Learning About Screening Using an Online or Live Lecture

Does It Matter?

Anderson Spickard III, MD, MS, Nabil Alrajeh, MS, PhD, David Cordray, PhD, Joseph Gigante, MD

OBJECTIVE: **To determine the impact of an online lecture versus a live lecture on screening given to medical students who are participating in an outpatient clerkship.**

DESIGN: **Prospective, randomized, controlled study.**

PARTICIPANTS AND SETTING: **Ninety-five senior medical students in a primary care medicine clerkship based at university and distant clinic sites.**

INTERVENTION AND MEASUREMENTS: **Forty-eight medical students were randomized to the live lecture on screening (live lecture group), and forty-seven medical students were randomized to the online lecture on screening (online lecture group). Outcome measures included students' knowledge, use of time, and satisfaction with the lecture experience.**

RESULTS: **Compared to students in the live lecture group, students in the online lecture group demonstrated equal post-intervention knowledge of screening ($P = .91$) and expended 50 minutes less time to complete the lecture. Online lecture students who used the audio feed of the lecture were equally satisfied with the lecture as the live lecture students. Without the audio feed, online lecture students were less satisfied.**

CONCLUSIONS: **An online lecture on screening is a feasible, efficient, and effective method to teach students on outpatient clerkships about principles of screening.**

KEY WORDS: **medical education; students; lecture; education measurement; computer-assisted instruction.**
J GEN INTERN MED 2002;17:540–545.

W ithout question, the landscape of medical education is changing. New discoveries have led to a plethora of information resources that challenge all learners in medicine. Technological advancement keeps pace with the rate of change of information by providing rapid and efficient access to information. Meanwhile, trainees in medicine travel farther from the inpatient medical center to receive instruction.[1] Students in medicine, nursing, pharmacology, and other disciplines are required to complete community rotations to achieve learning goals

in the new context of medical care delivery, the ambulatory setting.

Given these trends, many medical educators are considering the use of computer-aided instruction (CAI) as a means of delivering curricular content to students. Uses range from the simple, such as the display of syllabus notes on web pages, to the complex, such as the use of multimedia interactive learning modules.[2–4] Enthusiasm for CAI depends on whether implementation of the new technology is feasible and whether its use is associated with gains in the efficiency of delivering course material and gains in the learning of students.

Many of our medical students complete rotations at distant sites and return to campus to attend required lectures. For these students, we asked the question whether we could deliver medical curricular content, specifically a core lecture, to our students on a clerkship at distant sites through the Internet using a relatively simple CAI tool.

METHODS

Objective

We hypothesized that medical students could expend less time but demonstrate equal satisfaction and knowledge in participating in an online versus a traditional classroom lecture on screening.

Setting

We conducted a prospective randomized controlled trial that involved senior medical students at Vanderbilt University who were participating in the Primary Care Medicine Core Clerkship. Students complete the clerkship by working with primary care physicians either at the Vanderbilt Medical Center or at distant practice sites in the region or in other states. Students are required to complete a core set of lectures on the Vanderbilt campus to gain credit for the course. If they are unable to attend the lectures during the clerkship, they attend makeup lectures at another time during the year.

Participant Eligibility and Enrollment

Approval for the study was obtained from the Dean of the Vanderbilt University School of Medicine and the Institutional Review Board. All senior medical students at Vanderbilt were eligible for the study. Enrollment occurred during each monthly orientation to the 4-week Primary Care Clerkship during the academic year from July 1999 to May 2000.

Informed Consent and Randomization

Students signed a consent form that specified that 1) participation and performance in the study were confidential and had no bearing on the final grade, 2) the clerkship co-directors and clinic preceptors were blinded to student participation and student group assignment in the study, and 3) students must learn the material of the lecture as part of their clerkship requirement. Block randomization was accomplished by randomizing a new group of students into 2 groups during orientation to each 4-week block. Students chose an envelope that included the group assignment and the student's identification number for tracking of results. To maintain randomization, students in the live lecture group were not given access to the online lecture until they had completed the live lecture. In addition, an assistant blinded to the study protocol and outcomes checked the list of participants in the live lectures against a list of students randomized to the online lecture. There were no instances of participation in the live lecture by a student randomized to the online lecture.

