Weekly Open-Book Open-Access Computer-Based Quizzes for Formative Assessment in a Medical School General Pathology Course

Margret S. Magid, M.D.¹, Merril K. Schindler, M.A., M.S.L.S.²

1department of pathology
2Gustave L. and Janet W. Levy Library

 (+)1-212 241-7459 (+)1-212-289-2899 margret.magid@mountsinai.org

ABSTRACT

Frequent testing, enhanced by computer delivery, provides a valuable means of formative assessment through timely review of course material, prompt feedback and image display. We introduced weekly computer-based quizzes in a medical school General Pathology course. Quizzes were released for several days following each of five weekly modules and represented, in total, 10% of the final grade. We hoped to further enhance the value of frequent computer quizzes by introducing two new features. First, we offered independent, open-access scheduling with the ability for the student to take the quizzes independently at any site with Internet access. By doing so we conserved in-class hours for instruction and eliminated the need for a designated testing site and faculty supervision. Secondly, we permitted an open-book format to encourage directed course review and decrease stress. Data were generated on specific questions from the quiz and on the entire quiz by individual student and class. Prompt analysis of results permitted timely remediation of problematic topics and identified students at academic risk. In the three years of study, student feedback has been highly favorable, particularly with regard to reinforcing understanding of topics for study, including both those formally presented in class and those assigned as independent learning, and for the impetus to “keep up” with course material. Performance on the final examination showed a statistically significant improvement after introduction of the quizzes. Weekly quizzes, enhanced by self-scheduled computer delivery and open-book format, are a valuable teaching tool for formative assessment.

INTRODUCTION

Formative assessment, that is, the process by which feedback is utilized to modify ongoing teaching and learning, has been lauded as an important process to enhance learning¹. Feedback is a critical aspect of formative assessment, in that it provides information on what the student has learned and can dictate both teacher- and student-generated actions to correct deficiencies. Frequent testing has been suggested as one of the ways formative assessment can be offered to students, as it provides both the means by which material can be assessed shortly after delivery and also the possibility of timely feedback to both students and teachers²,³. Other reported benefits of frequent evaluation include improvement of student performance on final examinations⁴-⁸ and reduction of examination anxiety²,³.

Computer-based delivery of frequent quizzes is a further modification recently introduced within the medical school curriculum⁹,¹⁰. The computer offers the additional advantages of immediate scoring and feedback, thus increasing the effectiveness of formative assessment, as well as increased flexibility and enhanced linking of images and other media⁹,¹⁰.
We introduced weekly computer-based quizzes into the General Pathology course at Mount Sinai School of Medicine and currently have three years experience with this initiative. Our various objectives reflected both the previously recognized advantages of frequent testing, as well as specific issues within the General Pathology course.

The General Pathology course is a six-week course given in the spring of the first year to approximately 120 students. The course introduces basic pathological concepts, introduced as weekly modules covering Cell Injury, Inflammation and Repair, Hemodynamic Disorders, Immunopathology, and Neoplasia, within a 22-hour schedule utilizing a mixed lecture/small group/laboratory format. The course concludes with a two-hour, in-class, multiple-choice examination. Due to limitations of time, there had been no additional formal testing within the course. An inordinate eighty-five percent of the student's final grade was determined by performance on the final examination, with instructor evaluation of student laboratory/small group performance contributing the remaining 15%.

There were several disadvantages inherent in the course format. The students had no opportunity prior to the end of the course to formally assess their retention and understanding of course material. The students were also concerned about related practical considerations – while a few sample questions from previous examinations were available to the students for review, details on examination question content and level of difficulty were not obvious. Another consequence of the limited course hours was that only a portion of the General Pathology material that the faculty deemed important could be presented by formal didactic instruction. The students were required to acquire the remaining material through their textbook readings – a task that appeared to the faculty to be indifferently performed by many students. An additional issue for the pathology faculty was that General Pathology runs contemporaneously, at least in part, with several other first-year courses, including Immunology, Pathogenesis and Mechanisms of Host Defense, Epidemiology, and the Art and Science of Medicine. The students were inclined to devote the major part of their studying to these larger courses at the expense of General Pathology.

In sum, specific objectives we hoped to address in weekly quizzes included: 1) Provide formative assessment to medical students on material both presented in class and assigned for independent study; 2) Reduce the importance of the single final examination on the student grades; 3) Provide examples of topics, and their level of detail, considered to be important by the faculty; 4) Encourage weekly review of the course material (particularly in view of the competition presented by the other major courses); and 5) Provide image reinforcement.

