Message from the President

Edward P. Finnerty, Ph.D.
President, IAMSE

As 2007 comes to an end, it seems appropriate to recap some of the highlights of the year for IAMSE.

This past July IAMSE visited Cleveland for another successful meeting. This year the highlights, in addition to a wonderful visit to the science center, included some excellent pre-conference faculty development courses and workshops, presentations on quality teaching, promoting students communication skills, teaching methods and assessments. The various focus sessions further elaborated on the plenary sessions as well as a variety of other timely and interesting topics. The culminating debate, "Is Medicine a Science" lead to some very pointed and timely conversations during the post-debate discussions sessions.

IAMSE was present and participated in the AMEE meeting in Trondheim, Norway this past August. In addition to hosting a luncheon for those interested in IAMSE we sponsored several sessions. From reports they were well received. We intend to participate again in 2008 in Prague. If anyone plans to attend please contact Julie Hewett (julie@iamse.org).

The Webcast Audio seminar series concluded this past November with the final of our two-part series on academic leadership. Tom Schmidt did a masterful job of bringing it all together. Next Spring's series looks to be another valuable offering. Nehad El Sawi has taken over the reigns as Chair of the WAS committee relieving Jack Strandhoy to focus on his role as Secretary of IAMSE. Our special thanks to Jack. Please respond to Nehad's calls for ideas and input so we can continue to offer a quality service to our members.

The membership committee has initiated the first Colleague-to-Colleague series. This is a project that came out of the Strategic Planning process and discussions at the annual meeting. This first session focused on faculty development and how to help faculty become better teachers.

As mentioned above, the annual meeting ended with the debate regarding medicine as a science. Related to that is the role and value of the sciences in medical education. This is a timely project being undertaken by IAMSE and described in the last issue of JIAMSE. To date the coordinating committee has been established and a first round of letter inviting a variety of constituent organizations to participate. Please contact your professional associations, e.g. Society for Neuroscience, American Association of Anatomists, etc. to inquire if they have responded and are participating. Our plan is to begin the study with a focus on the U.S. As we progress we would like to incorporate a more international perspective from our European and Asian members.

The National Board of Medical Examiners (NBME) has unveiled their proposal for restructuring the licensure examination process and replacing USMLE Steps 1 and 2 with a Gateway examination. IAMSE has carefully considered the issue and the proposal and submitted a response to the NBME. A copy of the letter is posted on the IAMSE website (www.iamse.org) for your review. You will note that in our response, we have encouraged the NBME to make use of the wisdom and experience of our membership in their deliberations on this issue. I would encourage all of you to please accept if asked to contribute.

A new initiative for IAMSE is a collaboration with HEAL (Health Education Assets Library; www.healcentral.org). This initiative will involve a major revision of the HEAL site to take advantage of the web 2.0 technology. New features will include a Wiki configuration for greater interaction and sharing of resources and information. Ample opportunities will be available for IAMSE members to contribute. Watch for more details this Spring.

The strength of IAMSE is its membership and for IAMSE to thrive, we need to look to increasing our membership base. IAMSE has welcomed affiliation with related associations that promoted the mission of IAMSE; The Group for Research in Pathology Education (GRIPE) being one example. Our meeting this July in Salt Lake City will be the first incorporating the members of the former Slice of Life. Along these lines, the Team Based Learning Consortium (TBL) will be joining us as well. We look to these groups as becoming productive and active members of IAMSE.

Please mark your calendars for 12th Annual IAMSE meeting, July 26-29, 2008 co-hosted by University of Utah School of Medicine and Slice of Life. Veronica Michaelsen and her committee have been working diligently to develop an exciting, informative and fun meeting. Also, for long-term planning, the 2009 IAMSE meeting will be held in early July in Leiden, Netherlands.

The Medical Educator's Resource Guide

John R. Cotter, Ph.D.

The websites reviewed by the Medical Educator's Resource Guide in this issue - one from the Medical University of South Carolina, another from the University of Iowa College of Medicine and a third from the University of Oklahoma Health Sciences Center –are used by the schools for laboratory instruction in histology.

The digital images of all three websites resemble the histological specimens that would be studied with a microscope. The websites differ substantially, however, in the way the content of the materials is presented and displayed to an online audience. The website at the University of Oklahoma is structured as an atlas and uses static labeled digital images. The other two websites utilize virtual microscopy and are structured to resemble a microscope laboratory. As Robert Ogilvie, Professor Emeritus of Cell Biology and Anatomy and Professor of Pathology and Laboratory Medicine at the Medical University of South Carolina, pointed out in his 2006 IAMSE audio seminar, virtual microscopy uses computer software to simulate the process of viewing glass microscope slides with a microscope. He also stressed that virtual microscopy may or not use virtual slides.*

At Dr. Ogilvie's institution, the images are viewed with WebMic, a virtual microscope that scans and magnifies the snapshots of digitized histological materials. The students are guided in their study of the images by a laboratory manual that explains the relevant aspects of the morphology contained within the images and the steps that are needed to be taken if the images are to be thoroughly examined. In contrast, the website at the University of Iowa utilizes virtual slides. The difference is subtle and, according to Dr. Ogilvie, related to the digitization of the entire specimen contained on a glass microscope slide.

With virtual slides, the users can, if they wish, explore any sector of a slide and the cellular details of a specific area of the slide at higher magnifications. The technological difference is apparent when comparing the magnification features of the two virtual microscopy websites. There is also some perceivable but modest difference in the loading times for the images at higher magnifications when the field of view is moved.

If you know of a website that teachers and students working in the medical sciences would find useful, please consider submitting a review to The Medical Educator's Resource Guide. You can do so by contacting us by e-mail (jrcotter@buffalo.edu).

Interactive Histology Atlas. University of Oklahoma Health Sciences Center

http://w3.ouhsc.edu/histology/

The Interactive Histology Atlas from the University of Oklahoma's College of Medicine is a collection of digital images taken of microscope slides selected from the school's teaching collection. The slides are organized into organ systems and tissue type laboratories, making navigation easy and intuitive.

A total of twenty different laboratories are available for review, each containing digital images from as few as one (blood and hematopoiesis) and as many as nineteen microscope slides. The actual number of images is greater with the major structures contained within the microscope slides being presented separately within the context of the individual slides.

A short explanation that describes the particular function and microanatomy of an organ or tissue introduces some of the topics. The structures contained on the slides are listed and linked to labeled images. The digital images of the light microscopy slides are of high resolution allowing for easy recognition of cellular details illustrated by the images. The images are organized in a way that allows the user to contemplate a structure before an example appears on the computer screen.

Users looking to find a histology site that can be used as a study aid will find this one helpful. (*Reviewed by Philip Brondon, M.S., University at Buffalo*.)

The Virtual Slidebox. Department of Pathology. The University of Iowa.

http://www.path.uiowa.edu/virtualslidebox/

The appeal of virtual microscopy is that digital images can be examined in much the same manner as real specimens would be examined with a microscope. Thus, the accessibility of histological slides is extended beyond the laboratory environment.

At the University of Iowa, the concept has been applied to normal histological materials and pathological specimens. The digital slides can be viewed through a range of magnifications thereby allowing the examination of an entire specimen. Additionally, the brightness can be adjusted and a specific area of the slide selected via a 'toolbox' located within a separate frame. The histology of some of the specimens is described using a format reminiscent of an atlas or laboratory manual. The labeling of the images is minimal and predominantly associated with the 'Comparative Search Tool'. Working with the specimens, therefore, requires some histological knowledge and/or accompanying texts. At the University of Iowa, the virtual slides are used as one part of a package of instruction that also includes labeled digital images and microscope slides. **

The 'Comparative Search Tool' offers the potential of side-by-side slide analysis. This is helpful when learning to distinguish between tissues of similar morphology, normal and pathological tissues or tissues from different species of animals. Additionally, some of the comparative slides are annotated and allow the user to hide or show labels while viewing a slide.

The loading times for the images can sometimes be delayed, but the time is not excessive given the overall quality of the images. The best quality and color was obtained on a computer monitor with a resolution setting of 1280 x 1024, though a resolution of 1024 x 768 still performed well.

Conceptually, this online resource is well organized and easy to navigate through portals on the main page. The website is a good supplement to texts, classroom instruction and/or microscopy. The comparative capabilities of the website are very helpful. The website is best viewed with a monitor resolution of 1280 x 1024 or higher, though lower resolutions do not negate the overall usefulness of the site as an ancillary learning tool. (*Reviewed by René Lisjak, M.A., University at Buffalo.*)

WebMic and Companion Manual of Histology Exercises. Medical University of South Carolina.

http://people.musc.edu/~vslide/webmic/allgspez/WebMicGenOrg.html

The site is a virtual microscope laboratory for histology. It consists of a collection of digital images (WebMic) that is available through the American Association of Medical Schools and MedEdPORTAL*** or by contacting Dr. Robert Ogilvie at ogilvieb@musc.edu. Local installation and use of WebMic is possible by contacting Dr. Ogilvie. The supplemental laboratory manual mentioned in the introduction to the Resource Guide that directs the students in their use of WebMic is available for purchase.

WebMic, which covers all aspects of a course in histology, is remarkable for the way the collection of nearly 160 digitized specimens reproduces the viewing of microscope slides with a microscope and facilitates the identification of structures contained within the digital images. In the microscope mode of the application, two images are displayed on the monitor screen. The smaller of the two images provides an overview of a specimen and is used as a reference to focus on areas that are examined at a greater magnification in the second larger image. The field of view of the second image, which mimics the way specimens are seen through a microscope, can be selected from the smaller image or by dragging the larger one. In the full screen mode, the field of view is not limited. It displays the entire specimen. In either mode, users can view the images without labels or with labels applied to select structures and test their ability to identify structures. In some instances, the images are annotated with information about key structures or terms applied to structures in the digital images. These features make the application an extraordinary learning tool.

The combination of virtual microscopy and user interaction at this site is innovative and a valuable addition to what is offered at most on-line histology websites. (*Reviewed by John R. Cotter, Ph.D., University at Buffalo.*)

*The audio seminar "Implementing Virtual Microscopy in Medical Education" was given by Robert W. Ogilvie, Ph.D., on May 16, 2006 and accessed on November 1, 2007. The talk shall be found on the International Association of Medical Educators website at http://www.iamse.org/development/2006/was_2006_spring.htm.

** Heidger Jr., P.M., Dee, F, Consoer, D., Leaven, T., Duncan, J., Kreiter, C. Integrated approach to teaching and testing in histology with real and virtual imaging. *Anatomical Record (The New Anatomist).* 2002; 269 (2):107-112.

***MedEdPORTAL is located at www.aamc.org/mededportal.

COMMENTARY

Recognising the Role of Medical Scientists in Undergraduate Medical Education: The Experiences of One Medical Faculty

Michelle McLean, Ph.D., M.Ed.,[1] F. Christopher Howarth, Ph.D.[2]

[1]Department of Medical Education
[2]Department of Physiology
United Arab Emirates University
FMHS, UAEU
P.O. Box 17666
Al Ain
United Arab Emirates

Phone: (+)1- 971-3-713-7137 Fax: (+) 1-971-3-767-2167 Email: mcleanm@uaeu.ac.ae

ABSTRACT

From the authors' experiences at a medical faculty where students are introduced to research early in their studies, members of the traditional pre-clinical medical science departments (Anatomy, Physiology, Biochemistry, Pharmacology and Medical Microbiology) have contributed significantly to the development of a research ethos amongst students, not only by supervising curriculum research, but also by providing extra-curricular research opportunities for students locally and abroad. Some of this work has resulted in scholarly achievements such as conference presentations and journal articles. Thus, medical scientists, through their supervision and mentorship of student research, have contributed to the preparation of young medical students for their clinical experience by developing skills that foster life-long learning. These important contributions should be duly acknowledged.

Abraham Flexner's 1910[1] report changed the face of medical education across the globe. Flexner advocated that analytical reasoning and research (in his view, clinical) should be included in a physician's training. Questions arising during patient care should serve as stimuli for inquiry and research to improve both teaching and patient care.[1] Following this famous 1910 report, not only did research became an important part of the life of an academic, but the 'traditional' medical curriculum evolved to reflect the preclinical study of foundational medical sciences of Anatomy, Physiology, Biochemistry, Pathology, Pharmacology and Microbiology, followed by a period of clinical clerkships. Medical scientists were appointed (presumably on their research scholarship) to teach the "science" of medicine, while physicians took the responsibility of training students in the "art" of clinical diagnosis, treatment and management.

Although medical practitioners need to use evidence to inform their clinical decisions[2], it is only relatively recently that the importance of student research opportunities in the medical curriculum has been appreciated. Even today, despite the paradigm shift in medical education to more integrated, student-centred and outcomes-based programmes, research opportunities for medical students vary from negligible[3], through the inclusion of special study modules[4] or perhaps extra-curricular summer research[5], to the other extreme, where a research dissertation is required for graduation.[6-9] The latter is the case for some German medical faculties and a handful of North American institutions that have adopted a scientific model of learning, involving the discovery of new knowledge. Perhaps the high priority given by medical educators and curriculum developers to designing a 'core' curriculum over the past few decades of curriculum reform has resulted in considerably less time being devoted to important issues such as research, ethics and social responsibility.