Intervention

The control group ($n = 48$) participated in a live lecture on screening; the intervention group ($n = 47$) participated in an online lecture on screening. The same faculty member who has given the lecture in the clerkship for the past 5 years conducted the live lecture. The live lecture occurred in a conference room and included 10 to 15 participants. The lecturer began with a didactic introduction followed by the presentation of 4 patient cases. The lecturer then led a group discussion of the 4 cases. During the discussion, the lecturer provided structured information about approaches to the patient cases that included 3 major concepts about screening derived from the United States Preventative Services Task Force (USPSTF) on screening: the burden of suffering of the disease, the efficacy of the screening tests for the disease, and the effectiveness of early detection in the treatment of the disease.[5] A handout that outlined the material covered during the lecture was provided at the conclusion of the session.

The online lecture presented the same information, in the same order, as the live lecture. It consisted of an Internet-based Power Point slide presentation from the same lecturer with the option of audio. It began with a didactic introduction of the same material as the live lecture. Then the same 4 cases were presented to the learners with the same discussion questions for consideration that were posed in the live lecture. Next, identical information about approaches to the 4 cases following the USPSTF format was displayed. Following the online lecture, students were able to review the slide material at any time. In the online group, there was no interaction between lecturer and students through online chat rooms or other online discussion groups. Students completed the online lecture on their own time at the site of their choosing, such as home, school, or the office practice site.

Outcomes

The primary outcome variables were students' knowledge of screening, time used to complete the lecture assignment, and satisfaction with the lecture presentation. Pre-intervention variables were obtained to determine the demographic profile of the 2 study groups.

Pre-intervention Measures. The pre-intervention measurement battery included 1) a survey on demographic attributes of participants, 2) a survey on computer skills, and 3) a pre-test on knowledge of screening. The computer skills survey queried participants' access to computers, their attitudes about the use of computers for their education, and their skill and attitudes in using computers. Computer skills of participants were determined by asking them to rank their skill level on 15 computer tasks. A summary score on the 15 computer tasks was derived. The summary scale showed high internal consistency (Cronbach's $\alpha = 0.88$) and was used to classify each participant's skill level as either beginner, intermediate, or advanced. A beginner computer user would be, for example, very familiar with creating and moving files, have some familiarity with creating handouts, but have little understanding of using e-mail attachments or virus scans. An advanced computer user would be, for example, quite adept at creating power point slides and efficient at browsing the Internet, writing html language, and using spreadsheets. Finally, participants were asked to rate the importance (not important, important, very important) of each of the 15 computer tasks in order to determine their attitudes regarding the importance of these computer tasks. The summary scale for ratings of importance also revealed adequate reliability (Cronbach's $\alpha = 0.82$).

The pre-test on knowledge of screening was a 12-item exam that had been used in the course for 1 year. The exam included 10 multiple-choice items and a discussion question worth 2 points that covered information about screening that would be presented in the subsequent lecture. Analysis of the internal consistency of this pre-test measure of screening revealed only modest reliability (Cronbach's $\alpha = 0.38$), uniformly low inter-item correlations (φ coefficients), and an average correct response over the 12 items of 47%.

Post-intervention Measures. Post-intervention outcome measures included measures of time, satisfaction, and a post-test of knowledge of screening. Students were asked to report their travel time to and from the lecture, time to access the online lecture on the computer, time to participate in the lecture, and time used in reviewing reference materials after the lecture. Students rated their satisfaction with their lecture on a 5-point Likert scale, and they provided narrative comments about their experience with the live lecture and the online lecture.

The post-test knowledge of screening was an open book exam that has been used in the course for 1 year.

