Clinical experience and examination performance: is there a correlation?

Gary L Beck,[1] Mihaela T Matache,[2] Carrie Riha,[1] Katherine Kerber[1] & Fredrick A McCurdy[3]

CONTEXT The Liaison Committee on Medical Education (LCME) requires there to be: '…comparable educational experiences and equivalent methods of evaluation across all alternative instructional sites within a given discipline'. It is an LCME accreditation requirement that students encounter similar numbers of patients with similar diagnoses. However, previous empirical studies have not shown a correlation between the numbers of patients seen by students and performance on multiple-choice examinations.

OBJECTIVE This study examined whether student exposure to patients with specific diagnoses predicts performance on multiple-choice examination questions pertaining to those diagnoses.

METHODS The Department of Pediatrics at the University of Nebraska Medical Center has collected patient logbooks from clerks since 1994. These contain information on patient demographics and students' roles in patient care. During week 7 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 analysed using Minitab statistical software to determine dependence between patient encounters and test scores. Subjects comprised a convenience sample of students who completed the clerkship during 1997–2000.

RESULTS Our analysis indicates that performance on a multiple-choice examination is independent of the number 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 the successful completion of clerkships.

KEYWORDS clinical clerkship/*standards; educational measurement/*methods; professional–patient relations; paediatrics/*education; ambulatory care; community medicine/education; teaching materials; Nebraska.

*Medical Education 2007: **41**: 550–555*
doi:10.1111/j.1365-2923.2007.02764.x

INTRODUCTION

Year 3 medical student clerkships in the USA are expected to meet 2 essential goals: to provide students with an adequate quantity and quality of clinical exposure, and to increase students' knowledge of the broader aspects of medicine. In order to satisfy these requirements, more medical schools are sending increasing numbers of students to community sites to complete the clinical components of their training, both as a result of the reduced numbers of hospitalised patients and in order to emphasise managed care models.

According to the requirements of the Liaison Committee on Medical Education (LCME), the accrediting authority for medical education in the USA and Canada, clerkships with more than 1 site must provide equivalent experiences. Although it is difficult

[1]Department of Pediatrics, University of Nebraska Medical Center, Omaha, Nebraska, USA
[2]Department of Mathematics, University of Nebraska at Omaha, Omaha, Nebraska, USA
[3]Department of Pediatrics, Texas Tech University Health Sciences Center at Amarillo, Amarillo, Texas, USA

Correspondence: Gary L Beck MA, Department of Pediatrics, University of Nebraska Medical Center, 982184 Nebraska Medical Center, Omaha, Nebraska 68198-2184, USA. Tel: 1 402 559 7351; Fax: 00 1 402 559 5137; E-mail: gbeck@unmc.edu

Overview

What is already known on this subject

Logbook data are 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 paediatric clerkship multiple-choice examination and patient encounter numbers related to examination topics. Our findings demonstrate that increasing patient encounters does not improve examination performance.

Suggestions for further research

Future research might study whether student roles in patient encounters improve student knowledge acquisition, and develop methods to evaluate experiential knowledge acquisition during clinical courses in order to better assess medical student performance.

to assess equivalency, having students maintain logbooks has been shown to be 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 a relationship between student exposure to patients and overall multiple-choice examination (MCE) performance,[5] which is considered the objective benchmark for successfully completing a clerkship.

Students who completed their Year 3 paediatric clerkship at the university and those who completed it in community-based practices reported significant differences in their overall experiences.[5–7] They also reported that community-based sites provided a richer experience and that students logged a greater volume of patients. However, after completing a standardised MCE and a structured oral examination, no discernible differences between students could be determined according to 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 an MCE. As all study participants had completed essentially identical medical education and training within the same environment and physical resources until Year 3, their prior education may be considered equivalent. Clerkship settings were apportioned to 2 tracks: the more traditional university-based experience and the private practice community experience. All the students had the opportunity to take the MCE review during week 7 of the clerkship. This arrangement provided the opportunity to study any correlation between the demonstration of knowledge and patient exposure.

METHODS

Design

All Year 3 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 co-ordinator oversaw all administrative tasks, attended all meetings pertaining to curriculum design decisions, and facilitated the consistency of data collection across all clerkship training sites.