This situation is, however, changing. The increasing use of evidence-based medical practice in a cyberworld of readily accessible information requires a sound understanding of scientific methodology[2], analytical and communication skills as well as an ability to critically appraise a vast literature.[10-13] Research experience is therefore increasingly being recognized as integral to students' understanding of medicine, and its integration into medical and health science curricula is now becoming widespread.[3,13-18]

Including student research in the medical curriculum has several positive spin-offs. In the first instance, patients should benefit from practitioners who critically use evidence to inform their clinical decisions.[2] Improved practice and health care should enhance the reputation of the medical profession. Furthermore, in a profession plagued by declining numbers of academic physicians and "endangered" disciplines, research experience may influence career pathways, motivating some medical graduates to enter academic medicine or become involved in post-graduate research.[14,16-19] Students too should benefit from their research experience. As they design research projects, generate hypotheses and analyze and interpret their data, they develop generic or transferable skills that prepare them for the challenges of their future careers in clinical medicine.[11-13,20] Research will also hopefully foster a life-long pursuit of knowledge, perhaps prompting post-graduate studies. It has been reported that research experience strengthens students' residency applications, particularly in North America.[3,9]

Recognizing the value of scientific and clinical research on the development of clinical reasoning skills and on the practice of evidence-based medicine, research projects were introduced in 2001 into the second year (Year 2 of Medical Sciences) of the undergraduate medical curriculum at the Faculty of Medicine and Health Sciences (FMHS) at the United Arab Emirates University (UAEU). The annual intake into the Faculty ranges from 29-60 students, while the staff complement is stable at around 90 members. Either individually or in groups of 2-5, students use protected time over a 4-5 month period for research. In addition, in 2003, extra-curricular summer research opportunities locally and abroad for students at any level of study were formally introduced. This extra-curricular research has become increasingly subscribed. The percentage of Year 1 students participating rose from 34.5% in 2003 to 58.5% in 2006. For the same period, participation amongst Year 2 students increased by 31.8% (from 11.8% to 43.6%). From humble beginnings, the Faculty's undergraduate research programme (curricular and extra-curricular) has developed from strength to strength, the merit of which has been recognized by a recent external review committee - *"Involvement of the medical students in research is impressive. Continuation of this programme is highly encouraged"*.

As it had been approximately 5 years since the introduction of mandatory second year research projects and 3 years since summer research was first offered, an audit was called for (Table 1). Data collected from several sources (e.g. lists from the student project co-ordinator; year books; emails to faculty) took into account, amongst other things, types of projects offered and the disciplines of supervisors, as well as the scholarly outcomes (e.g. conference presentations, publications). In this discussion, we will utilize some of these data to support our view that medical scientists, through their supervision of undergraduate student research, can foster a research ethos and promote a scholarship of discovery.

Immediately apparent from Table 1 should be the consistent contribution of medical scientists from the traditional pre-clinical disciplines of Anatomy, Physiology, Biochemistry, Medical Microbiology and Pharmacology to the supervision of both mainstream (51.3% of projects) and extra-curricular (76.5% of projects) research projects. Medical scientists accounted for 60.2% of Year 2 and 81.0% of the summer project supervisors, respectively. Projects have generally been laboratory-based, with many involving animal models to investigate human clinical conditions such as diabetes, cancer and neurological disorders. Apart from the supervisory contributions of members of Community Medicine, who supervise project work in the senior clerkship year, participation by clinical faculty has been erratic. This warrants exploration. A possible explanation may relate to their service responsibilities.

The importance of the medical sciences as the foundational underpinnings of the training and education of physicians cannot be disputed. As Finnerty (2004) pointed out, *"the understanding of "basic sciences" is used everyday when a physician confronts a patient and attempts to generate a diagnosis and treatment plan"*.[21] By experiencing "science in action" during laboratory-based research, students witness first-hand the ramifications of a disruption of normal physiological and biochemical processes in the human body. Biomedical and scientific research thus has the potential to provide authentic and contextually relevant learning experiences for students. It is our view then that medical scientists in our Faculty have been able to incorporate such learning experiences into the undergraduate curriculum through their supervision of student research. According to Nancy Malkiel, a Dean at Princeton, *"the research experience challenges and stretches students in ways that cannot be replicated even in the most rigorous and demanding coursework"*.[22]

As medical scientist faculty members generally do not have clinical responsibilities, they can (as is the case in our faculty) contribute to the scholarship of an institution by promoting a research ethos amongst students. Many of our students have presented their research findings at international conferences (e.g. the Gulf Co-operation Council Medical Students' and the Young European Scientists' conferences), where some have won prizes. A number of students also appear as co-authors on peer-

Table 1. Second year and extra-curricular faculty-based research projects in terms of projects undertaken and supervision. Year 2 projects were introduced in the 2001/2 academic year. Extra-curricular research, formally introduced in 2003, is undertaken in the summer between two academic years. The academic staff complement is ± 90 faculty members.

	Curriculum research: Year 2 projects		**Extra-curricular summer research**		
Academic year (*n* = Year 2 students)	**Project description (*n* = number of projects)**	**% supervisors from different disciplines (*n* = number of supervisors)**	**Year (*n* = Year 1-6 students)**	**Project description (*n* = number of projects)**	**% supervisors from different disciplines (*n* = number of supervisors)**
2001/2 (*n* = 34)	(*n* = 11) Laboratory: 27.3% Community: 27.3% [#]Other: 45.4%	(*n* = 12) 33.4 pre-clinical 58.3 clinical 8.3 medical education	**Not yet introduced**		
2002/3 (*n* = 29)	(*n* = 8) Laboratory: 25.0% Community: 25.0% Other: 50.0%	(*n* = 13) 53.8 pre-clinical 46.2 clinical	**2003** (*n* = 224)	(*n* = 26) Laboratory: 57.7% Clinical: 19.2% Community: 3.9% Other: 19.2%	(*n* = 24) 75.0 pre-clinical 20.8 clinical 4.2 medical education
2003/4 (*n* = 35)	(*n* = 11) Laboratory: 37.5% Community: 62.5%	(*n* = 9) 66.7 pre-clinical 33.3 clinical	**2004** (*n* = 219)	(*n* = 19) Laboratory: 94.7% Other: 5.3%	(*n* = 12) 83.4 pre-clinical 8.3 clinical 8.3 medical education
2004/5 (*n* = 39)	(*n* = 11) Laboratory: 100%	(*n* = 10) 90.0 pre-clinical 10.0 clinical	**2005** (*n* = 234)	(*n* = 27) Laboratory: 74.1% Clinical: 7.4% Other: 18.5%	(*n* = 25) 72.0 pre-clinical 28.0 clinical
2005/6 (*n* = 60)	(*n* = 12) Laboratory: 66.6% Clinical: 16.7% Community: 16.7%	(*n* = 14) 57.1 pre-clinical 42.9 clinical	**2006** (*n* = 242)	(*n* = 21) Laboratory: 81.0% Clinical: 4.7% Other: 14.3%	(*n* = 16) 93.7 pre-clinical 6.3 clinical
Average (*n* = 39)	**(*n* = 10) Laboratory: 51.3% Clinical: 3.3% Community: 26.3% Other: 19.1%**	**(*n* = 12) 60.2 pre-clinical 38.1 clinical 1.7 medical education**	**Average** (*n* = 230)	**(*n* = 23) Laboratory: 76.5% Clinical: 8.0% Community: 1.0% Other: 14.5%**	**(*n* = 19) 81.0 pre-clinical 15.9 clinical 3.1 medical education**

[#]*Other* includes medical education, clinical/laboratory, community/laboratory

reviewed articles. This is the ultimate proof of scholarship.[23] This is a win-win situation for everyone: students' residency applications are strengthened, faculty promotion or tenure is potentially boosted through increased productivity and the prestige of the academy is enhanced.

In addition to promoting scholarship, medical scientists in our Faculty have also, by introducing young students to the rigors of scientific and biomedical research and reasoning, inculcated skills to equip students for their future studies and ultimately, their professional practice. To this end, students must independently (but under supervision) plan, organize, conduct experiments and subject their research findings to public scrutiny. Developing a project proposal, generating an ethics application, presenting it to the Ethics Committee, writing a progress report, creating a poster and finally, preparing a manuscript in the form of a journal article contribute not only to students' communication skills but also to their information-handling, organizational, technical and numeracy skills. These have been identified as core skills for life-long learning.[4] In addition, as students generally work in groups, they also learn to interact with colleagues, thereby potentially contributing to better management of learning in terms of giving and receiving feedback, accepting responsibility and, hopefully, reflecting on their contributions to the group.[4] We have supporting evidence. During the past academic year, two second year students' research project involved canvassing their senior colleagues' perceptions of their research experience. When asked about the contribution of their research (curricular and extra-curricular) experience to the development of their generic skills (e.g. data analysis, information technology skills, time management, etc.), most students indicated a value in excess of 50%. Students from other institutions have also testified to the value of research on skills development, particularly critical thinking and analytical skills.[10,13,20] In terms of the impact of research on their studies and future careers as medical doctors, research allowed one FMHS student to *"understand diseases better"*. He *"also developed a sense of seriousness in my [his] life that made me [him] function better currently and hopefully in my [his] future studies"*. For another student, research helped *"bring up hypotheses for new research topics, especially topics in the clinical years"*, while for a third,*" it helps understanding my [her] career and analyzing results before accepting them"*.

It is our opinion that the medical scientists (and a handful of clinical researchers) at our institution have promoted in young students the ability to critically evaluate evidence, a practice that will inform their future clinical decisions. They will also have fostered skills that will enable students to be better communicators and team members. For some students, this research experience may motivate them to pursue a career in academic medicine, thereby potentially contributing to a new generation of urgently required clinician-researchers.[17,18] As the first students exposed to curricular research are still in their early residency training, only time will tell whether they will become tomorrow's clinical researchers. The research seed has, however, been sown.

Thus, at a time in medical education and health care when most clinicians spend a considerable proportion of their time attending to patients and supervising interns and residents, medical scientists have an important role to play in preparing young medical students for clinical practice through Boyer's[24] scholarships of discovery, application, integration and teaching. In integrated curricula, where early patient contact is advocated, research can contribute to the development of a number of skills that these young students will use in their studies and in clinical practice. Critical analysis, clinical reasoning and communication skills would be amongst these. In addition, providing research opportunities, including international exposure for those who have excelled (as happens at our institution), may potentially "career pipeline" some students into clinical disciplines in dire need of researchers.[17] The valuable role of medical scientists in promoting a research ethos and scholarship in the institution through their supervision of student research should be duly acknowledged and rewarded.

REFERENCES

1. Flexner, A. Medical education in the United States and Canada: A Report to the Carnegie Foundation for the Advancement of Teaching. New York: Carnegie Foundation, 1910.
2. Sackett, D.L., Rosenberg. W.M.C., Gray, J.A.M., Hayes, R.B. and Richardson, W.S. Evidence based medicine: What it is and what it isn't? *British Medical Journal.* 1996; 312: 71-2.
3. Rhyne, R.L. A scholarly research requirement for medical students: the ultimate problem-based learning experience. *Academic Medicine*. 2000; 75: 523-524.
4. Murdoch-Eaton, D.G., Ellershaw, J., Garden, A., Newble, D., Perry, M., Robinson, L., Smith, J., Stark, P., Whittle, S. Student-selected components in the undergraduate medical curriculum: a multi-institutional consensus on purpose. *Medical Teacher.* 2004; 26: 33-8.
5. Schor, N.F., Troen, P., Kanter, S.L., Leving, A.S. The Scholarly Project Initiative: Introducing Scholarship in Medicine through a longitudinal, mentored curriculum programme. *Academic Medicine.* 2005; 80: 824-31.
6. Altunbas, A., Cursiefen, C. Research activities of medical students in Germany using as an example the Wurzburg University Clinic. *Deutschse Mediese Wochenschrift.* 1998; 15: 617-20.
7. Cursiefen, C., Altunbas, A. Contribution of medical student research to the Medline™-indexed publication of a German medical faculty. *Medical Education.* 1998; 32: 439.
8. Dewey, M. Students' evaluation of research during medical studies: Medical dissertation in Germany. *Medical Education.* 2003; 37: 278-80.
9. Chongsiriwatana, K.M., Phelan, S.T., Skipper, B.J., Rhyne, R.L., Rayburn, W.F. Required research by medical students and their choice of a women's health care residency. *American Journal of Obstetrics and Gynaecology.* 2005; 192: 1478-80.
10. Frishman, W.H. Student research projects and theses: should they be a requirement for medical school graduation? *Heart Disease.* 2001; 3: 140-4.
11. Whittle, S., Murdoch-Eaton, D.G. Development of lifelong learning and self-evaluation skills through special study modules. *Medical Education*. 2001; 35: 1066-90.
12. Whittle, S.R., Murdoch-Eaton, D.G. Student-selected projects: Can they enhance lifelong learning? *Medical Teacher.* 2002; 24: 41-4.
13. Joubert, G. Research by pre-clinical undergraduate medical students. *Medical Education.* 2006; 40: 459-89.
14. Friedrich, M.J. A novel programme seeks to take clinical scientists off the endangered species list. *Journal of the American Medical Association*. 2003; 290: 1019-20.
15. Zier, K., Stagnaro-Green, A. A multifaceted programme to encourage medical students' research. *Academic Medicine.* 2001; 76: 743-7.
16. Solomon, S.S., Tom, S.C., Pichert, J., Wasserman, D., Powers, A.C. Impact of medical student research in the development of physician-scientists. *Journal of Investigative Medicine*. 2002; 51; 149-56.
17. Halpain, M.C, Jeste, D.V., Trinidad, G.I, Wetherell, J.L., Lebowitz, B.D. Intensive short-term research training for undergraduate, graduate, and medical students: early experience with a new national-level approach to geriatric mental health. *Academic Psychiatry.* 2005; 29: 1. http://ap.psychiatryonline.org. Accessed 7/22/2006.
18. Wagner, R.F., Ioffe, B. Medical student dermatology research in the United States. *Dermatology Online Journal*. 2005; 11: 8.

http://dermatology.cdlib.org/111.commentary/msresearch/wagner.html, Accessed 7/22/2006.

19. Segal, S., Lloyd, T., Houts, P.S., Stillman, P.L., Jungas, R.L., Greer, R.B. The association between students' research involvement in medical school and their postgraduate medical activities. *Academic Medicine.* 1990; 65: 530-3.
20. Jacobs, C.D., Cross, P.C. The value of medical student research: The experience at Stanford University School of Medicine. *Medical Education.* 1995; 29: 342-6.
21. Finnerty, E.P. Why the basic sciences? *Journal of the International Association of Medical Science Educators.* 2004; 12(1). http://www.iamse.org/artman/publish/article_170.shtml. Accessed 8/24/2007.
22. Malkiel, N.W. (unknown). Spotlight: undergraduate research, scholarship, and creative activity. http://www.reinventioncenter.miami.edu/Research%20spotlight.html. Accessed 9/3/2007.
23. McLean, M., Howarth, F.C. Does undergraduate student research constitute scholarship? Drawing on the experiences of one medical faculty. *The Journal of Scholarship of Teaching & Learning.* In press.
24. Boyer EL. *Scholarship Reconsidered: Priorities of the Professoriate*. San Francisco: Jossey-Bass, 1990.