In addition to the above, we chose to introduce two elements into our computer quiz program that had not been previously described and that we hoped would make it easier to administer the quizzes and improve the learning experience for the students. First, we offered the students self-scheduling of quizzes and the ability to take them independently at any site with Internet access. By this means we could avoid both use of class time for quizzes and the necessity of arranging for Medical School space and personnel to administer the exam. Secondly, we allowed the students to refer to their books during the quiz, which we believed would facilitate directed learning in a less stressful environment.

MATERIALS AND METHODS

One quiz was designed for each of the five major modules of the course, and was rewritten each year. Each quiz consisted of 10-20 case-based questions that were frequently linked to both gross and microscopic images. Our computer system is WebCT™. Each question was created using an html table, which allowed it to be transferred to the exam with its format intact. This allowed images with legends to be included, which is more difficult when done directly within the system. Answer format was multiple-choice (single-best-answer) or matching (Figure 1). General Pathology, as every course at the Mount Sinai School of Medicine, has its own website within the WebCT™ course management system, available on our Intranet. Access by students to the site is secured by login and password. Each quiz was released on the course website following the in-class formal presentation of the week's material, which occurred on Mondays and Tuesdays of each week. The quizzes could be accessed at the student's convenience from any computer – indeed, one student took the quiz while at home in California!

Sunday evening. Although not required, the quizzes were emphasized to the students as an important tool for self-study. The quizzes were to be taken independently, but were open-book. During the open period, students were allowed unlimited access to the questions, until such time that they formally submitted their results. Each Monday morning, the quizzes were scored by computer and released to the students with their grades and the correct answers. On Monday evenings, second-year medical student teaching assistants met with the General Pathology students to review the previous module's material and go over questions they might have about the computer quiz from the previous week. A value of 10% of the final grade, or 2% per quiz, was selected to provide a "safety cushion" toward the final grade, yet minimize the harm to the final grade if a student was unable to complete one of the quizzes due to conflicting academic or personal obligations. Each quiz remained available for review for the remainder of the course, even for those students who had not formally
Group and individual student compliance and performance were analyzed with the WebCT™ toolset. Data analysis provided the discrimination factor for each question, as well as the frequency that each answer was selected, which permitted analysis of the students' thought processes. The duration of time each student had the quiz site open was also documented. Specific questions relating to the computer quizzes were incorporated into the student course evaluations, required at the end of the every course at Mount Sinai School of Medicine, for analysis of student opinions. Quiz and final examination grades were reviewed to determine if the individual student performance on quizzes within a given year correlated with performance on the final examination. Additionally, the whole class performance on the final examination (means and standard deviations) was analyzed by a 2-sample t-test comparing the two years prior to the introduction of quizzes to the years afterwards.

The Associate Dean of Undergraduate Medical Education, Mount Sinai School of Medicine, submitted and received Institutional Internal Review Board (IRB) waiver to use any student survey data for publication.

RESULTS

Student compliance in taking the quizzes ranged from 95-99%. Results were generated on specific questions and the entire quiz by individual student and class, with the mean grades for individual quizzes ranging from 86 to 98%. We observed that student compliance in taking the test was highest in the first year (99%), and for the subsequent two years dropped a little to and stayed stable at 95%. We believe that the exceptionally high student compliance in the first year reflects the novelty of the computer quizzes in the curriculum. In subsequent years other courses in the first year of the medical school followed the example of General Pathology and introduced computer quizzes. The mean performance on the five quizzes in the first year (94%) was slightly higher than in the following two years (90 and 89%, respectively), probably for similar reasons.

Analysis regarding performance by the upper and lower 25% of the class allowed calculation of a discrimination factor for each question. The results led to several interpretations: 1) when both the upper and lower 25% of the class scored high, the concept was deemed to have been learned by all; 2) when the upper 25% performed well and the lower 25% poorly, the high discrimination factor indicated a concept that was more difficult and worthy of additional emphasis for the weaker students (Table 1); 3) when both the upper and lower 25% scored poorly, the conclusion was either that the topic was inadequately covered by the course and presented by the textbook in a confusing manner, or that the question was poorly written. If it was determined that a question was poorly written, it was discarded and the students were not penalized for incorrect answers.

In the infrequent event that a student elected not to take a quiz, the score for that quiz was recorded as zero, which was added to the rest of the scores to arrive at a final quiz grade. It was very unusual that a student scored less than 75% on any individual quiz. Review of the student performance in this circumstance indicated that the problematic questions were generally based on assigned readings and not on topics discussed in class. When we looked at the time that the student devoted to the quiz, we frequently found that the student spent less than an hour, indicating in our opinion less of an inability to answer specific questions than a disinclination to spend the time reading the appropriate section in the textbook. As noted earlier, following the quizzes students were provided with correct answers and the assistance of teaching assistants (and faculty, if necessary) for clarification of problematic questions.