Students in the live lecture group were allowed to use their lecture handouts and other personal materials, but not the online lecture material, to assist them in completing the post-test. Students in the online group were allowed to review the online lecture and other personal information resources to assist them in completing the post-test. The test included 4 discussion questions that solicited approaches to 4 hypothetical patients. Students had to demonstrate understanding of the 3 key concepts about screening (burden of suffering, efficacy of screening, and effectiveness of early detection) derived from the lecture information in their approach to these 4 hypothetical cases, for a maximum of 12 concepts.

Two raters who were blinded to each other and to student group assignment graded the post-tests. Disagreements about the grading of an exam question were resolved after discussion. Prior to rendering a final grade for each question, reliability of the post-test grading was determined. Reliability was assessed in 2 ways (inter-rater agreement and internal consistency) and involved 2 forms of grading (item-by-item grading and summary grade for each hypothetical case). The item-by-item scoring revealed an inter-rater agreement of 77%. When the items were summated to create a knowledge score, the internal consistency of the resulting scale was 0.45. That is, although the 2 raters were fairly consistent in their ratings of individual items, as whole, the individual items reflected only moderate internal consistency. Because both indices of reliability were lower than is desirable, the narratives were re-rated using a 4-point summary score for each of the hypothetical cases, resulting in a total possible post-test score of 16. The summary rating scheme revealed higher inter-rater agreement (97%) and higher internal consistency (Cronbach's α = 0.77) than the item-by-item scoring protocol. Because the psychometric properties were better for the summary ratings (as compared to the item-by-item scoring) they were used in the outcome analyses.*

Statistical Analysis

For all measurements, a 2-tailed *P* value of <.05 was considered significant. The baseline comparisons between groups were analyzed with χ^2 analysis for differences in proportions and *t* tests for differences in means. The post-intervention comparisons for time, satisfaction, and post-test knowledge were analyzed with χ^2 analysis for differences in proportions, and *t* tests and analysis of variance for differences in means. Multiple regression analyses were conducted to examine the influence of covariates (age, level of computer skill, attitudes toward computer tasks, beliefs in the role of computers in medical education, pre-test knowledge of screening, and time elapsed between the intervention and taking the post-test) on key outcome variables (e.g., knowledge and satisfaction). All statistical analysis was performed using SAS software (version 6; SAS Institute, Inc., Cary, NC).

As noted earlier, 48 and 47 participants were randomly assigned to the live or online conditions, respectively. We considered a 10% difference between the groups in the post-test of knowledge to be educationally significant.[6,7] The study was designed with a sample size of 90 students (45 students per treatment group) to have 80% power (2-sided test at α = 0.05) to detect a 10% difference in the post-test scores of knowledge between the groups.

RESULTS

One hundred out of a possible 105 students in the senior class agreed to participate in the study. Ninety-five students completed the study protocol. There were no statistical or educationally meaningful differences between the groups on all baseline items including demographic items, access to the computer, attitude about the use of computers for education, skill in using the computer, attitudes about the importance of 15 computer skills, and pre-test of knowledge of screening. See Table 1 for a baseline comparison of the groups.

Knowledge

The main outcomes of the study are shown in Table 2. The post-test on knowledge of screening was taken on average 2 weeks after the lecture (live group 2.2 ± 1.6 weeks, online group 1.8 ± 1.0 weeks; *t* [81] = −1.5; *P* = .93). There

---

* These analyses point to the importance of assessing the psychometric properties of scales that are used in these types of studies. Although raters agreed moderately well on the item-by-item grades, as an index of knowledge, the items did not represent a reliable index. Using the holistic summary rating (1 to 4 for each narrative) improved the internal consistency of the knowledge scale. Increasing the reliability of the scale increases the sensitivity of the analyses in detecting between-group effects if they exist.

Table 1. Baseline Comparisons of the Groups

	Live Lecture (*N* = 48)	Online Lecture (*N* = 47)	*P* Value
Mean age, y	26.7	26.2	.32
Male gender, %	60	65	.62
Site of clerkship (away at distant site), %	34	35	.88
Computer access at home, %	72	83	.13
Computer attitude ("The effect of computers on my education is beneficial or highly beneficial."), %	96	92	.55
Computer ability, %			
Beginner	30	17	.27
Intermediate	45	58	
Advanced	25	25	
Pretest knowledge (SD)*	6.0 (1.8)	6.4 (2.1)	.27

* The pretest contains multiple-choice items and a discussion question for a possible 12 points.