JIAMSE

Journal of the International Association of Medical Science Educators

Volume 17 Number 2 2007

IAMSE

Role of Medical Scientist in UGME

Classroom versus Online Learning

Pre-course Online Quizzes

Audience Response System

Pop Quizzes- Evidence Based Strategy

Pre-entry Program- Effect on Performance

IAMSE on the Web
www.iamse.org

ISSN: 1550-8897

JIAMSE

The Journal of the International Association of Medical Science Educators

Volume 17 **Number 2** **2007**

COMMENTARY

Classroom Versus Online Learning: Experience with Pathology Remedial Course

Sharon L. Wenger, Ph.D.[1], Scott Cottrell, Ed.D.[2], Barbara S. Ducatman, M.D.[1]

[1]Department of Pathology
[2]Department of Community Medicine and the Office of Medical Education
P. O. Box 9203
West Virginia University
Morgantown, WV 26506-9203

Phone: (+)1- 304-293-0446 Fax: (+) 1-293-3976 Email: swenger@hsc.wvu.edu

ABSTRACT

A remedial course for pathology was created and administered during the past four summers; two years as a classroom lecture course, and two years as an online course. Currently, West Virginia University School of Medicine is the only U.S. medical school that offers a summer remedial experience for a second-year medical school pathology course.

Objective: The objective of this study is to ascertain any differences in student learning between the online and the lecture-based remedial courses.

Plan of study: At the end of the summer courses, we asked students to evaluate the course and provide input about the course regarding what was done well and what could be improved. Weekly and shelf examination scores were also compared between students in the lecture-based and online remedial courses.

Results: Students in both the lecture-based and online remedial courses participated to the same extent in the learning experiences ($p<0.9$). Weekly student examination scores between the lecture-based and the online course were similar ($p<0.09$). However, the National Board of Medical Education (NBME) pathology shelf board scores were significantly lower for the online course ($p<0.001$). Students' comments on how to improve the learning experiences addressed similar themes for both the lecture-based and the online courses. Students noted, however, that a unique benefit of the online course was the ability to view lectures at their own pace.

Conclusion: Overall, students had more positive comments about the online learning experience, which could be offered at a lower cost due to reduced faculty time. However, future modifications of the online learning experience will be necessary in order to maximize student performance on the standardized pathology shelf examination.

INTRODUCTION

A summer remedial course was developed by pathology faculty at West Virginia University. For the first two years, this course was designed as a didactic, lecture-based course. As with most academic departments, clinical responsibilities have increased, subsequently limiting time that the faculty has for teaching, particularly during the summer months. In order to ameliorate the demand for faculty time, lectures were recorded from the traditional second-year medical school pathology course. These lecture recordings made it possible to create a remedial course online. Although several lecturers had to rerecord lectures initially, overall, it greatly reduced faculty time, and we could therefore reduce tuition costs. We compared students' experiences in the classroom and the online course by examining differences between three sources of information: course evaluation results, students' performance on weekly examinations, and students' performance on pathology shelf examination.

MATERIALS AND METHODS

Classroom and online courses were set up in a similar manner. Both courses were designed to cover one chapter per day using Robbins (Kumar et al., 2005), a standard pathology textbook, for a four hour lecture. In the classroom, faculty members were available for students' questions the day of the lecture. PowerPoint presentations were printed and handed out to students. For the online course, PowerPoints could be printed, and recorded PowerPoints with voice over were available, the latter were compressed using software for online usage (Impatica Inc, Ottawa, Ontario, Canada, and Camtasia, TechSmith Corporation, Okemos, MI). In addition, daily quizzes for each chapter were available for self study, and students could post questions to the faculty on the website. Weekly examinations were given in both the lecture-based and the online remedial courses. In class, all questions were reviewed with the students after the examination; online, the questions and answers were made available to the students for 15 minutes, followed by an hour in which they could ask questions of the faculty in a "live" chat room. In order to pass the course, students had to pass the seven weekly examinations at 75% or better, as well as pass the NBME pathology shelf examination at the 12th percentile or higher. The scoring is based on standardization of national performance of medical students taking the examination, ranked in percentiles. For the 12th percentile rank, 12 percent of all medical students scored at or lower than this score.

Prior to taking the final shelf examination, students were asked to provide feedback regarding the course. The questions asked were: 1) What percentage of lectures did you attend/view?, 2) Did you feel the course helped you learn the material?, 3) What did you like about the course?, and 4) What suggestions would you make to help improve the course? Anonymous questionnaire responses were compiled regarding the classroom and online experiences. A Student t-test was used to compare the average of weekly examination scores, shelf scores and answers to the questions regarding lecture attendance or viewing, and adequacy of course for learning the material.

RESULTS

Most students in both the classroom and the online remedial courses reported attending or viewing greater than 75% of lectures ($p<0.9$). Almost all of the students felt that the course helped them learn the material regardless of the means of presentation ($p<0.9$). Average scores for the weekly examinations were similar for both the classroom and online courses ($p<0.9$). However, shelf examination scores for the online summer course (average 32.6 percentile) were significantly lower than for the classroom remedial courses (average 54.2 percentile) ($p<0.001$). Five of the 32 students who took the course online failed the shelf examination, while all who attended the classroom setting had passed (Table 1). Over all 4 years, 14 of 15 remediation students from our institution had scored an average of less than the 7th percentile (2nd -11th %ile), while their repeat remediation shelf scores averaged 50th percentile (22nd-77th percentile), similar to overall student performance during the normal semester length course. While eleven of these 15 students passed the examination portion of the semester course, four students received a failing score for the course examinations; one who was on academic probation, and two who had barely passed several other courses.

Table 1. Frequency of Pathology Shelf Scores for Students in the Lecture and the Online Remedial Courses

Shelf Board Percentile	Lecture Course n	Online Course n
<12	0	5
12 to 30	1	12
31 to 50	7	9
51 to 70	6	6
>70	3	0

$t=3.987$, $df=49$, $p<0.001$

Students commented positively about the course in four separate areas: faculty, lecture material, exams and overall (Table 2). They noted some advantages for the online course. The daily online quizzes were helpful and the structuring of the course, in particular not having to relocate for eight weeks, allowed them to work at their own pace.

Student comments for suggestions to improve the remediation course differed for the lecture and online courses (Table 3). While there were issues with faculty teaching in the classroom, there were no such comments regarding online recorded lectures. Students in the classroom felt that more emphasis should have been placed on clinical aspects, and more time spent on systems pathology, while students online had more issues related to presentation of material on the website, such as lack of available downloadable audio files separate from the voiced over PowerPoint (which was corrected for the second year online). Students were also more critical about weekly examinations in the classroom setting than they were online. Overall, students felt the organization could be better regardless of the venue.

Table 2. Student Comments for the Question: What did you like about the course?

Theme	Lecture Course	Online Course
Faculty	Individual attention	Chatroom discussion, student questions and answers from faculty posted online
Lecture Material	Stressed important material	Audio PowerPoints, followed textbook
Exams	Review of weekly exam, board-style questions	Quizzes, weekly exams, review of weekly exams, quality of questions
Overall	Small class, informal, accommodating staff	Excellent flow of course, well organized, online, own pace, did not have to relocate

DISCUSSION

Some of the advantages of the online course for the students were the reduction in tuition cost, not having to relocate for two months during the summer, and the flexibility of looking at material at their own pace. An advantage for the faculty was reduction of teaching time, which minimized time away from their clinical activities. A major disadvantage was that students had to be self-motivated to keep up with their studies. The overall decline in the shelf scores supported the fact that this may not be the best learning option for all students.

In order to provide a good online course, Wong and colleagues recommended ten areas of activity that must be addressed (Wong et al., 2003), which included need, course material, staffing, and active learning features. Since there have been more inquiries about the remedial course over the past four years than enrollment, there is demand for this service. Remediation allows students to continue their medical school education with delay of only a few months, rather than an entire year. During our normal teaching of the pathology course, we have placed material on a website developed for medical student use for all our medical school courses, where students can access material for the course, including PowerPoint slides and recorded PowerPoint lectures for student review. Since this material was available, it lent itself to the online venue. For online material, some lectures were not included, or lectures were rerecorded to condense the material covered during the school year. Limiting the number of recorded lecture hours to approximately four per day seemed to work well, allowing students time to go through the material at least once, and review textbook material before taking the quiz.

Table 3. Student Comments for the Question: What suggestions would you make to improve the course?

Theme	Lecture Course	Online Course
Lecture Material	More emphasis on clinical aspects, condense lectures, more time on organ based material	Handouts of PowerPoints, both PowerPoint and audiotaped lecture available, post in advance, audio for MP3
Exams	Proofread questions, more exam reviews	Proofread questions, take shelf exam closer to home
Overall	Better organization and communication, reduce cost	Better organization, Reasonable number of lecture hours per day, conference sessions before exam

Other online courses have been developed to be more interactive as far as the learning process (Velan et al., 2002a; 2002b). One study that evaluated paper and online post-graduate study in professional programs reported that while students enjoyed online interactive exercises, they preferred paper text material (Burgess et al., 2005). We had a similar experience with our students, who wanted the plain PowerPoints to print out in order to take notes while listening to the recorded lectures. In a study comparing interactive versus non-interactive web-based modules, medical students and residents showed significantly better learning for the former, which was also rated higher (Kerfoot et al., 2006).

Chao (2003) evaluated interactive functions for learners that are necessary for content. These included links to related educational sites, learning materials and multimedia presentations. Patel and colleagues (2006) described the development of a histology atlas for their medical school course, which highlights areas of interest using a JavaScript rollover function. We found the disk accompanying the Robbins textbook (Kumar et al., 2005) includes case studies for about two-thirds of the chapters. These case studies present patient history and laboratory results, followed by questions and answers about gross and histological specimens with a rollover function to highlight specific areas of interest. We will include review of these case studies for our online course to enhance the student learning experience, as well as help prepare them for the shelf board examination. We also intend to make faculty available to the learners in a chat room prior to the weekly examinations, to allow for more student-instructor interaction, another area of interaction important for web-based learning (Chou, 2003).

Certainly, for any course, improvements should be made each year. These changes we have made have been based on student input, with the intent of improving student learning. Even though our experience with the online course was overall positive, our main goal is to help students learn the material, and evolving into a more interactive course should help those students who require more self discipline.

REFERENCES

1. Burgess, G., Holt, A. and Agius, R. Preference of distance learning methods among post-graduate occupational physicians and hygienists. *Occupational Medicine* (London). 2005;55:312-318.
2. Chou, C. Interactivity and interactive functions in web-based learning systems: a technical framework for designers. *British Journal of Educational Technolog*y. 2003;34:265-279.
3. Kerfoot, B.P., Conlin, P.R. and McMahon, G.T., 2006. Comparison of delivery modes for online medical education. *Medical Educati*on. 2006;40:1137-1138.
4. Kumar, V., Abbas, A. K. and Fausto, N. (Eds). Robbins and Cotran pathologic basis of disease, 7th ed. *Philadelphia*, *Elsevier Saunders*. 2005.
5. Patel, S. G., Rosenbaum, B. P., Chark, D. W. and Lambert, H. W. Design and implementation of a web-based, database-driven histology atlas: technology at work. *The Anatomical Record*. 1006;289B:176-183.
6. Velan, G.M., Killen, M.T., Dziegielewski, M. and Kumar, R.K. Development and evaluation of a computer-assisted learning module on glomerulonephritis for medical students. *Medical Teacher*. 2002a;24:412-416.
7. Velan, G.M., Kumar, R.K., Dziegielewski, M., and Wakefield, D. Web-based self-assessments in pathology with Questionmark Perception. *Pathology*. 2002b;34:282-284.
8. Wong, G., Greenhalgh, T., Russell, J., Boynton, P. and Toon, P. Putting your course on the web: lessons from a case study and systematic literature review. *Medical Education*. 2003;37:1020-1023.

SHORT COMMUNICATION

Pre-Course Online Quizzes: a Facile Way to Get Students up to Speed at the Start of Your Course

Frazier T. Stevenson, M.D.,[1] Uldis N. Streips, Ph.D.[2]

[1] Department of Medicine
University of California Davis
School of Medicine
Sacramento, CA 95817
[2] Department of Microbiology and Immunology
School of Medicine
University of Louisville
Louisville, KY 40292

Phone: (+)1- 502-852-5365 Fax: (+)1-502-852-7531 Email: unstre01@gwise.louisville.edu

ABSTRACT

Course directors often note that medical science classes struggle with beginning medical courses when the material introduces novel concepts and terminology or requires recall of previously learned material. Pre-class quizzes induce students to review past courses and pre-read before the class is initiated, fostering easier entry into the subject matter.

Students in professional schools are often challenged to apply previously learned, as well as some novel material, in required professional courses. Based on their understanding of the objectives of prerequisite courses, instructors often assume that students have understood prior concepts. However, this assumption is often incorrect. The earlier courses may not have taught what was assumed, or students may have learned material superficially for a summative exam, but never internalized it for later application. Frequently, there is a long time lag between the prerequisite course and the new one. These mistaken assumptions may continue within a course—the instructor assumes learning of initial material prior to moving on with more advanced content, but students may not have done so. Students may not pre-read the required texts and may then struggle as the course quickly moves on into even more complex ideas and material. These challenges were met at two professional schools by using online pre-quizzes at the start of required courses in Microbiology/Immunology (University of Louisville School of Medicine, U of L) and Nephrology (University of California Davis School of Medicine, UCD). Online quizzes have been used in several curricula to stimulate review or provide formative feedback[1-10]. Those described pre-quizzes differed from those in this manuscript because ours are given at the beginning of major courses.

At UCD a 30 question online quiz was distributed 3 days before the start of the year 2 Nephrology course. The quiz covered normal renal histology and physiology that had been taught one year earlier in required year 1 courses and was designed to reflect the key concepts needed for success in understanding renal pathophysiology. Students had 2 weeks to complete it, taking it at most once per day and keeping their best score. WebCT software was used for administration. Questions were scrambled on each administration to prevent students generating a standard key. The quiz was open book, but students were asked to complete it independently. Questions were a mix of factual recall and short vignettes requiring understanding or application of knowledge (sample questions are shown in Table 1 with correct answers bolded).