The timely generation of test results and analysis permitted prompt remedial action as indicated – either by email to the class or in-person by the Course Director at the next lecture. The Course Director noted that an unexplained failure of a student to take more than one quiz often correlated with academic or personal difficulties that potentially dictated various levels of intervention, ranging from contact by email or personal conference with the Course Director, to involvement of the medical school deans.

We followed two different methods of evaluating student feedback. In the first two years that we offered computer quizzes, we asked two relatively general questions: “Is it worthwhile to continue the quizzes as a means of teaching and reinforcing selected course material?”, and “Is it worthwhile to continue the quizzes as a percentage of your final grade?”. Affirmative answers to both questions averaged 90% of the class, and we did not find that there was any significant change from the first year to the second year. In the third year of our program, we posed different and more specific questions to the students and utilized a Likert score in an attempt to generate a more detailed analysis (Table 2). As demonstrated, responses were generally favorable. The strongest level of agreement was seen for the role of quizzes in reinforcing understanding of topics for study, including both those formally presented in class and those intended as
independent learning. Surprisingly, the value of quizzes in decreasing anxiety for the final examination, either by giving examples of questions or by providing partial credit toward the final course grade, was deemed the least important justification. In free text many students volunteered that the quizzes forced them to keep up with the course material. The teaching assistants reported that attendance was low at their weekly reviews of the quizzes; most students volunteered that the feedback they had received on incorrect answers was sufficiently clear to make additional review unnecessary.

The mean and standard deviation of the student performance on the final examination for the years prior to the introduction of the weekly quiz program were 86.5 and 5.90, respectively, and for the years after were 89.8 and 5.45. By a 2-sample t-test the improvement in the mean grade was significant at a p value of <0.001. Correlation between quiz grades and performance on the final examination for individual students could not be assessed, since the grade spreads on both the quizzes and final examination were too narrow to permit analysis.

DISCUSSION

The use of frequent quizzes has been recommended for a variety of didactic purposes that relate in various ways to formative assessment. In the literature, the effect of frequent testing on final grades has been generally beneficial, although not uniformly so. However, other educational objectives of frequent testing have been identified, including providing early identification of individual students with problems or more generally problematic course material requiring clarification, reducing student anxiety about examinations by encouraging regular studying and providing familiarity with types of questions asked, and generally engendering an improvement in student approaches and attitudes toward learning by encouraging a consistent pattern of studying. Furthermore, in several studies students commonly favored frequent quizzes, which they claimed help them keep up with work, and were more likely to favorably rate courses offering frequent testing.
Figure 1. Question from the Quiz on Cell Injury. This matching question was designed for reinforcement of material presented in the class and the textbook. It includes a microscopic image for identification.

A 5-year-old boy received a small intestinal transplant two weeks ago for short gut syndrome following neonatal necrotizing enterocolitis. He is now experiencing diarrhea. An endoscopic biopsy of small intestinal mucosa is performed to evaluate the cause of his graft dysfunction.

Match the following image taken from this biopsy with either necrosis or apoptosis.

Then choose either necrosis or apoptosis for each of the following characteristics of cell death.

Matching pairs:

- Genetically determined internal mechanism of cell death.
- Deletion of self-reactive T cells in the thymus during development.
- Incites an acute inflammatory response
- Fas receptor activation
- More frequently involves multiple cells in continuity

Choose match
Necrosis
Apoptosis
The computer is recognized as a powerful educational tool, and has been utilized in innumerable on-line didactic exercises and tutorials. Within the last ten years computer-based quizzes have been introduced into medical school courses at a few institutions. These institutions observe that computer-based quizzes have certain advantages over paper-based testing, including flexibility of delivery that allows closer integration with instructional material, ease of incorporating images and other media into questions, providing immediate feedback, and facilitation of item banking. The programs of computer-based quizzes described in the literature offer mastery type of quizzes at special supervised computer testing centers at specific times that are arranged in advance.