Students at all sites had the opportunity to take the examination review. This was administered as an actual examination with a time limit of 90 minutes. Once completed, students returned the scoring sheets and were given an 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 50–475 miles from the medical school campus. In scheduling the clerkship rotations, students were able to select a CP site or the university site. The clerkship co-ordinator completed the schedule according to students' requests, site availability and previous academic performance. As long as a student had not repeated a course during the first 2 years, requests for a community site were granted. Students who chose the community sites for their clerkship experience

were provided with living provisions so that they would encounter little additional financial hardship relative to students who remained at the university.

Sample

Study participants included Year 3 students who completed their 8-week paediatric clerkship during the 3 years from 1997 to 2000. Each academic year consists of 6 clerkship groups with approximately 20 students in each rotation. A total of 243 students completed the course over the 3-year period – 174 at the university and 69 at CP sites. From these, 154 logbooks were returned, coded and entered into a secure database – 117 from university-based and 37 from CP rotations.

Students maintained logbooks of their patient encounters. These were returned to the clerkship co-ordinator on the last day of the course. Patient logs included data on the observed patient's age, primary diagnosis, and the student's role in the encounter. Logbook entries totalled 20 464 for this time period; university students reported seeing 9962 patients (an average of 85 patients per student over 8 weeks) and CP students reported 10 502 patients (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, UT, USA). This software allows the user to enter exact words or phrases to obtain International Classification of Diseases (ICD)-9 codes, standardised alpha-numeric code numbers for specific diagnoses used for patient billing. Initially, this coder's work was thoroughly reviewed by 1 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 1090 different ICD-9 codes and students at CP sites logged 953 different ICD-9 codes.

Evaluation tools

During the 3 years of this study, students took an examination review consisting of an MCE in week 7 of the clerkship. Students were given 90 minutes to complete the examination. The MCEs were graded and results 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 emphasise knowledge of each of these diagnostic entities. This allowed 1 of the co-authors (FAM) to assign a single ICD-9 code to each test item to correlate with the logbooks.

For their final examination, students took the National Board of Medical Examiners (NBME) Subject Examination, a national standardised 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 paediatric medicine. Each of these test questions was not available for coding with the ICD-9 code.

Validity/reliability

The MCE was administered to 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 examination 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 according to the standardised curriculum for paediatric clerkships.

Analyses

The statistical analyses of the data consisted of contingency tables, which test dependence of categorised 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 1-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 analysed.

RESULTS

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

Various statistical analyses were performed on the available sample. Students were arbitrarily grouped according to 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 5 groups (≥ 90%, 80–89%, 70–79%, 60–69%, < 60%). We initially reviewed descriptive statistics to obtain a general overview of the data.

Contingency tables were used to summarise categorised data, such as the number 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 the variables tested were independent of one another. We found that patient exposures and examination scores on both MCE (UNMC students $\chi^2 = 14.672$, CP students $\chi^2 = 6.255$; both less than the test statistic of 21.026) and NBME (UNMC students $\chi^2 = 9.595$, CP students $\chi^2 = 11.303$; both 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 scores and patient exposure (Table 1).

Given the structure of Year 3, students completing their first clerkship in paediatrics had little or no prior clinical experience in paediatrics. Because of this, we applied the same testing using contingency tables and ANOVA for students who completed the clerkship at the beginning of the academic year and students who finished 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 was independent of patient encounters.

As students at CP sites tended to see a greater volume of patients, we applied similar tests to those described above for the 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 the number of 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 was computed at $r = 0.192$ for the MCE and at $r = 0.189$ for the NBME. This indicates a weak association between patient exposure and examination results. Analyses looking at test items coded V20.2 (health care 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 summarised in Table 2.

DISCUSSION

The revision of the paediatric curriculum at UNMC was met with a great deal of resistance when it was unveiled in 1994. Faculty members were concerned about a shift in focus to more ambulatory training because it was felt that students would not have enough patient exposure. To ensure that adequate numbers and types of patients were being seen, students were required to maintain a logbook of patient encounters. For the purposes of accreditation, the educational experiences and evaluation methods for this decentralised clerkship were carefully structured.