Table 2. Knowledge, Time, and Satisfaction Outcomes

	Live Lecture (N = 48)	Online Lecture (N = 47)	P Value
Post-test knowledge (SD)*	10.8 (2.8)	10.7 (3.0)	.91
Time (minutes)†	113	61	<.001
Satisfaction‡	4.4		.002
With audio		4.5	
Without audio		3.8	

* *The post-test contains discussion questions for a possible 16 points.*
† *Time is the sum of time to travel to the lecture, access the lecture, and participate in the lecture.*
‡ *Satisfaction is measured on a five-point Likert scale: 1 = very dissatisfied; 5 = very satisfied.*

was no difference in the summary post-test scores between the live group (10.8 ± 2.8) and the online group (10.7 ± 3.0); $t(93) = -0.11$; $P = .91$). Multiple regression analyses showed that age, level of computer skill, attitudes toward computer tasks, beliefs in the role of computers in medical education, pre-test knowledge of screening, and time between the intervention and taking the post-test were unrelated to status on the summary post-test measure of knowledge of screening ($F(8,86) = 1.33$; $P = .24$). Only 1 variable, the time it took to answer the questions, was related to post-test performance; the zero-order correlation was substantial ($r = .52$; $P < .0001$), suggesting that this may not be a chance finding.

Time

The sum of the time to travel to the live or online lecture, access the lecture, and participate in the lecture was higher (mean = 113 minutes) for the live lecture group than for the online lecture group (mean = 61 minutes) ($P < .001$). All of the time saved derived from less time participating in the lecture. Students in the live group used more time in traveling to and from the lecture (19.3 minutes in the live group versus 7.9 minutes in the online group), but no time accessing the computerized lecture, which involved downloading the lecture from the Internet to a computer (0 minutes in the live group versus 12.5 minutes in the online group). Both groups used the same amount of time in reviewing reference materials (26.1 minutes in the live group versus 22.5 minutes in the online group; $P = .43$) and in answering the questions (19.7 minutes in the live group versus 17.9 minutes in the online group; $P = .38$).

Satisfaction

Ninety-six percent of the live lecture participants versus 81% of the online lecture participants were satisfied or highly satisfied with their lecture ($P = .03$). But only 31% of the students randomized to the online group used the audio feed of the lecture. Ninety-three percent of the students in the online group who used audio were satisfied or highly satisfied with their lecture experience compared to only 76% of the students in the online group who did not use the audio feed. Across the 3 groups—live ($n = 47$), online with audio ($n = 15$), and online without audio ($n = 33$)—the average ratings for satisfaction were 4.4, 4.5, and 3.8, respectively ($F_{2,94} = 6.90$; $P = .002$). Using multiple regression analysis to examine satisfaction of only those participants in online condition reveals that the presence of audio feed increased satisfaction ($b = 0.753$; $t = 2.62$, $P = .012$), but satisfaction was negatively influenced by the level of computer sophistication ($b = -0.558$; $t = -2.47$; $P = .018$). Other variables, such as total time devoted to the learning material, familiarity with and importance of computer-related tasks, and post-test performance were unrelated to satisfaction within the online group. The overall regression equation accounted for 30% of the variance in satisfaction scores ($F_{7,47} = 2.84$; $P = .017$).

A review of the comments from the students with more advanced computer skills assigned to the online group revealed that students appreciated the movement of the lecture to the online format, but they wanted the "bugs to be worked out." All of the negative sentiment was directed at improving the audio feed, shortening the time to download the lecture from the Internet, and improving the quality of the slides. No student requested more sophisticated display of information with graphics, video, or Internet links.

DISCUSSION

In this study, students who participated in an online lecture on screening saved time and achieved knowledge scores similar to those of students who participated in a traditional lecture on screening. Not surprisingly, students in the online group were less satisfied when audio was absent from the lecture.

While many universities and businesses have adopted an Internet-based slide presentation and audio feed of lectures to convey didactic material, there exist few head-to-head comparisons of the computerized lecture format with the traditional live classroom lecture format.[8] Instead, computer instructional modules that include simulations, individual exercises, and quizzes with feedback have been compared to traditional lectures.[9–11] Comparison studies of CAI have been criticized for lacking background information about study participants and for making comparisons between more instruction with less instruction or between distinctly different interventions developed by different authors using outcome measures that have not been validated.[12,13]

Design flaws that have plagued other studies of CAI were addressed in this study in the following ways. Randomization and characterization of students allowed for the balance of potential confounders, including participants' demographic background, access to computers, skill in using computers, and attitudes about the use of computers for education. The same teacher developed a similar instructional design of both interventions in terms of how the concepts about screening were defined, the order in which they were presented, and the patient cases

to which they were applied.[14] During the lecture, similar questions for learner reflection were posed to both groups. This helped learners to determine what they needed to know about screening and to find that information provided to them from the lecture. Finally, outcome measures were monitored for internal consistency.

The intervention in this study was feasible. Most studies of CAI involve robust multimedia "virtual learning environments" with text, graphics, animations, videos, or self-directed quizzes that require scores of hours to develop.[3,15,16] We estimate that it took 3 hours to develop the lecture presentation on the Internet, including 2 hours to dictate the lecture and 1 hour to upload the lecture. Newer software packages allow immediate audio capture of live presentations and faster mounting to the Internet. Current technology also allows immediate video capture of lectures for the Internet so that students are able to access video lectures anywhere at anytime. Students avoid having to check out or purchase videotapes from their instructors and they are able to start and stop and skip back and forth through the video material with the ease of a mouse click. Widespread use of video lectures awaits further proof of effectiveness and technological advancements to improve the speed and quality of Internet transmission of video material.[17]

The intervention in this study was efficient. End users saved time to achieve the learning goal of the lecture. The lecturer, a practicing internist, saved time and money by replacing lecture time with clinic time during the months following the study, when the live lecture was officially replaced by the computerized version of the lecture in the course.

Limitations caution the interpretation of these study results. First, the interventions contained the same instructional design, but they were not identical. The live lectures differed among sessions, depending on the audience discussion that was generated. The online lecture did not contain an electronic audience discussion component. Therefore, the lecture experience was shorter for the online participants. However, this did not lead to a difference in knowledge measures between the groups. It is possible that group discussion is less necessary to achieve knowledge goals from an introductory lecture on screening than may be necessary with more advanced lecture topics. Next, the relatively short period of time between the intervention and taking the post-test (2 weeks) allows the assessment of the impact of the instruction, but does not allow measurement of lasting effects.[18,19] Third, we have concerns about the students in the online group who did not access the audio feed. Subgroup analysis of all demographic variables showed 2 differences between the students in the online group who did not use audio compared with students who did use audio. Fewer non–audio users had computers at home compared to audio users (76% vs 100%; $P = .04$), and fewer non–audio users had advanced computer skills compared to audio users (18% vs 40%; $P = .05$). It is possible that students failed to use the audio feed because

of lack of access to and lack of sophistication with audio material transmitted over the Internet. These students spent 15 minutes less time reviewing the lecture than students who used the audio feed (data not shown). Listening to the lecturer, students in the audio group were likely to sit through the 45-minute presentation. Without audio, students likely advanced the slides at their own pace and moved quickly through the presentation. Our study results showed that the students who did not use audio were less satisfied with the experience. The study was not sufficiently powered to determine post-test knowledge differences between audio and non-audio computer users. We are conducting a follow-up randomized trial of an audio version versus a non-audio version of a computerized lecture to test the hypothesis of improved satisfaction and learning with the addition of the audio feed.