Our computer-based tests differed from these programs by having an open-book format and open scheduling. Students were free to access the quiz at any time (within a five-day "window") and from any location with an Internet connection, providing scheduling flexibility to the students, eliminating the necessity for proctored supervision, and conserving in-class hours for instruction. The open-book format was selected to encourage reading and study in a stress-free setting. Specific questions relating to weekly reading assignments encouraged review of the material, which could be done during the course of the open-book examination. Inclusion of such subject matter, which had not been presented formally in course contact sessions, encouraged the self-directed study that is implicit in the current emphasis on lifelong learning. The students were instructed to work independently. Some of the faculty expressed skepticism that the students would refrain from collaborating on answers. While this was a risk, there were several reasons that collaboration was not likely: 1) the quizzes were "low stakes"; 2) Mount Sinai School of Medicine operates on an Honor System; and 3) the students knew that images and questions from the quizzes, in shuffled context, would appear on the final examination. Prompt feedback heightened the value of the quizzes for formative assessment of the students. Individual students had the opportunity to clarify promptly areas of confusion with teaching assistants or with the course faculty. For the instructors, immediate analysis of student answers also permitted timely intervention on occasional topics shown to be problematic to the class. The computer-based format also provided the opportunity to introduce and reinforce images, which is particularly desirable in a Pathology course. While the Course Director anticipated identifying students in academic trouble following poor performance on the quizzes, an unexpected result of the process was identification of individuals with personal difficulties who had repeatedly failed to take the quizzes. Student compliance was high and feedback was favorable. In evaluation surveys, students specifically praised the quizzes for highlighting topics deemed important in the course and providing them with a means for regular self-assessment and motivation to keep up. We were unable to correlate the individual student performance on quizzes with their performance on the final examination, due to generally high grades. However, we noted a statistically significant improvement in the mean class grades on the final examination in the years following introduction of the computer quiz program.
Based on our experience with the computer quizzes, we made several changes in subsequent years to facilitate smooth functioning of the program. We conducted more rigorous pre-quiz review to eliminate ambiguous questions. A set of detailed instructions on how to access the quiz and submit answers was added as an introduction to the first computer quiz. We were able to anticipate and warn students about procedural difficulties, such as the inability to access quizzes if the students had blocked “pop-ups” on their computers. The option of make-up quizzes was eliminated as it was too disruptive to the timetable of the quizzes. (We found that the students accepted this policy if it was explained in advance.)

CONCLUSIONS

By both student and faculty assessment, weekly computer-based quizzes provided valuable formative assessment in General Pathology. We found that the two new features we introduced into our computer quiz program enhanced its usefulness for both faculty and students. The introduction of independent self-scheduling of quizzes conserved in-class hours for didactics and eliminated the need for faculty supervision and a designated testing site. Open-book format encouraged directed weekly review of course material, including topics to be learned by independent reading assignment, in a stress-free environment. Favorable student evaluations demonstrate that computer-based open-book quizzes provide enriched curriculum support and student satisfaction. Mean class performance on the course final examination showed a statistically significant improvement following introduction of the weekly computer quizzes, providing an objective measure of its beneficial effect on learning.

Acknowledgment: The authors thank John Doucette, PhD., for assistance in statistical analysis.

REFERENCES

Open-ended feedback comments by students in 2007 provided evidence that the formative assessments were achieving their aims. Relevant examples included assertions that each assessment "provides an opportunity to correct misconceptions and actually learn/relearn concepts." and that there "is integration of a variety of concepts learnt throughout the course", which "gives good insight into students' current performance and areas of potential improvement which can be addressed for the actual exam." In order to determine whether students benefited more from formative assessments in our vertically integrated medical program during their first or second year, the analysis was stratified according to year of commencement (Table 4). We found no consistent
Purpose The Hofstra Northwell School of Medicine (HNSOM) uses an essay-based assessment system. Recognizing the emphasis graduate medical education places on the United States Medical Licensing Examination (USMLE) Step exams, the authors developed a method to predict students at risk for lower performance on USMLE Step 1. Method Beginning with the inaugural class (2015), HNSOM administered National Board of Medical Examiners (NBME) Customized Assessment Service (CAS) examinations as formative assessment at the end of each integrated course in the first two years of medical school. Chicago. Warner, Zachary B. "Adoption of Computer-Based Formative Assessment in a High School Mathematics Classroom." In Cases on Emerging Information Technology Research and Applications, ed. Mehdi Khosrow-Pour, D.B.A., 333-348 (2013), accessed July 19, 2019. doi:10.4018/978-1-4666-3619-4.ch016. Export Reference. This case follows a high school mathematics teacher who is new to the classroom and is looking to adopt computer-based formative assessment as a part of his curriculum. Working within the confines of the school environment, this requires navigating a shrinking budget, colleagues that do not share his value of technology, restricted time, student issues, and limited resources. Formative assessment—Students receive computer generated feedback on their performance, and teachers can measure the effectiveness of their teaching. Summative assessment—Candidates have to pass the test to progress in a course or gain a particular qualification or accreditation. A recent survey of UK medical schools and royal colleges revealed a growing interest in the use of computer based testing (unpublished data from electronic survey by W Irish and P Cantillon, 2004). Administration of open-ended test questions by computer in a clerkship final examination. Acad Med 2001; 76:835-9. OpenURL CrossRef PubMed Web of Science.