Table 1 Mean examination scores versus patient encounters

	MCE		***NBME***	
Number of patients seen	***UNMC***	***CP***	***UNMC***	***CP***
≤ 50 patients	63.95 (SD 8.39)	70.27 (SD 11.08)	73.04 (SD 6.33)	73.33 (SD 3.20)
51–100 patients	65.08 (SD 7.71)	69.24 (SD 13.24)	74.88 (SD 6.60)	73.00 (SD 7.62)
101–150 patients	67.19 (SD 8.73)	71.25 (SD 10.40)	76.42 (SD 8.02)	73.25 (SD 5.37)
> 150 patients	71.15 (SD 10.12)	68.85 (SD 8.68)	77.69 (SD 9.38)	73.88 (SD 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
MCE = multiple-choice examination; NBME = National Board of Medical Examiners; UNMC = University of Nebraska Medical Center; CP = community practice; SD = standard deviation

Table 2 Comparison of ICD-9 coded multiple-choice examination test items and average number of patients seen

Diagnosis (ICD-9 CM code)	*Mean correct score on pre-examination*	*Average number of 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
Health care maintenance (V20.2)	7.53 of 10	16.86 (Range 0–119)	0.032

*r = correlation between patient exposure and correctly answered questions
NEC = necrotising enterocolitis

When students began completing the paediatrics clerkship in clinics throughout Nebraska, the difference in clinical experiences was quickly apparent by the volume of patients students were logging. On average, students who participated in the CP training track logged an average of 163 patients, whereas 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 on the CP training track had a mean score of 73.45 (± 5.895) on the NBME Pediatric Subject Examination, and university track students scored a mean of 74.93 (± 7.211). Scores on the 60-point MCE averaged 41.68 (± 6.20) for CP track students and 39.48 (± 5.08) for university-based students. Performance on these examinations 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 209 ± 17, Step II 221 ± 18; class of 2000: Step I 213 ± 18, Step II 218 ± 21; class of 2001: Step I 216 ± 17, Step II 225 ± 20).

From the statistical analyses of patient encounters and examination performance, the results implied that examination performance on both the MCE and NBME was not dependent on the number of patient encounters logged. These results indicate that 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 (health care maintenance) and MCE performance on questions pertaining to this code, there was no demonstrated improvement in examination performance with increased patient encounters.

The limitations of this study should be addressed. Firstly, the use of historical controls may be questioned. Student performance on Medical College Admission Tests, as well as student performance during the first 2 years of medical school, may be a confounding variable that was not taken into consideration. Knowing that MCE results did not constitute part of student grading most likely impacted performance on that examination, which was probably taken less seriously than the actual NBME, which represented 30% of grade marks.

Another limitation is that this analysis did not take into account the roles played by students in the patient encounter (e.g. active versus passive roles). It focused solely on the number of encounters recorded. The amount of time spent with each patient may also influence student learning and may correlate with performance on a standardised examination. Again, this information was not collected.

Finally, this study involved the paediatrics clerkship at a single institution. Therefore, results may not be generalisable. 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.

Given that accreditation standards require quantified criteria for the types of patients seen during a clerkship (LCME ED-2 requirement), great care must be taken in the analysis of students' experiences. These results clearly show that, regardless of the numbers and types of patients seen, students performed similarly on knowledge-based examinations. Previous studies[8–10] have demonstrated experiential and didactic knowledge to be independent, although both are of huge importance. In terms of the grading and evaluation of

clinical courses, evaluations are integral to grading but greater emphasis continues to be placed on objective examination performance. Future investigation will include the development of more reliable mechanisms for assessing experiential knowledge acquisition during clinical courses, which, in conjunction with didactic knowledge, should provide for better assessment of medical student performance on clerkships.

Contributors: GLB and FAM were co-principal investigators for this project, and contributed towards collecting logbooks, completing statistical analyses and writing the manuscript. MTM oversaw the statistical testing, made recommendations on how best to analyse the data, co-wrote the Methods and Results sections and contributed towards the editing of the entire manuscript. CR organised the logbook and examination data, prepared the data for analysis, offered suggestions for approaches to the data analysis, and contributed towards the review and editing of the manuscript. KK was responsible for coding all patient logbooks and entering the information into a database, and contributed towards the editing of the manuscript.

Acknowledgements: none.
Funding: none.
Conflicts of interest: none.
Ethical approval: the University of Nebraska Institutional Review Board considered this study exempt under 45 CFR 46:101b, category 4. The IRB number is 140–04-EX.

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 versus 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 centre 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.

Received 6 July 2006; editorial comments to authors 10 October 2006, 21 November 2006; accepted for publication 4 December 2006