Because a new technology is available doesn't necessarily mean that one should use it. Technology should be used to solve real, defined problems. This study demonstrates that an online lecture that includes didactic and case-based material is efficient and associated with similar post-test scores of knowledge compared to a live lecture in delivering curriculum on screening to senior medical students at local and distant sites. This study also demonstrated that a live lecture on screening remains an effective means of teaching clinical clerks. Students learn from this and are highly satisfied even if it takes extra time to complete. Although most of the students in this study were satisfied with the online lecture, not all were. Future studies will address students' concerns, learning styles, and determinants of their satisfaction, as well as the cost utility associated with the use of the CAI tools.

REFERENCES

1. Kassirer JP. Preserving quality as the site of medical education shifts. Acad Med. 1998;73(10 suppl):120–2.
2. Hilger AE, Hamrick HJ, Denny FW Jr. Computer instruction in learning concepts of streptococcal pharyngitis. Arch Pediatr Adolesc Med. 1996;150:629–31.
3. Carr MM, Reznick RK, Brown DH. Comparison of computer-assisted instruction and seminar instruction to acquire psychomotor and cognitive knowledge of epistaxis management. Otolaryngol Head Neck Surg. 1999;121:430–4.
4. Pusic M, Johnson K, Duggan A. Utilization of a pediatric emergency department education computer. Arch Pediatr Adolesc Med. 2001; 155:129–34.
5. US Preventive Services Task Force. Guide to Clinical Preventive Services. 2nd ed. Baltimore, Md: Williams and Wilkins; 1996.
6. Christenson J, Parrish K, Barabe S, et al. A comparison of multimedia and standard advanced cardiac life support learning. Acad Emerg Med. 1998;5:702–8.
7. Fincher RE, Abdulla AM, Sridharan MR, Houghton JL, Henke JS. Computer-assisted learning compared with weekly seminars for teaching fundamental electrocardiography to junior medical students. South Med J. 1988;81:1291–4.
8. Umble KE, Cervero RM, Yang B, Atkinson WL. Effects of traditional classroom and distance continuing education: a theory-driven evaluation of a vaccine-preventable diseases course. Amer J Public Health. 2000;90:1218–24.
9. Maki RH, Maki WS, Patterson M, Whittaker PD. Evaluation of a

Web-based introductory psychology course: I. Learning and satisfaction in on-line versus lecture courses. Behav Res Methods Instrum Comput. 2000;32:230–9.

10. Rogers DA, Regehr G, Yeh KA, Howdieshell TR. Computer-assisted learning versus a lecture and feedback seminar for teaching a basic surgical technical skill. Am J Surg. 1998;175:508–10.

11. Bachman MW, Lua MJ, Clay DJ, Rudney JD. Comparing traditional lecture vs. computer-based instruction for oral anatomy. J Dent Educ. 1998;62:587–91.

12. Keane DR, Norman GR, Vickers J. The inadequacy of recent research on computer-assisted instruction. Acad Med. 1991;66:44–8.

13. Clark RE. Dangers in evaluation of instructional media. Acad Med. 1992;12:819–20.

14. Dick W, Carey L. The Systematic Design of Instruction, 4th Ed. New York: HarperCollins Publishers; 1996.

15. Bridges AJ, Reid JC, Cutts JH III, Hazelwood S, Sharp GC, Mitchell JA. A comparative study of computer-assisted instruction for rheumatology. Arthritis Rheum. 1993;36:577–80.

16. Devitt P, Worthley S, Palmer E, Cehic D. Evaluation of a computer based package on electrocardiography. Aust N Z J Med. 1998;28: 432–5.

17. Wofford M, Spickard A III, Wofford J. The computer-based lecture. J Gen Intern Med. 2001;16:464–8.

18. Summers AN, Rinehart GC, Simpson D, Redlich PN. Acquisition of surgical skills: a randomized trial of didactic, videotape, and computer-based training. Surgery. 1999;126:330–6.

19. D'Alessandro DM, Kreiter CD, Erkonen WE, Winter RJ, Knapp HR. Longitudinal follow-up comparison of educational interventions: multimedia textbook, traditional lecture, and printed textbook. Acad Radiol. 1997;4:719–23.

◆