The quiz counted 13% of the final course grade, more than normally used for a formative quiz. However, the multiple

Table 2. Student evaluation of UCD Nephrology prequiz (two

Table 1. Nephrology prequiz sample questions.

1. Subjects A and B are 70-kg males. Subject A drinks 2 L of distilled water and Subject B drinks 2 L of isotonic NaCl (140 meq/L in Na^+). As a result, subject B will have
   a. greater change in ICF volume
   b. greater change in plasma osmolarity
   c. higher positive free water clearance
   d. higher urine osmolarity
   e. lower ECF volume

Which of the below molecules has the lowest permeability through the glomerular basement membrane (GBM)?
   a. large anionic molecule
   b. large neutral molecule
   c. large cationic molecule
   d. small anionic molecule
   e. small neutral molecule
   f. small cationic molecule

2. In renal compensations for respiratory acidosis the kidney
   a. increases bicarbonate reabsorption
   b. decreases ammoniagenesis
   d. compensates by loweringthe blood pH
   e. compensates mostly in the first 5 minutes
3. Which of the below would cause the greatest increase in the glomerular filtration rate (GFR)?
   a. Afferent arteriolar constriction
   b. Efferent arteriolar constriction
   c. Renal artery constriction
   d. Renal vein constriction

4. Compound X reduces the freezing point of water but does not cause water to shift out of cells when added to the extracellular fluid. Which one is true about Compound X?
   a. It exhibits neither osmolality nor tonicity
   b. It exhibits both osmolality and tonicity
   c. It exhibits osmolality but not tonicity
   d. It exhibits tonicity but not osmolality

attempts allowed and the score provided at the end of each attempt allowed the student to use it formatively to stimulate additional learning. Specifically, over the two years the pre-quiz has been given, students took the quiz a mean of 3.9 times (range 1-11, n = 192), with a mean final correct score of 94.3% (n = 192). Assessment of the value of the exercise by students has been positive (Table 2), both regarding its value in reviewing the one year distant course material, and in its value in understanding current Nephrology concepts.

At U of L the 10 question electronic quiz (sample questions shown in Table 3; correct answers bolded) covering the first two chapters of the required text was sent to the class the week before class started in Microbiology and Immunology. Since immunology is the first course subject and is complex in both nomenclature and concepts, it was communicated to the students that the answers could be readily found in the first two chapters of the required book and that the quiz represented 0.5% of their total grade in the course. We also indicated that the quiz is open book but must be done independently. The native curiosity of medical students to do

Question	mean	n
The physiology online prequiz focused my review of normal physiology.	6.06	144
The physiology online prequiz helped me to master later course content.	5.82	143

Modified 1-7 Likert type scale, 1 = strongly disagree, 7 = strongly agree

Table 3. Medical Microbiology sample questions

1. An antigen is:
    a. always a protein
    b. reactive with a single class or isotype or antibody
    c. unable to be endocytosed
    d. composed of epitopes that bind specifically to antibodies
    e. a molecule too small to react immunologically
2. Dendritic clles are known to:
    a. activate T cells in lymph nodes
    b. be of erythroid progenitor lineage
    c. destroy pathogens
    d. be a kind of granulocytes
3. The antibody molecule has variable and constant regions. The constant region:
    a. is identical in amino acid sequence among all classes of antibodies
    b. has antigen binding capacity
    c. is present only on heavy chains
    d. can bind to eukaryotic cell surface
    e. is coded by two separate genes
4. B cells:
    a. secrete but don't bind antibody
    b. can mature to antibody secreting plasma cells
    c. recognize only pathogen derived peptides in context of MHC Class I
    d. rearrange heavy chains only in response to antigen presence
    e. are activated by macrophages to produce antibody
5. The Major Histocompatibility Complex:
    a. is responsible for transplant rejection
    b. contains genes for antibody synthesis
    c. is involved in immune responses to microorganisms replicating inside mammalian cells (i.e viruses and certain microbes)
    d. is not on lymphocyte surfaces
6. The innate response is characterized by:
    a. specific B and T lymphocytes
    b. inflammation and complement activity
    c. absence of microbicidal activity
    d. immunological memory
    e. antigen specificity

questions correctly, the possibility of getting some positive credit before the course started, and the fact that they would ultimately get it all correct yielded 100% response from the class of 146 two years in a row. An analysis of both years showed that 60% did it correctly the first time, while it took 23% 2 tries, 12 % 3 tries, and 5% more than that to get all the answers correct. No formal survey of the class has been taken regarding the value of the quiz, but anecdotal evidence suggests that the class enjoys this exercise and appreciates the credit points. Instructors report far fewer elementary questions and apparent confusion in the introductory lectures. Similar instructor feedback was obtained anecdotally from UCD Nephrology discussion leaders, who reported increased student precision in vocabulary and in application of physiology principles during the initial discussion sections on electrolyte disorders.

In summary, self-paced pre-quizzes given during the initial phase of these courses stimulated learner review of past material, focused them on concepts essential to understanding the upcoming course, and provided formative feedback of their understanding and retention of past curricular content. They were also useful for course instructors to assess the variability of retention from prerequisite courses. While the quizzes at the two institutions differed in length (30 questions UCD vs. 10 U of L), in weight towards the final grade (13% vs. 0.5%) and in how many times the quiz could be taken (12 vs variable) student and faculty acceptance was similarly positive. At UCD the pulmonary and neurology pathophysiology courses are now using similar pre-quizzes based on the success of the exercise in nephrology. While online quizzes are now widely endorsed for assessment during the progress of conventional and electronic curricula [2-7, 9, 10], we additionally endorse such quizzes for motivating students to succeed early in courses and to integrate course content with previously learned material.

References

1. Narloch, R., Garbin, C.P., and Turnage, K.D. Benefits of Prelecture Quizzes. *Teaching of Psychology.* 2006;33 (2):109-112.
2. Schwartz R.W., Donnelly, M.B., Nash, P.P., and Young, B. Developing students' cognitive skills in a problem-based surgery clerkship. *Academic Medicine.* 1992;67:694-696.
3. Hammoud, M.M., and Barclay, M.L. Development of a Web-based question database for students' self-assessment. *Academic Medicine.* 2002;77:925.
4. Andersen, J.D. An Electronic Leap: Quizzes on the Web. *Teaching Professor.* 2004;18:4.
5. Schwartz, R.W., Donnelly, M.B., Sloan, D.A., and Young, B. Knowledge gain in a problem-based surgery clerkship. *Academic Medicine.* 1994;69:148-151.
6. Kibble, J. Online quizzes are a valid and acceptable formative assessment tool. *FASEB Journal.* 2007;21:479.
7. Tse, M.M., Pun, S.P., and Chan, M.F. Pedagogy for teaching and learning cooperatively on the Web: a Web-based pharmacology course. *Cyberpsychological Behaviour.* 2007;10:32-37.
8. Byrnes, M., and Byrnes, J.F. Quizzes Are the Right Answer. *Teaching Professor.* 2007;21:3.
9. Kibble, J. Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance. *Advances In Physiological Education.* 2007;31:253-260.
10. Bernardo, V, Ramos, M.P., Plapler, H., De Figueiredo, L.F., Nader, H.B., Ancao, M.S., Von Dietrich, C.P., et al. Web-based learning in undergraduate medical education: development and assessment of an online course on experimental surgery. *International Journal of Medical Information* 2004;73:731-742.

Developing Case-based Assessments to Integrate Basic and Clinical Sciences and Improve Perceived Relevance

Hettie Till, M.Sc., M.Med.Ed., D.Ed.,
Rahim Karim, B.Sc., D.C., FCCRS(C)

Undergraduate Education
Canadian Memorial Chiropractic College
6100 Leslie Street
Toronto, Ontario
Canada M2H 3J1

Phone: (+)1-416-482-2340 Fax: (+)1-416-488-0470 Email: htill@cmcc.ca

ABSTRACT

Students find it difficult to see the relevance of, and have trouble integrating, some of the basic science course material when taught in isolation from the clinical courses. Integration of disparate material is perceived to be a challenge and this problem is compounded by assessments in which a single discipline is tested at a time – mostly in a multiple choice format. As such, the theory assessments in Years I and II were changed into a case-based, integrated format with extended-matching and short-answer type questions. This article describes the development of case-based theory exams involving simultaneous testing of all material taught in a module. Specific steps involving the process of building common cases across disciplines, question construction, faculty coordination and quality assurance are described.

This paper also addresses the quality of these integrated exams as experienced over the first academic year of implementation. Statistical analyses of internal consistency, such as Cronbach's alpha of between 0.75 and 0.89, and split-half values of between 0.69 and 0.93 for the different exams were found. Correlations between modular results, as well as between Year I discipline-based results and Year II modular results, indicated some measure of construct validity. This result was supported by backward stepwise multiple linear regression analyses.

The new assessment format is resource intensive, but addresses issues of clinical relevance and integration. The fairly sound psychometric proportions of the exams support the high-stake decisions that we have to make on the basis of our examination results.

INTRODUCTION

Traditional undergraduate health education exemplifies non-integration where quite often every hour, on the hour, a different lecturer walks into the classroom and teaches his or her own subject as though it is the most important subject in the program. These lecturers cover widely different content material and they have different teaching styles and varying expectations of the learners. How does a student take all this information, bring it together, and combine the knowledge and skills, particularly from the basic sciences, into the reality of professional practice?[1-4] This challenge to the student becomes even more serious when single courses are assessed separately, mostly in a multiple-choice question (MCQ) format of varying question writing depth and style,[5-7] scattered throughout the year and often competing with each other[7,8] , though at some schools Block Testing has obviated this particular problem.[7]

A new curriculum at the Canadian Memorial Chiropractic College (CMCC) introduced in 1999 was specifically designed to improve horizontal and vertical integration by arranging the first three years of a four-year health professional curriculum according to anatomical region and by the early introduction of clinical sciences. However, formal[9,10] as well as informal student feedback indicated that the taught curriculum was not as well integrated as anticipated. In order to get an idea of how well our curriculum was integrated, we evaluated it against the integration ladder developed by Harden (2000)[3]. This integration ladder, which is a tool that can be used for curriculum planning and evaluation, has 11 possible steps from the lowest level of integration, called Isolation, through Awareness, Harmonization, Nesting, Temporal Coordination, Sharing, Correlation, Complementary, Multidisciplinary, Interdisciplinary, to Transdisciplinary, the highest level of integration. This review indicated that we were still mostly operating at Step 1 or "Isolation" - a situation where individual lecturers organize their own courses, determining the content as well as the depth to which the course will be covered and developing their assessments in isolation without knowing what others are teaching.

The students also indicated that they had difficulty in seeing the relevance of much of the taught material and they expressed dissatisfaction with assessment methods that tested mostly recall and short term memory through the commonly used multiple choice question (MCQ) assessment format. Assessment-related stress caused by a large number of separate course examinations was also a concern.[11]

In order to improve the process of integration, instruction in Year III was changed at the beginning of the 2002-2003 academic year into a modular system with nine modules each culminating in a week of theory and practical assessments. This modular system has now been in operation for four academic years. At the beginning of the 2005-2006 academic year, a modular system was also introduced in Years I and II with each year being divided into four modules. This model is similar to that described by Streips et al[7] in that courses are taught as before during the module, but, instead of discipline-based assessments scattered throughout most of the year, each module is followed by two weeks of both theory and practical assessments.

Assessment strongly affects student learning, and clinical cases can help students associate course material with "real" patient situations and thus improve relevance and retention.[2] Therefore we decided to change the modular theory assessments into a case-based format using extended-matching (EMQs) and short-answer (S-A) questions instead of MCQs.[12,13] It was anticipated that these assessments would also lead to better sequencing and integration of course material when faculty collaborate on the development of cases and trans-disciplinary questions. This article describes the development and implementation of integrated modular theory exams for Years I and II of the program.

.

MATERIALS AND METHODS

The modular system for Years I and II included two weeks of assessment at the end of each of four modules (Figure 1), except for the last module where the assessment period was three weeks long to accommodate an Objective Structured Clinical Examination (OSCE). In order to give the students the opportunity to have a complete break from their studies over the winter and March break periods, the first two modules were shorter than the last two modules of the year (Figure 1). Although the assessment period at the end of each module also included practical examinations as well as Objective Structured Practical Examinations (OSPEs), this paper describes only the three theory examinations written at the end of each module.

Each theory exam included questions from all courses. If a module, for example, was nine weeks long, Exam 1 covered predominantly the work taught during weeks one to three, Exam 2 predominantly the work taught during weeks four to six, and Exam 3 predominantly the work taught during weeks seven to nine. The number of exam questions per course was prorated to the instruction time, usually two to three questions per lecture hour. The questions could be either in S-A or EMQ format. Each exam had a total score of at least 100. The results of the three exams were averaged for a final theory percentage per module.

Development of the theory exams

The group responsible for developing the exam questions consisted of all faculty teaching in a particular module. Shortly after the start of each module, a meeting was held with all relevant teaching faculty and/or course coordinators. The learning outcomes of all courses taught during the module were made available to the attendees and at the start of the meeting, each attendee briefly described the content to be covered in his/her course during the module. This helped the individual faculty members better understand what was being taught overall. At the conclusion of this part of the meeting, the clinical conditions to be covered in the four cases for each exam (twelve in total) were decided. All the cases had to be relevant to practice.

Completed cases were circulated to all examiners to edit, review and modify so that they could link their own questions to them. For example, if nutrition questions on carbohydrates were to be asked, the faculty member might add to the case that the patient was a single parent who often bought fast food, or was a diabetic. Each faculty member had to spread his/her questions across all four cases. They could also add extensions to a case such as additional physical exam findings, a return visit sometime in the future, or a question asked by a patient. Clinical relevance had to be emphasized and some of the basic science faculty sought assistance from clinicians in this process. This process lead to the development of a modified form of EMQs where the four cases per exam served as the themes with varying numbers of questions linked to them (See Appendix 1). In order to minimize the opportunity for guessing the correct answer, all option lists had to contain seven to twenty-six options.[15]

Once all the questions were submitted, they were reviewed for submission accuracy and then strategically sequenced after each of the four cases to create a "real life" patient situation. This meant that questions relating to history and

Figure 1. Outline of the Year I modular system

MODULE 1		MODULE 2			MODULE 3			MODULE 4	
Lectures and Labs	2 weeks exams	Lectures and Labs	2 weeks exams	WINTER BREAK	Lectures and Labs	2 weeks exams	MARCH BREAK	Lectures and Labs	3 weeks exams OSCE

basic sciences were placed before the physical examination questions, followed by diagnosis, treatment and/or other management questions. At this stage, potential problems such as not sending a pregnant patient for X-rays, were addressed. The final step in this process was peer review by at least three faculty members.

Writing the tests

The "traditional" examination procedures were not changed. Students were randomly assigned to seats and three hours were allowed to complete each exam.

Quality assurance of results

Analyses of item difficulty and point biserials of all the extended-matching questions in each exam were conducted. Poorly discriminating questions were flagged and referred to the relevant faculty members for review and comments. After receiving responses from the faculty members, corrections were made which might either be the addition of an extra correct option worth a full or part mark, or the deletion of a few poorly discriminating questions. On the final number of questions in each exam, difficulty levels, Spearman-Brown, Kuder-Richardson's 20 and 21, Standard errors, as well as Cronbach's alpha were calculated by means of the Performance Evaluation Technology (PET Version 4.1399, TDA Inc., Burlington, ON, Canada). All other statistical analyses were performed by means of the Statistical Package for the Social Sciences (SPSS) Version 14 programme.

Overall pass/fail performances for each exam were determined. After calculating averages to obtain one final theory percentage from the three exams, descriptive statistics of the results were calculated. An absolute standard was applied to the results in that all students who scored below 60% failed. The percentages obtained for each individual discipline in each exam was also calculated and published in separate category reports to assist the students in their self-directed learning and to flag poorly performing students in each discipline for the relevant lecturers. Although not used for promotion purposes, the individual scores per discipline were also tracked throughout the year.

Calculations of Pearson product moment correlation coefficient, r, was completed in order to describe the magnitude of relations between the modular theory results per year of study. For the Year II students we also had the results from their Year I individual course assessments available and correlation coefficients between their Year I final course results (independent variables) and their Year II Modular theory results (dependent variables) were calculated. From these same results a backward stepwise multiple linear regression was performed for the four outcome (dependent) variables: the average theory results obtained in Year II Modules 5, 6, 7, and 8. The eight input (independent) variables included in the final analysis were the final Year I results in Anatomy, Body Mechanics, Diagnosis/ Orthopaedics, Histology, Neuroanatomy, Radiology and Radiation Physics, as well as the Year I GPA.

RESULTS

Integration

Apart from promoting cooperation among faculty members,[4] the assessment development process had many other positive outcomes, such as helping us move upwards on Harden's Integration Ladder.[3] The format of the exam planning meetings not only helped the individual faculty members better understand overall educational outcomes, but each person could also review whether his/her course material was optimally placed in a specific module.

Incorrectly placed material could then be moved before the start of the next academic year and unnecessary duplication of the material removed.[4] This portion of the planning meeting moved us from Step I, or "Isolation" to Step 2, or "Awareness" and on to Step 3, or "Harmonization" on the integration ladder.[3]

During the previous academic year, the exam planning meetings also stimulated lecturers to talk about prerequisites for their courses. It became apparent that "Nesting" (Step 4) was required and lecturers started working together to ensure an enriched learning experience for the students by talking about skills relating to other courses. For example, although the inflammatory response was normally only taught in Year II, a basic lecture on this topic, taught by a basic scientist, is now included as part of the Functional Recovery and Active Therapeutics I course. For the second academic year of using this system, some Histology lectures and a lab were moved from Anatomy in Year I to Immunology in Year II even though they are still taught by the Histology lecturer. Changes such as these are the beginning of "Sharing" or Step 6 on the integration ladder.[3]

Relevance

Another positive outcome of the adapted form of EMQs from four common clinical cases per exam, was the fact that twelve cases per module, together with red and yellow flags, covered many of the major clinical disorders likely to be encountered in the anatomical areas covered by a module of five to nine weeks. This exam format also showed the students how the preclinical courses, in particular the basic sciences, fit into professional practice.[2,4,6]

Feasibility

The theory examinations proceeded smoothly and the students soon became familiar with the EMQ and S-A question formats. They indicated that the three hours allocated per exam were more than enough. The mean for every module in the two years was over 70% and the pass percentage was consistently over 90% (Table 1).

Reliability

Depending on the number of short answer questions included in each exam, and after elimination of poorly discriminating questions, the number of EMQs per exam varied between 70 and 100 (Tables 2 and 3). The internal consistency (Cronbach's alpha) was between 0.75 and 0.89, and both split-half (Spearman-Brown) and Kuder-Richardson's 20 values were between 0.69 and 0.93 for the exams written. (Tables 2 and 3).

Validity

To investigate construct validity for the modular theory examinations in Year I, the final results were correlated with each other. The Pearson correlations between the results of Modules 1, 2, 3 and 4 were all significant and between 0.75 and 0.87. Correlations with other assessments in Year I was also calculated and gave statistically significant correlations between 0.63 and 0.78 with the final Biological Sciences OSPE, between 0.40 and 0.52 with the final Radiology OSPE, and between 0.44 and 0.50 with the end of year OSCE.

Similar results were obtained with Year II results (Table 4) with Pearson correlations between the results of Modules 5, 6, 7, and 8 of between 0.76 and 0.79. For the Year II students we also had their Year I results available which gave significant correlations with the modular results in Year II (Table 4). The Year I GPA correlated between 0.82 and 0.86 with the different modules with the correlation increasing over the year from 0.82 for Module 5 to 0.86 for Module 8. Correlations between the Year II modular theory results and the Year I Technique practical assessment results were only between 0.12 and 0.25 and not significant at the 0.01 level.

The backward stepwise multiple linear regressions all reached convergence in seven iterations or less. After six iterations, 69% of the variance in Module 5 ($R^2 = 0.687$) was explained by the independent variables Year I GPA, Radiation Physics and Radiology (in order of importance). After six iterations, 69% of the variance in Module 6 ($R^2 = 0.691$) was explained by Year I GPA, Diagnosis and Orthopaedics, and Body Mechanics. After four iterations, 71% of the variance in Module 7 ($R^2 = 0.711$) was explained by Year I GPA, Radiation Physics, Radiology, and Anatomy. After seven iterations, 74% of the variance in Module 8 ($R^2 = 0.737$) was explained by Year I GPA and Radiology.

.

DISCUSSION

This report describes the development and introduction of integrated modular theory examinations into an undergraduate health professional curriculum. Similar to the block testing introduced by two medical schools in the USA and described in a recent paper,[7] our modular system was designed to progress from frequent individual discipline assessments, by grouping assessments into a dedicated period at the end of each module. This was to allow the students an uninterrupted period of five to nine weeks to concentrate on their learning, "digest" the content, and move towards an understanding of the relevance of and integration between the different courses taught during the module.

The assessment development process

The process followed to develop these integrated theory exams was found to be resource intensive from both faculty as well as support perspectives. The value of

Table 1. Descriptive statistics of modular results

	YEAR I				YEAR II			
	Mod 1	**Mod 2**	**Mod 3**	**Mod 4**	**Mod 5**	**Mod 6**	**Mod 7**	**Mod 8**
N	193	186	186	185	178	170	175	177
< 60%	13 (7%)	12 (6%)	18 (10%)	16 (9%)	16 (9%)	15 (9%)	10 (6%)	9 (5%)
> 80%	55 (28%)	38 (20%)	40 (22%)	32 (17%)	52 (29%)	34 (20%)	52 (30%)	29 (16%)
% passes	93%	93%	90%	91%	91%	91%	94%	95%
Highest %	92%	93%	96%	95%	92%	91%	92%	89%
Lowest %	48%	43%	36%	44%	39%	45%	50%	47%

changing to case-based questions is well supported by the literature[2,4,6,8] and evidence exists that EMQs have many advantages over MCQs. EMQs are viewed as a more "fair" format, with better item discrimination, [12] and with the ability to measure degrees of expertise.[13] However, it was anticipated that faculty reaction to this change might not be positive.[4, 6, 7, 11, 14]

Modular assessments are not new to our institution as a modular system had been in place for Year III since the beginning of the 2002-2003 academic year. In the Year III modular theory exams each examiner developed his/her own cases and questions in the EMQ or S-A format and assessment integration was mostly of the "stapled" variety.[2] The implementation of the Year III modular system went smoothly with faculty members quickly adapting to the new assessment format, even though some concerns regarding increased workload was expressed to the Dean.

The recent assessment changes in Years I and II, however, affected many more faculty members, particularly from the basic sciences. Some of the faculty members were initially not comfortable with the new format and most expressed concern at the increased workload caused by having to develop new questions.[4,6,7,11,14] As the academic year progressed, some faculty members appeared to become more positive which supports the experience at other institutions implementing change.[7] The majority of the faculty adapted and wrote creative questions, perhaps realizing that they were starting to develop a new bank of questions which would decrease their workload over the next few years. This result is similar to that reported in the literature [11] indicating that exam setting time was reduced in subsequent years.

Exposure to the case-based questions very soon lead to a request from the students that all courses should now be taught in a case-based format and for some of the faculty members this request added significantly to their stress levels. To assist faculty members in this task, faculty development workshops on case-based teaching have since been offered.

As Harden (2000)[3] describes, moving up the integration ladder requires more central organization and resources and more "communication and joint planning between teachers from different subjects." In this regard we have found our exam planning meetings very useful. Over the year, as faculty members began to understand the process better, the meetings were getting shorter and more work was done by e-mail. However, one meeting at the beginning of each module was an absolute necessity. It is important to note that these meetings should be driven by

Table 2. Year I - Statistics of the extended-matching question portion of the modular theory exams

	No. of Students	No. of EMQs	Difficulty	Spearman-Brown	Kuder-Richs's 20	Standard Error	Cronbach's alpha
Module 1							
Paper 1	193	91	.72	0.85	0.84	4.21	0.84
Paper 2	192	71	.71	0.80	0.79	3.77	0.79
Paper 3	193	85	.70	0.85	0.83	4.14	0.83
Module 2							
Paper 1	189	87	.75	0.80	0.78	4.0	0.78
Paper 2	189	85	.67	0.79	0.81	4.26	0.81
Paper 3	189	100	.73	0.83	0.84	4.33	0.84
Module 3							
Paper 1	188	85	.70	0.86	0.85	4.09	0.85
Paper 2	188	84	.70	0.86	0.85	4.11	0.85
Paper 3	189	86	.72	0.89	0.89	3.99	0.89
Module 4							
Paper 1	185	93	.70	0.74	0.76	4.35	0.76
Paper 2	186	97	.70	0.80	0.84	4.43	0.85
Paper 3	185	71	.72	0.87	0.83	3.70	0.83

someone from the Dean's Office, in our case the relevant Education Coordinator. From our experiences, ownership of the process is crucial to success. Adequate administrative assistance was also a necessity, as putting together these integrated exams is time-consuming and requires a good understanding of the process. Skilled administrative assistance saved the reviewers many hours of work.

Reliability

The quality of the integrated assessments

In order to allow wide sampling of content across the discipline material taught in a module, and to reduce the error caused by task variability to a minimum[8] we had decided on three theory exams of at least 100 marks each to be written at the end of every module, thus a total of 300 marks or more. We also decided to use only the average of the three exams for promotion purposes, rather than the total of the individual discipline marks tracked from assessment to assessment over the year.[4,7] This was in an attempt to average out unreliability problems that might occur in one or more areas of the assessments and which might be compounded when tracking individual discipline results over the year. When a final discipline result is made up from small numbers of questions at a time, unreliability could become a serious problem.[8] The fairly good measures of internal consistency found in our individual modular exam papers (Tables 2 and 3) supported us in this decision. Adding together the results of three exam papers (235 - 272 EMQs plus S-A questions), with a testing time of nine hours[8] and using the average of the three papers, should also improve reliability and support the making of high-stakes decisions on the basis of these results.

Validity

Validity is not easy to show. However, the process followed in developing the questions should help with a measure of content validity as it ensured that the questions used in a module came from all courses taught in that module.[15] Although processes are currently underway to have the submitted questions checked against the learning outcomes[16] to ensure a good spread, so far this has not yet been concluded. However, the fact that assessments took place at five to nine week intervals and that at least 300 questions had to be developed at the end of each module, could be indications that most work taught during the module would be assessed.

Bridge et al[16] described four primary principles on which the development of a content-valid test should be based. These include review by experts in the field, as well as the writing of high quality test items. Content validity of our exams would therefore also have been positively affected by the fact that we implemented peer review of the exams, as well as by workshops designed to help faculty develop their question writing skills.

The correlations found between the results that both Year I and Year II students achieved in the four modules per year, as well as between Year II modular results and the final marks that the same students obtained in their Year I exams the year before, are similar to others reported in the

Table 3. Year II - Statistics of the extended-matching question portion of the modular theory exams

	No. of Students	**No. of EMQs**	**Difficulty**	**Spearman-Brown**	**Kuder-Richs's 20**	**Standard Error**	**Cronbach's alpha**
Module 5							
Paper 1	183	79	.71	0.88	0.87	3.89	0.87
Paper 2	184	86	.75	0.85	0.87	3.91	0.87
Paper 3	184	70	.70	0.86	0.84	3.73	0.84
Module 6							
Paper 1	182	85	.74	0.81	0.80	3.97	0.81
Paper 2	182	76	.68	0.85	0.84	3.95	0.84
Paper 3	183	88	.72	0.84	0.83	4.14	0.83
Module 7							
Paper 1	181	87	.73	0.75	0.74	4.06	0.79
Paper 2	181	75	.72	0.79	0.71	3.81	0.79
Paper 3	171	80	.77	0.77	0.69	3.68	0.75
Module 8							
Paper 1	181	80	.75	0.74	0.75	3.84	0.75
Paper 2	181	98	.24	0.81	0.93	4.44	0.81
Paper 3	181	90	.75	0.78	0.75	4.04	0.75

Table 4. Bivariate correlations of students' Year I final course results (independent variables) and their Year II Modular theory results (dependent variables)

Year I final course results	**Year II final modular theory results**			
	Module 5	**Module 6**	**Module 7**	**Module 8**
Body Mechanics	.64	.67	.62	.65
Technique	.12	.25	.18	.14
Anatomy	.53	.55	.52	.58
Histology	.67	.62	.56	.67
Neuroanatomy	.67	.63	.65	.66
Diagnosis & Orthopaedics	.70	.72	.64	.71
Professionalism & Ethics	.47	.55	.44	.42
Health Promotion	.30	.36	.34	.31
Biochemistry	.22	.22	.27	.28
Research and Biometrics	.38	.39	.46	.35
Radiology	.48	.58	.50	.49
Radiation Physics	.70	.58	.55	.63
Year I GPA	.82	.83	.83	.86
Module 5	1.00	.78	.76	.79
Module 6	.78	1.00	.79	.79
Module 7	.76	.79	1.00	.78
Module 8	.79	.79	.78	1.00

All correlations are significant at the 0.01 level (2-tailed), except for Technique

literature, and indicated some measure of construct validity.[15]

Of interest was the fact that, although still significant, the correlations between the practical scores in Year I and the Year II modular theory results of the same students were much weaker (0.12-0.25) than those shown between Year I theory results and Year II modular theory results. As different traits were measured in the practical and theory assessments, this result could also be an indication of validity in the new theory assessment method.[8]

Many published results are limited by relying only on univariate or bivariate statistical procedures. So, in order to further investigate the separate effects of the Year I results directly upon the criterion variables in terms of their regression coefficients, a multiple regression equation was developed for the average modular theory result in each of the four modules in Year II. The results obtained from these regression equations also indicated some measure of construct validity in the new format of assessment.

Limitations

Our new assessment format has limitations. Although most faculty members involved had previously attended workshops on EMQ writing, those workshops were held some time before the initiation of test preparation. To assist faculty members in improving their question writing skills two workshops were offered at the start of the new academic year. Faculty members in courses such as Professionalism, Ethics, and Health Promotion, decided not to mark S-A questions for more than 180 students and therefore tried to write EMQs on work which could not easily be linked to patients and they found it difficult to develop seven or more options. This result is similar to that reported by other authors.[15] We will be working with the faculty members involved to try and develop more creative ways to assess their courses.

CONCLUSIONS

According to van der Vleuten (1996)[8] "perfect assessment is an illusion". However, we are cautiously optimistic that our new assessment format is starting to address the perennial issues of relevance and integration, particularly of the basic sciences. For the first time this year, feedback from Year I and II students has shown an improvement in understanding the relevance of material taught in the basic sciences. The fairly sound psychometric proportions of the exams described in this paper support the high-stake decisions that we have to make on the basis of our examination results. The improvements in integration that came about through the assessment planning meetings were an added positive outcome.

REFERENCES

1. Anderson, M.B., Swanson, A.G. Educating medical students – the ACME-TRI report with supplements. *Academic Medicine.* 1993; 68(Suppl.):S1-46.
2. Schmidt, H. Integrating the teaching of basic sciences, clinical sciences, and biopsychosocial issues. *Academic Medicine*.1998; 73(9):S24-S31.
3. Harden, R.M. The integration ladder: a tool for curriculum planning and evaluation. *Medical Education*. 2000; 34:551-557.
4. Ber, R. The CIP (comprehensive integrative puzzle) assessment method. *Medical Teacher.* 2003; 25(2):171-176.
5. Hudson, J.N., Vernon-Roberts, J.M. Assesment – putting it all together. *Medical Education.* 2000; 34:947-958.
6. Hudson, J.N., Tonkin, A.L. Evaluating the impact of moving from discipline-based to integrated assessment. *Medical Education.* 2004; 38:832-843.
7. Streips, U.N., Virella, G., Greenberg, R.B., Blue, A., Marvin, F.M., Coleman, M.T., Carter, M.B. Analysis on the effects of block testing in the medical preclinical curriculum. *Journal of the International Association of Medical Science Educators.* 2006; 16:10-18.
8. Van der Vleuten, C.P.M. The assessment of professional competence: developments, research and practical implications. *Advances in Health Science Education.* 1996; 1:41-67.
9. Till, H. Identifying the perceived weaknesses of a new curriculum by means of the Dundee Ready Education Environment Measure (DREEM) Inventory. *Medical Teacher.* 2004; 26(1): 39-45.
10. Till, H. Climate Studies: Can students' perceptions of the ideal educational environment be of use for institutional planning and resource utilization? *Medical Teacher.* 2005; 27(4):332-337.
11. Tonkin, A.L., Hudson J.N. Integration of written assessment is possible in a hybrid curriculum. *Medical Education.* 2001; 35:1072-1073.
12. Case, S.M., Swanson, D.B., Ripkey, D.R. Multiple-choice question strategies: Comparison of items in five-option and extended-matching formats for assessment of diagnostic skills. *Academic Medicine.* 1994; 69(10):S1-S3
13. Beullens, J., Struyf, E., Van Damme, B. Do extended-matching multiple-choice questions measure clinical reasoning? *Medical Education.* 2005; 39:410-417.
14. Van der Vleuten, C.P.M., Schuwirth, L.W.T., Muitjens, A.M.M., Thoben, A.J.N.M., Cohen-Schotanus, J., Van Boven, C.P.A. Cross institutional collaboration in assessment: a case on progress testing. *Medical Teacher.* 2004; 26(8): 719-725.
15. Beullens, J., Van Damme, B., Jaspaert, H., Janssen, P. Are extended-matching multiple-choice items appropriate for a final test in medical education? *Medical Teacher.* 2002; 24(4):390-395.
16. Bridge, P.D., Musial, J., Frank, R., Roe, T., Sawilowsky, S. Measurement practices: methods for developing content-valid student examinations. *Medical Teacher.* 2003; 25(4):414-421

Appendix 1

Case and sample questions

Jennifer, a 45-year-old woman, complained of "pins and needles" and a dull ache in her right shoulder, radiating down the postero-lateral surfaces of arm, forearm and hand. The pain started 2 months ago and had progressively worsened, affecting her ability to work full time and often awakened her during the night. She also indicated that she lost about 12 pounds since the pain started. Her family history revealed that her grandmother suffered from breast cancer.

Examination revealed that the patient's shoulder had lost its rounded contour and the entire upper limb was adducted. The forearm was more pronated and flexed than usual. The entire upper limb assumed a configuration analogous to "waiter's tip" position. You then listened to her chest, and examined her thyroid gland. On examination of both breasts, a large mass was found in the upper lateral quadrant of the right breast fixed to the surrounding connective and muscle tissues. The right nipple was seen to be higher than the left and inverted. A small dimple of skin was noted over the lump. Palpation of the axilla revealed enlarged, firm lymph nodes. The left breast was unremarkable. A lateral radiograph of the cervical spine showed metastases in the bodies of the 5th, 6th, 7th cervical vertebrae. A blood analysis revealed moderate anaemia. Biopsy of the lump tissue was recommended.

A diagnosis of carcinoma of the right breast with metastases into the cervical spine was made.

1. You evaluate Jennifer's right shoulder and on x-ray note that the humeral head is positioned more superior than what would be expected. This could cause impaction of the head of the humerus onto what ligament? (List 2)
2. Which muscle is the first to be invaded by the patient's tumorous mass? (List 5)
3. Which nerve is compromised if the patient's shoulder lost its rounded contour? (List 6)
4. Invasion of the basement membrane is the first step in the metastasis of this cancer from the breast to the cervical spine. Cancer cells must express a specific protein receptor in order to begin this process. Which protein receptor must the cancer cells express in order to penetrate the basement membrane? (List 10)
5. What is the key characteristic of cancer cells that differentiates their growth pattern from that of normal cells? (List 9)
6. Jennifer wants to know: Which of the currently used cancer immunotherapies is the safest and most successful? (List 16)
7. What allows tumour cells to escape the attack of specific cytotoxic T lymphocytes? (List 16).

(List # refers to the relevant Option List)

Pop Quizzes: Evidence Based Strategy for Medical Students?

Sara G. Tariq, MD[1], Allison M. Johnson, MD,[1] Carol R. Thrush, EdD[2], Majka B. Woods, PhD[2], Virginia A. Johnson, EdD[2], James A. Clardy MD[3]

[1]Department of Internal Medicine
[2]Office of Educational Development
University of Arkansas for Medical Sciences
4301 W. Markham Street
Little Rock, AR 72205 USA

Phone: (+)1-501-526-5968 Fax: (+)1-501-603-1228 Email: tariqsarag@uams.edu

ABSTRACT

Pop quizzes have been utilized in non-medical academic settings to ensure student preparedness and improve test scores. This study is the first to examine the utility of pop quizzes in an undergraduate medical education (UME) setting. Three consecutive years of sophomore medical student data (n = 409) were used to compare performance indicators for students who were administered pop quizzes versus those who were administered scheduled quizzes and those who took no additional quizzes. Indicators examined were course test scores, semester exam scores, cumulative grades and scores on the National Board of Medical Examiner's (NBME) Introduction to Clinical Diagnosis subject exam. Students who took pop quizzes performed the poorest on semester exams, the NBME exam and course final grades. ANOVA and post hoc test analyses showed that all three groups differed significantly on NBME scores and semester exams, with those who completed no quizzes scoring the highest, followed by those who took scheduled quizzes and then those who took pop quizzes. This study suggests that mandatory pop quizzes have questionable value in a UME clinical skills course.

INTRODUCTION

Educators in almost every academic environment are faced with the common problem of students who come to classes unprepared. In medical education, instructors also encounter challenges in presenting multiple and complex academic subjects that require students to learn a voluminous amount of material in a relatively short period of time. Pop quizzes have been utilized by educators in non-medical academic settings to ensure student preparedness[1], promote critical thinking skills[2] and improve subsequent test scores and grades[3]. A small number of previous research studies in non-medical educational settings have examined the effectiveness of using pop quizzes. Results have demonstrated that pop quizzes improve students' preparation for exams, their subsequent test scores, and final course grades[3]. In addition, some studies have found that pop quizzes may be perceived as useful and even liked by some students[3-6]. However, no published studies in undergraduate medical education settings have examined how pop quizzes affect medical student performance.

Anderson[3] has studied the frequency of quiz administration in a behavioral science course and its effect on medical student study behavior, but did not examine pop quizzes. Anderson used a prospective experimental design within a single course and had a small sample of students (n = 10) record their study behaviors under various quiz versus no quiz administration sequences. This study found that students studied more during the weeks in which they were being quizzed regularly, but that "cramming" behaviors occurred in both conditions and there was no relationship between quizzing and performance on exams. Streips et al[7] noted that frequent administration of quizzes/exams to medical students may hinder the retention of information by not providing ample time for integration of information learned, resulting in a "study and forget" cycle.

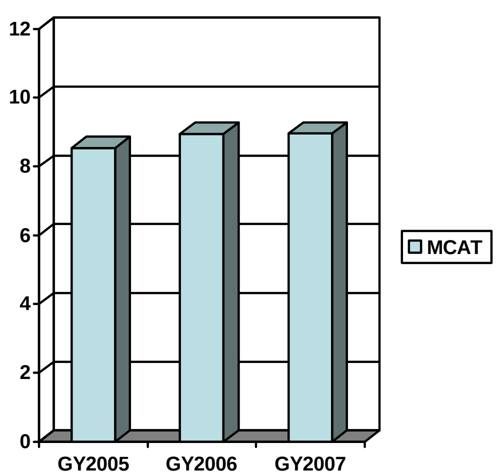
Are pop quizzes effective in undergraduate medical education (UME) settings? In an effort to address this question, the current study explored whether pop quizzes were associated with medical student performance. Specific objectives of this study were to examine whether the quiz format (pop quizzes, scheduled quizzes, or no quizzes) used in an Introduction to Clinical Medicine-II (ICM-II) course for sophomore medical students was associated with differences in: a) ICM-II test scores, semester exam scores, and cumulative course grades, and b) scores on the National Board of Medical Examiners' (NBME) Introduction to Clinical Diagnosis subject examination.

MATERIALS AND METHODS

Participants, Design and Instruments. This study compared three cohorts (2002-2004) of sophomore medical student course data (*n* = 409) collected in an ICM-II course at a medical school in the mid-south United States. The ICM II course is a year-long course that runs parallel to the basic science curriculum; the focus is systems-based and utilizes clinical vignettes for teaching and testing. This study was approved by the institutional review board at the medical school where the study was conducted.

The academic performance of the three cohorts was fairly similar as indicated by average scores on the MCAT. (See Figure 1).

Figure 1. Student MCAT Scores



As shown in Table 1, three consecutive years of student data were used to compare scores on ICM-II course tests, semester exams, cumulative grades and scores on the NBME Introduction to Clinical Diagnosis subject examination. The first cohort of students (*n* = 139) were not administered any additional pop or scheduled quizzes. The second cohort of students (*n* = 132) completed five pop quizzes (10 points each) and the third cohort (*n* = 138) completed scheduled quizzes (10 points each) instead of pop quizzes. All students completed four scheduled 50 point tests, fall semester and cumulative spring exams and the NBME subject examination. There was a total of 700 points possible at the end of the year for all three cohorts. During this three year time period, the structure of the pop quizzes, quizzes and exams did not change, nor did the Course Director. The content of material taught in the course did not change significantly. The only difference in the exam scores was in the first cohort, whose spring semester exam and NBME scores were worth 175 points, instead of 150 points for the second and third cohort. This change was made in response to student feedback that the NBME carried too much weight on their final scores. The weight of the NBME was reduced from 25% of the final grade to 21% of the final grade.

The pop quiz and scheduled quiz questions were identical across both years. The format of the questions were the same; all of the questions were balanced in terms of recall versus application of information. The majority of the quiz questions were presented in clinical vignette format, similar to the NBME style of question-writing. This was designed with the distinct purpose of preparing the students as much as possible for these types of questions. Below is an example question taken from the pop quizzes:

A 45 yr old woman presents to your clinic with a chief complaint of left sided chest pain and shortness of breath that began one day prior. She states that the pain is constant and worse with deep breath. She just returned from a trip to Singapore 2 days ago. She denies fever, chills, nausea, vomiting, or cough. She is a smoker (2 ppd for 25 years), and has had hypertension for 10 years. The pain is not worse with exertion and does not radiate. There is no associated nausea or dizziness. On physical exam, her vital signs show a hr 110, rr 20, and bp 154/99 in both arms. She is afebrile. You note clear lungs, and a 2/6 holosystolic murmur at the left mid clavicular line that radiates to the axilla that the patient states is old. There is no S3, jvd, or peripheral edema.

1) What is the most likely cause of her chest pain?
   a. Coronary artery disease
   b. Pulmonary embolism **
   c. Aortic stenosis
   d. Congestive heart failure

Analyses. Descriptive statistics including means and standard deviations for the ICM-II quiz totals (pop and scheduled), test score totals, and semester exam totals were computed. Semester exam raw scores were converted to z-scores due to a slight variance in the total points possible over the timeframe of analysis for this study (ZExam). Also, due to changes in the way NBME reported scores during the timeframe of analysis for this study, the NBME scores were transformed to standard z-scores for comparative analyses (ZNBME) following methods previously described[8]. A one-way analysis of variance (ANOVA) with Tukey post hoc tests was used to evaluate and compare mean test scores, overall course grades, ZExam scores, and ZNBME scores between years. A one-

Table 1. Student Cohorts and Measures Observed

Year 1 (2002-2003)	Year 2 (2003-2004)	Year 3 (2004-2005)
No quizzes	5 Pop Quizzes	5 Reg. Scheduled Quizzes
4 Prescheduled Tests	4 Prescheduled Tests	4 Prescheduled Tests
Semester Exams (Fall/Spring)	Semester Exams (Fall/Spring)	Semester Exams (Fall/Spring)
NBME Clinical Diagnosis Exam	NBME Clinical Diagnosis Exam	NBME Clinical Diagnosis Exam
Final Course Grade	Final Course Grade	Final Course Grade

way ANOVA with no post hoc tests was used to compare the quiz scores between the cohorts who received either scheduled or pop quizzes.

RESULTS

Descriptive statistics for quizzes, exams and test score totals across the three cohorts are presented in Table 2. One way ANOVA F-tests indicated the presence of significant differences for the quiz totals ($F_{1,269}$ = 14.87, p < 0.01), course test totals ($F_{2,406}$ = 4.82, p < 0.01), ZExam scores ($F_{2,406}$ = 130.17, p < 0.01), overall course grades ($F_{2,406}$ =8.51, p < 0.01) and ZNBME scores ($F_{2,405}$ =344.65, p < 0.01). Tukey post hoc tests (with alpha set at p < 0.05) indicated that test scores were significantly higher for the groups who had scheduled or pop quizzes compared to those who had no pop or scheduled quizzes. Both groups who took quizzes (pop or scheduled) scored significantly higher on the ICM-II test scores than students who had no quizzes. Significant differences were found between all three groups on course Exam scores (ZExams) and NBME scores (ZNBME), with students who completed no quizzes scoring the highest on NBME and ICM-II Exams, followed by those who took scheduled quizzes and then those who took pop quizzes. The students who took the pop quizzes had significantly lower cumulative final course grades than those who took scheduled quizzes, but there was no significant difference in final course grades between the pop quiz cohort and those who did not take quizzes.

Table 2. Descriptive Statistics for Course Quizzes, Tests, Exams, Final Grades, and NBME Scores by Cohort

	No Quiz Cohort			Pop Quiz Cohort			Scheduled Quiz Cohort			Cohort Differences, p < 0.05†
Measures	***n***	***Mean***	***(SD)***	***n***	***Mean***	***(SD)***	***n***	***Mean***	***(SD)***	
Quiz Scores	-	-	-	132	43.90	(4.20)	138	45.20	(2.20)	scheduled quiz > pop quiz*
Prescheduled Test Scores	139	178.80	(9.80)	132	181.80	(9.70)	138	182.10	(9.30)	pop quiz, scheduled quiz > no quiz†
Z Semester Exams (Total)	139	0.85	(0.83)	132	-0.59	(0.75)	138	-0.29	(0.76)	no quiz > scheduled quiz > pop quiz†
ZNBME Clinical Diagnosis Exam Scores	139	1.09	(0.67)	132	-0.73	(0.60)	138	-0.40	(0.56)	no quiz > scheduled quiz > pop quiz†
Final Course Grades	139	88.40	(5.00)	132	88.10	(4.70)	138	90.20	(3.80)	scheduled quiz > pop quiz†

† Tukey post hoc tests
* Oneway ANOVA

DISCUSSION

This study examined whether pop quizzes in a UME clinical skills course were associated with improved performance. We did not find evidence to indicate that pop quizzes were useful for medical students in an ICM-II course. Of the 3 cohorts studied, students who took pop quizzes performed consistently worse in all areas examined except for performance on ICM-II course tests.

ICM-II test scores were highest for the group who were administered scheduled quizzes. Cumulative course grades were also highest for the group who completed scheduled quizzes while NBME scores were highest among the group who had no quizzes.

To our knowledge, this is the only published study that has examined the utility of pop quizzes in an undergraduate medical student population. In reviewing the empirical research on pop quizzes, we found that some studies demonstrated pop quizzes improve subsequent test scores[9]. However, the current study found that pop quizzes did not result in significantly improved final course grades. Graham's study found that students in higher education settings favored the use of pop quizzes because it motivated them to study[9]. However, anecdotal observations by the professor who taught the ICM-II course indicated that students experienced increased stress and frustration on the days that pop quizzes were administered. Not only did this seem to hinder interactive class discussions, but it also may have shifted the perceived locus of control for learning away from the goal-oriented learners. Medical students have traditionally been viewed as highly self-motivated learners and the implementation of pop quizzes may have altered their normal study habits and learning styles.

Previous research has shown pop quizzes are useful when they are offered for extra credit[2] or in short answer format[9]. Our study used a traditional multiple choice format. We discovered that this format was not conducive to positive learning outcomes. Ruscio[10] has shown that pop quizzes which utilize short answer or mini-essay questions result in more student involvement and deeper discussion. This observation may be related to the fact that such formats are more likely to promote critical analysis and reflection than a multiple choice quiz format. Unfortunately, time and class size constraints may prohibit the use of short answer or essay question quizzes in the UME environment.

Although strengths of this study include the use of a large sample size and the examination of different formats of quizzes (pop vs. scheduled vs. none) across several cohorts, there are limitations. This was a retrospective study and is thus limited to a description of observations. We do not suggest explanatory or causal effects between the measured variables. The study was completed at a single institution for a single course, thus limiting the ability to generalize the findings. While we did not measure and control for potential differences in academic performance of individual students, average baseline MCAT scores indicate that the cohorts were not fundamentally different in terms of their test-taking abilities. The three years of medical school classes examined in our study were essentially chosen from the same demographic pool and with the same standards each year of this study.

Promoting a supportive learning environment with effective teaching and assessment methods is vital in successful undergraduate medical education where classes are large, learning is fast-paced, and there is a great deal of material to be covered. In this study, overall course grades were highest for the group who completed scheduled quizzes and lowest for those who completed pop quizzes which suggests that scheduled quizzes, rather than pop quizzes, may help students keep up with their course work, resulting in better final grades. NBME scores were highest among the group who had no quizzes and lowest for those who completed pop quizzes. For medical educators, this study highlights the importance of considering how intuitive interventions in the classroom setting (e.g. pop quizzes) could actually result in counter-intuitive outcomes. Educators who continually assess their educational interventions are well-positioned to provide evidence for the use of various strategies that promote optimal learning. Further critical analysis of traditional evaluation methods and their application in UME will ultimately provide data to help educators make informed course design decisions and ultimately better prepare medical students.

Acknowledgment: The authors would like to thank Karen Szauter, MD, for providing the inspiration for this study, Jennifer VanEcko, ICM-II course coordinator for her administrative expertise, and Elizabeth Hicks for research support.

REFERENCES

1. Kauffman, J. Pop quizzes: Theme and variations. *The Teaching Professor.* 1999; 13:4.
2. Carter, C., Gentry, J.A. Use of pop quizzes as an innovative strategy to promote critical thinking in nursing students. *Nurse Educator.* 2000; 25:155.
3. Anderson, J.E. Frequency of quizzes in a behavioral science course: An attempt to increase medical student study behavior. *Teaching of Psychology.* 1984; 11:34.
4. Kouyoumdjian, H. Influence of unannounced quizzes and cumulative exam on attendance and study behavior. *Teaching of Psychology.* 2004; 31:110-111.
5. Lay, W.E. Pop tests: Valuable instructional tools. *NISOD Innovation Abstracts.* 2002; 24:1.
6. Thorne, B.M. Extra credit exercise: A painless pop quiz. *Teaching of Psychology*. 2000; 27:204-205.
7. Durning, S.J., Pangaro, L.N., Sweet, J., Wong, R.Y., Sealey, M.L., Nardino, R., Alper, E., Hogan, K., Hemmer, P.A. Clerkship sharing on inpatient internal medicine rotations: an emerging clerkship model. *Teaching and Learning in Medicine*. 2005;17(1):49-55.
8. Streips, U.N., Virella ,G., Greenberg, R.B., Blue, A., Marvin, F.M., Coleman, M.T., Carter, MB. Analysis on the effects of block testing in the medical

preclinical curriculum. *Journal of the International Association of Medical Science Educators*. 2006;16(1):10-18.

9. Graham, RB. Unannounced quizzes raise test scores selectively for mid-range students. *Teaching of Psychology*. 1999; 26*:*271-273.

10. Ruscio, R. Administering quizzes at random to increase students' reading. *Teaching of Psychology*. 2001; 28:204-206.

Clickers: the Use of Audience Response Questions to Enliven Lectures and Stimulate Teamwork

Frazier Stevenson, M.D.

University of California
Davis School of Medicine
Medical Education Building (#33), Suite 3101
Sacramento, CA 96817 USA

Phone: (+)1-916-734-2769 Fax: (+)1-916-734-2769 Email: frazier.stevenson@ucdmc.ucdavis.edu

ABSTRACT

Health science educators are under increasing pressure to reduce traditional lecture time and build more interactive teaching into curricula. While small group exercises such as problem based learning achieve that aim, they are highly faculty intensive and difficult to sustain for many faculties. The commercial availability of easy to use audience response systems (ARS) provides a platform for increasing instructor interaction and engagement with learners. This article details my recent experience with ARS, and suggests its uses to increase lecture interactivity, build student teamwork, provide formative feedback, and energize both faculty and students.

INTRODUCTION

Recent medical education trends have emphasized the importance of increasing active learning for health science students. This trend has been driven by education literature emphasizing active learning, application, and analysis, rather than just memorization of facts, and by accreditation bodies[1]. Most education innovations have focused on adding new interactive techniques to curricula, such as problem-based or team-based learning, or use of standardized patients and simulations in small group exercises. Less attention has been given to how the traditional lecture might be enlivened and made more interactive.

For the past two years I have used an audience response system (ARS) in my core lectures in a second year required course in renal pathophysiology. My use of it was based on extensive literature, mostly from the undergraduate curriculum, touting ARS as a useful and stimulating addition to traditional teaching[2]. Among the advantages cited by these and other authors, I was most intrigued by these possibilities:

1. Formative assessment that assess students' understanding of my lecture material
2. Stimulating students to apply and analyze, not just memorize
3. Posing questions that demonstrate students' gaps in knowledge and set up subsequent lecture material
4. Providing a template for interactive discussion between students and between students and the instructor
5. Providing guidance for the instructor to see if topics are understood, or require additional time in the lecture
6. to make lecture fun

In this paper, I report on both my impressions and experience using ARS, and provide student feedback on the experiment.

MATERIALS AND METHODS

The ARS system used is the Interwrite PRS System, version 4.4 (Scottsdale, AZ). The system was used in a second year renal pathophysiology course in 2007 during a series of lectures on fluid and electrolyte disorders. Seven hours of lectures were given, and 20 ARS questions were asked during the lectures. Attendance at the sessions ranged from 35-60 students. Questions were delivered in one of two formats. In the first, a multiple choice single best answer or multiple best answer question was shown, and students were given 1-2 minutes to respond individually. After showing the class' pooled responses graphically, I asked students who answered various responses to defend their answers, and then elaborated, asked follow up questions, or resumed lecture. In the second protocol, I asked students to discuss the question with nearby students after they saw the initial class response data. This discussion usually lasted 2-3 minutes.

Students then re-entered their responses individually without comment from me. I then discussed the answers as above. Eight of the 20 ARS questions used in this report used the student discussion protocol, while 12 used individual student reporting only.

Student attitudes about ARS were surveyed in two ways. Routine end of course surveys were done on overall assessment of value, and the 54 responses were gathered by web based surveying by our office of medical education. Since class size was 94, this represents 56% of students. It is unknown how many ARS sessions were attended by these respondents. All likely attended at least one, since a possible option was "did not attend an ARS session".

In addition, I surveyed students about their preferences of ARS learning vs. other modalities, and about their more generalized impressions, by use of ARS surveys done in class at the end of the series of lectures. Depending on attendance and participation that day, these ARS surveys yielded 40-46 responses.

RESULTS/DISCUSSION

ARS as Formative Assessment

A weakness of traditional lecture is its disengagement from a given class' and individual learners' specific needs. The lecturer often exists in a bubble, delivering the same content regardless of context. Since students may have varying learning styles, daily curricular schedules, and degree of fatigue, greater instructor awareness of their comprehension and attention can lead to more stimulating and focused learning sessions. This sensitivity to learner needs increases learner attention and involvement.

ARS questions are very useful here as the punctuation of a lecture segment, in order to assess student comprehension. In order to do so, the questions should be conceptual, asking learners to apply principles given in the lecture block, and not simply ask them to recall a specific fact. Such questions are best done in the form of experimental or clinical vignettes, as is now done in USMLE licensing examinations. See Table 1 for examples of this type of ARS question. As discussed below, if ARS reveals that students have not mastered the concept, a lecturer may need to spend additional time on it, rather than moving on in a fixed schedule. For example, question 1 requires learners to synthesize the preceding hour of material on different types of metabolic acidosis, using the vignette and lab values to classify the disorder, and engage in two step thinking in identifying a cause of the identified disorder (here, non anion gap metabolic acidosis with hyperkalemia). Many students missed this question, and further questioning of them revealed many cognitive problems, including focusing on only one value or vignette item, lack of a systematic analysis of the acid base disorder, and reliance on memorized lists rather than global analysis. The time spent on this question, in which I modeled my approach to its solution, gave students a framework for success in solving these problems.

ARS in stimulating knowledge application and analysis

The lecture has traditionally been the reservoir of facts. Most books and presentations on "Powerpoint® Technique" emphasize clarity and presentation of bullet point slides, on an assumption that data presentation is the main objective of any lecture. Textbooks are usually written in the same manner--comprehensive and organized coverage of facts is the most common structural underpinning of most medical and science texts. But should this be the purpose of a lecture for our students? Secondary school teaching typically is far more interactive, even in groups of 30-40. It is only on arrival to college that we treat the students to the one-way lecture, on the assumption that this is somehow preferred for these mature learners. It is certainly efficient. But even mature learners need to be motivated, stimulated, and challenged to move beyond the Bloom cognitive process of remembering to that of understanding, analyzing, and applying[4]. This should be our goal for students entering the complex synthesis that characterizes clinical care, and these skills must be rehearsed before intensive clinical care begins.

The expert teacher must reach many types of learners, including those who first need the facts, as well as those who want the facts presented conceptually and contextually. ARS can effectively facilitate such a learning system. Students can be provided a well written text or syllabus that lays out the facts clearly, and introduces terminology. Then, the ARS "lecture" can follow up with explanation, explication, and exemplification. Students often state in course evaluations that they benefit most from these sessions if they have read the facts first, so that they can come to the ARS session ready to extend their knowledge.

The ARS questions, if written to emphasize understanding and application, give students an idea of the level of knowledge expected by the instructor and guide their subsequent study away from rote memorization. To do this, I feel it is important that the questions be challenging, so students are motivated to review and learn more after the session (see Table 1). My ARS questions are a mix of single best answer and multiple best answer, and the students' average correct response rates for each type in 2007 was 63% and 60%, respectively. Beginning or more insecure learners might benefit from less difficult questions that simply confirm memory of facts, but since I use these ARS sessions primarily to stimulate higher cognition, I feel that providing a false sense of mastery with easy questions undercuts the goal of motivating further study and self directed learning.

The use of clinical or experimental vignettes, amplified with the active learning of clicking on answers in the ARS format, can reformat the "lecture" into truly interactive

Table 1. Sample of ARS Question

1. A 62 year old man has Type 2 diabetes and hypertension. He comes to clinic complaining of diarrhea. He is on lisinopril, Dyazide (hydrochlorthiazide + triamterene) and metformin.

Na 138	K 5.2	Cl 112	HCO3 18
Glucose 220	BUN 23 Cr 1.3		

Serum pH = 7.34 pCO2 = 34
Urine pH = 4.6

Which of the following best explains his acidemia?

A. Diabetic ketoacidosis
B. Metformin
C. Dyazide
D. Diarrhea
E. Distal RTA
F. Lisinopril

2. A 50 year old man with a history of CHF has pulmonary and peripheral edema, and a blood pressure of 100/60. Which of the below is most likely (Select ALL that apply)?

A. His extracellular volume is decreased
B. His effective circulating volume is decreased
C. His total body sodium is increased
D. His serum sodium concentration is increased
E. His urine sodium concentration is increased

learning session in which students extend their factual knowledge into application and analysis, and set the stage for deeper learning at home. If one therefore reconsiders what a "lecture" is, then the pressure to cover all the factual material disappears. In this model, the lecture is an active, energizing supplement to the written syllabus or text.

<u>ARS for formative feedback (to students and instructor)</u>

Students commonly complain that lecturers assume knowledge that is either more or less advanced than their actual level. Since effective learning occurs best when built upon a base of preexisting understanding[5], the effective lecturer should assess this base regularly. This can be done in advance by reviewing the students' prior curriculum and the specific content of preceding lectures. However, ARS offers the advantage of real time assessment of student preparation and understanding. Normally this is done by an assessment question at the end of a lecture segment, ideally spaced about 20 minutes after a similar question, in order to minimize student lapses in concentration. However, an ARS question can also be used to <u>begin</u> a lecture segment, showing students what they do not know and provoking interest in the upcoming segment. This is especially useful if students have already "covered" a topic in a previous course or lecture. The question can frame how their knowledge will be <u>extended</u>, not just repeated, in the succeeding minutes. In this use of ARS, it is not necessary that students successfully answer the question. In fact, I frequently do not discuss the correct answer after showing the response. Instead, I mention that the upcoming lecture segment will clarify the issue, and generally return to the question later, either as a lecture slide, or as a re-take of the question by the class. To summarize, ARS provides useful formative feedback for instructors and for students. For students, it joins end-of-syllabus chapter review questions and online exams as ways for my students to practice challenging questions of the type that I will ask on summative exams.

<u>ARS as a vehicle for student peer interaction</u>

The most common way in which ARS is used is the sequence: lecture → ARS question → answer → instructor explanation. While engaging, this still keeps most students in a passive role. After reflecting on team-based learning strategies [6-8], I now often use the ARS system to stimulate student-student interaction. After having students individually answer the ARS question, I show them the class distribution of answers, without indicating the correct answer. Then I ask them to discuss their answer with nearby colleagues for 1-2 minutes, and then individually re-enter their answer. Students usually respond more accurately after such discussion (improving their correct response rate by 2-10%), even when the correct answer was initially a minority response. Perhaps additional reflection time improves response, or perhaps students with better understanding are persuasive in the brief interactions with their colleagues. In any case, students gain the satisfaction of benefiting from peer interactions in improving their own understanding. If students self-correct, I frequently offer little additional explanation after the peer discussions, since the students have gained understanding on their own. Most students enjoyed the addition of peer discussion to the ARS sessions, but this is variable: 49% preferred student-student interaction, 27% preferred individual ARS use alone, and 24% were undecided (n = 41). Thus ARS can provide a collegial learning process that echoes some goals of problem-based learning[9, 10], but now with a large class.

<u>ARS for instructor feedback</u>

A limitation of the lecture/transmission mode of teaching is its lack of real time feedback from learners. The lecture may have been delivered, but did learning occur? Traditional questions posed by the lecturer to the students often prompts more extroverted or knowledgeable students to respond, but this may not reflect the knowledge or engagement of the group as a whole. ARS provides an ideal medium to improve this student feedback to instructors (a vivid anecdote from a course in embryology teaching gives testimony to the lessons learned when student understanding is actually assessed)[11]. Regular use will tell the instructor whether points made were absorbed and understood. Low correct response rates on questions prompt the conscientious instructor to rephrase, repeat, or exemplify the poorly understood concept, so that learning occurs in the teachable moment. This inevitably “slows down” the lecture and may require the instructor to reduce the number of slides presented. However, if the traditional lecture is to be transformed into an interactive learning session, this “problem” is a good thing. Our students often complain that instructors may show in excess of 60 slides in a 50 minute lecture, and one lecturer at my institution has 120 scheduled for such a presentation. The feedback provoked by ARS can provide a needed brake on such excess.

<u>ARS to make lecture fun</u>

While learning should not be primarily an entertainment, enjoyment certainly belongs in any learning session. Humor, visual props, colorful slides, and animations are frequent lecture props, used by even traditional speakers to enliven the proceedings. However, these still remain mainly one-way, transmission oriented devices, in which the students remain observers, albeit more amused observers. ARS offers a platform for true interaction with students within the learning session, and provides a real sense that the teacher is interacting <u>with</u> learners, not just talking to them. This human contact allows a more personal interaction, even with a large group of students, and is a strong attractant for students who value the human interaction as key to learning (e.g. students with strong F domain in the Myers Briggs type indicator)[12]. Such students are often most put off by traditional lectures.

<u>Limitations and Challenges of ARS</u>

ARS is not an end in of itself. It is simply a new technological innovation that, if used well, can achieve the above aims. I list below several ARS pitfalls that should be avoided so that ARS does not detract from learning.

1. Overuse: One lecturer recently substituted ARS questions *en bloc* for his traditional lecture slides, without providing students with preliminary content via readings or other media. While the intent of session interactivity was appreciated, the students were made to answer ARS questions with only very limited knowledge, and resented the frustration of not being able to consolidate knowledge appropriately. Students surveyed after my ARS sessions strongly felt (92%, n=54) that three questions administered per 50 minute lecture was an ideal frequency, with the remainder evenly divided between wanting more and wanting less. They also felt that ARS works best on a base of factual knowledge, allowing them to explore its applications in a medical environment.

2. Overload: ARS cannot be grafted onto an already loaded slide presentation. Each slide takes 2-3 minutes at minimum, given the time to answer the question and to discuss the results. This often extends to 5 minutes or more. Obviously, pre-existing slides must be deleted to accommodate this, unless the session is lengthened, a rarity in the current minimalist lecture environment. This means that the instructor must prioritize the lecture content, using ARS to teach fewer concepts more deeply. Teaching fewer things with more depth, however, is a goal of most experienced teachers and leads to greater retention and application[5].

3. Poorly written questions: In order for ARS to best provoke and stimulate students, questions should contain uncertainty, controversy, or analysis/application of material. Simple factual recall questions do not do this well. For my second year medical students I use questions similar to, or more advanced than, USMLE Part I questions (Table 1). These are normally based on experimental or clinical vignettes that provoke the students to analyze and apply their knowledge. This approach has the additional advantage of preparing students for the more analytic questions ideally used on summative course and licensing examinations.

4. Inadequate faculty development: The availability of an ARS system usually leads to initial administrative and student enthusiasm, typically because it is first used by the extroverted “early adaptor” instructor who infuses it with excitement [13]. Once the glowing initial reviews come in, other instructors may use it, but sometimes without any real preparation or orientation other than on the technical aspects of building the session. This often leads to the above listed mistakes, or a stylistic discontinuity in which a lecturer uses ARS questions but does not really engage the students verbally or emotionally. Students may then pan the entire technique. To avoid this drawback, our

school provides regular lunchtime seminars for interested instructors in which experienced ARS users share tips and demonstrate effective practice. In addition, we have begun demonstrating ARS to entire departments at their faculty meetings so that all instructors can learn about ARS, thus enlivening a departmental course lecture curriculum systematically. Several initially reluctant instructors have told me that ARS helped them emerge from behind the podium and better engage the class, and improved their lecture technique generally. In these cases the technology facilitated a change in instructor behavior.

Student Response

These second year medical students rated the educational value of ARS questions highly (6.8 out of a 7 point score, n=54). More affective responses are quantitated in Table 2. Post course comments indicated that individual students valued different types/uses of ARS questions:

> The audience response system is great for gauging our comprehension of materials just presented, and helps to further cement our newly acquired knowledge by making us recall and actively apply it to complex scenarios. I think it's awesome!

> (The instructor) uses it the way it was meant to be used. He goes over the concepts and then puts a little twist into a question and then we can discuss it.

> I liked that he didn't give us a question about something we haven't seen yet.

We are currently doing a systematic study of faculty lecture evaluations pre- and post- incorporation of ARS to further assess this issue. ARS may also motivate greater student attendance (this is not required at my university). Lecture attendance in my course, which has declined for the past several years, subjectively increased this year (no precise data available). While it is not clear that this trend, if verified, is due to ARS alone, others have reported increased learner participation rates with institution of ARS[14]. Overall student exam scores have not changed with use of ARS, but the course already had a rich assortment of sample questions for students to use, so this is not surprising.

Table 2. Students answering “yes” to various descriptors of ARS sessions (n=44)

Added needed variety to the session	75.0%
Clarifying	70.5%
Confusing	25.0%
Distracting	4.5%
Provided feedback on my understanding	77.3%
Stimulating	72.7%
Tedious	2.3%

Conclusion

While no technology serves as a panacea for indifferent or poorly prepared instructors, appropriate use of ARS increases interactivity in large group learning sessions. It joins team-based learning as another formal option for instructors who feel that their sessions need to become more interactive. The reduction of formal lecture time has been encouraged by many accrediting bodies such as LCME, but should not be done for that reason alone. Declining student attendance at lectures nationwide shows that students are increasingly needing a rationale for attendance, and if not given one, will choose a distance learning strategy. In my view, given the wealth of current online and written resources for students, this is a justifiable view. Any time used for whole class presentations should have a clear rationale beyond simple presentation of facts, which can be done effectively at home. Is a lecture that duplicates preexisting written materials worth the time? Audience response systems is one means of taking a large group session to a more stimulating, interactive level, and provides a format for professional faculty to re-engage with students and return to the art of teaching, not just lecturing.

References

1. Kassebaum, D.G., Cutler, E.R., Eaglen, R.H. The influence of accreditation on educational change in U.S. medical schools. *Academic Medicine.* 1997;72:1127-1133.
2. Caldwell J.E. Clickers in the large classroom: current research and best-practice tips. *CBE Life Scientific Education.* 2007;6:9-20.
3. Knight J.K., Wood, W.B. Teaching more by lecturing less. *Cell Biology Education.* 2005;4:298-310.
4. Anderson, L.W., Krathwohl, D.R. (eds.). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of*

Bloom's Taxonomy of Educational Objectives. *New York,NY*: *Longman*.2001.
5. Bransford, J.D., Brown, A., Cocking, R.R. (eds.). *How People Learn: Brain, Mind, Experience, and School*. *Washington,D.C*:*National Academy Press*.2000.
6. Michaelsen, L.K. *Team-based learning for health professions education : a guide to using small groups for improving learning*. 1st ed. *Sterling,VA*:*Stylus Publications*.2008.
7. Koles, P., Nelson, S., Stolfi, A., Parmelee, D., Destephen, D. