtain a general overview of the data.

Contingency tables were used to summarize categorized data, such as numbers of patient encounters versus examination performance. Chi-square testing with a 0.05 level of significance was conducted on both the MCE and NBME examinations to determine if variables tested were independent of one another. We found that patient exposures and examination scores on both MCE (Chi-square for UNMC students = 14.672 and CP students = 6.255 were less than the test statistic of 21.026) and NBME (Chi-square for UNMC students = 9.595 and CP students = 11.303 were less than the test statistic of 21.026) were independent, indicating examination performance was not dependent on patient exposures. An ANOVA with a 0.05 level of significance further confirmed our findings that there was no statistical difference between mean MCE and NBME score and patient exposure (Table 1).

With the structure of the third year, students completing their first clerkship in pediatrics had little to no clinical experience in pediatrics. Because of this, we applied the same testing using contingency tables and ANOVA for students completing the clerkship at the beginning of the academic year and students finishing the clerkship at the end of the academic year. The results of the testing for both MCE and NBME for the different rotations indicated that test performance is independent of patient encounters.

Since students in CP sites tend to see a greater volume of patients, we applied similar tests as above for UNMC versus CP tracks to determine if the track had an impact on the relationship between patient encounters and grades. Based on the test results, there was no dependent relationship between the number of patients seen and test scores.

Finally, Pearson correlation analyses were performed to initially determine if there was any correlation between patients seen and overall examination scores. We assumed the data were regarded as a random sample from a bivariate normal population. The sample correlation coefficient for the MCE was computed at r=0.192 and for the NBME r=0.189. This is indicative of a weak association between patient exposure and examination results. Analyses looking at test items coded V20.2 (healthcare maintenance), the most frequent diagnosis seen by all students, and patient encounters showed a correlation coefficient of r=0.094, which indicates an extremely weak linear relationship between specific diagnostic exposure and examination performance. Additional MCE items are summarized in Table 2.

DISCUSSION

The revision to the Pediatric curriculum at the University of Nebraska Medical Center was met with a great deal of resistance when it was unveiled in 1994. A shift in focus to more ambulatory training concerned the faculty because it was felt the students would not have enough patient exposure. To ensure adequate numbers and types of patients were being seen, students were required to maintain a logbook of their patient encounters. For the purposes of accreditation, the educational experiences and evaluation methods for this decentralized clerkship were carefully structured.

When students began completing the Pediatrics clerkship in clinics throughout Nebraska, the difference in clinical experiences was quickly noted by the volume of patients students were logging. On average, students who participated in the community training track logged an average of 163 patients whereas the students at the university logged an average of 91 patients per clerkship. Given the significant differences in numbers and types of patients seen, we expected students who saw more patients to excel on the NBME Subject Examination.

On the contrary, students who completed the clerkship in the community training track had a mean score of 73.45 ($\pm$5.895) performance on the NBME Pediatric Subject Examination, and university training track students had a mean of 74.93 ($\pm$7.211). Scores on the 60-point MCE averaged 41.68 ($\pm$6.20) for community training track and 39.48 ($\pm$5.08) for university students. Performance on these exams may be attributable to the sound knowledge base of the students

as evidenced by their average score on USMLE Step I and II (Class of 1999: Step I average was 209±17, Step II 221±18; Class of 2000: Step I 213±18, Step II 218±21; and Class of 2001: Step I 216±17, Step II 225±20).

From the statistical analyses of patient encounters and examination performance, the results implied examination performance on both MCE and NBME was not dependent on the number of patient encounters logged. These results indicate performance on a knowledge-based examination was independent of clinical experience. When patient numbers increased, no concomitant increase in examination scores was noted. When more detailed analysis was completed on the more frequently recorded ICD-9 Code (healthcare maintenance) and MCE performance on questions pertaining to this code, there was no demonstrated improvement on examination performance with increased patient encounters.

The limitations of this study need to be addressed. First, use of historical controls may be questioned. Students' performance on MCATs as well as their performance during the first two years of medical school may be a confounding variable that was not taken into consideration. Knowing the MCE was not considered part of the grade most likely impacted performance on that examination, which was probably taken less seriously than the actual NBME, which was 30 percent of the grade.

Another limitation is that this analysis did not take into account the role the students played in the patient encounter (e.g., active versus passive role). The focus was solely on encounters recorded. The amount of time spent with each

patient may also be an influence on student learning which may correlate to performance on a standardized examination. Again, this information was not collected.

Finally, this study involved the pediatrics clerkship at one institution. Therefore, results may not be generalizeable. Clinical experiences and curriculum content vary widely from institution to institution, making a multi-institutional study difficult. The goal of this study was to demonstrate that regardless of patient encounters in various settings, students can still achieve passing scores on knowledge-based examinations.

In light of accreditation standards requiring quantified criteria for the types of patients being seen during a clerkship (LCME ED-2 requirement), great care and analysis of students' experiences need to be taken. Clearly from these results, regardless of the numbers and types of patients seen, students performed similarly on knowledge-based examinations. Previous studies (8-10) demonstrate experiential knowledge and didactic knowledge are independent, but both of incredible importance. For the purposes of grading and evaluation of clinical courses, evaluations are integral in grading but greater emphasis continues to be placed on objective examination performance. Future investigation will include developing more reliable mechanisms for assessing experiential knowledge acquisition during clinical courses which, in conjunction with didactic knowledge, should provide a better assessment of medical student performance on a clerkship.

REFERENCES

1. Butterfield PS, Libertin AG. Learning outcomes of an ambulatory care rotation in Internal medicine for junior medical students. *J Gen Intern Med* 1993; **8**:189-92.
2. Raghoebar-Krieger HMF, Sleijfer D, Bender W, Stewart RE, Popping R. The reliability of logbook data of medical students: an estimation of interobserver agreement, sensitivity and specificity. *Med Educ* 2001; **35**:624-31.
3. Châtenay M, Maguire T, Skakun E, Chang G, Cook D, Warnock GL. Does volume of clinical experience affect performance of clinical clerks on surgery exit examinations? *Am J Surg* 1996; **172**:366-72.
4. Patricoski CT, Shannon K, Doyle GA. The accuracy of patient encounter logbooks used by Family Medicine clerkship students. *Fam Med* 1993; **30**:487-9.
5. McCurdy FA, Beck GL, Kollath JP, Harper JL. Pediatric clerkship experience and performance in the Nebraska Education Consortium: A community vs. university comparison. *Arch Pediatr Adolesc Med* 1999; **153**:989-94.
6. McCurdy FA, Sell DM, Beck GL, Kerber K, Larzelere RE, Evans JH. A comparison of clinical pediatric patient encounters in university medical center and community private practice settings. *Ambul Ped* 2003; **3**:12-5.

7. Irigoyen MM, Kurth RJ, Schmidt HJ. Learning primary care in medical school: Does specialty or geographic location of the teaching site make a difference? *Am J Med* 1999; **106**:561-4.

8. Coughlin LD, Patel VL. Processing of critical information by physicians and medical students. *J Med Educ* 1987; **62**:818-28.

9. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: Theory and implication. *Acad Med* 1990; **65**:611-21.

10. Schmidt HG, Boshuizen HP. On the origin of intermediate effects in clinical case recall. *Mem Cognit* 1993; **21**:338-51.

Table 1. Mean Examination Scores Versus Patient Encounters

# Patients Seen	MCE		NBME	
	UNMC	CP	UNMC	CP
<50 patients	63.95 (Std. Dev. 8.39)	70.27 (Std. Dev. 11.08)	73.04 (Std. Dev. 6.33)	73.33 (Std. Dev. 3.20)
51-100 patients	65.08 (Std. Dev. 7.71)	69.24 (Std. Dev. 13.24)	74.88 (Std. Dev. 6.60)	73.00 (Std. Dev. 7.62)
101-150 patients	67.19 (Std. Dev. 8.73)	71.25 (Std. Dev. 10.40)	76.42 (Std. Dev. 8.02)	73.25 (Std. Dev. 5.37)
>150 patients	71.15 (Std. Dev. 10.12)	68.85 (Std. Dev. 8.68)	77.69 (Std. Dev. 9.38)	73.88 (Std. Dev. 5.86)

ANOVA resulted in p=0.965 for MCE and p=0.531 for NBME, demonstrating no statistical significance between examination scores, further validating Chi-square tests of independence between patient exposure and examination performance.

Table 2. Comparison of ICD-9 Coded MCE Test Items and Average Number of Patients Seen

Diagnosis (ICD-9 CM Code)	Mean Correct Score on Pre-Examination	Average Number Patients Seen	r*
Disorders of Fluid/ Electrolyte (276)	2.07 of 4	0.12 (range 0-1)	0.063
Specific Delays in Development (315)	1.58 of 2	0.01 (range 0-2)	0.084
Seizures (780.3)	1.66 of 2	0.51 (range 0-9)	0.089
Poisoning by Chemical NEC (977.9)	2.76 of 4	0.01 (range 0-1)	0.032
Healthcare Maintenance (V20.2)	7.53 of 10	16.86 (range 0-119)	0.032

*r=correlation between patient exposure and correctly answered questions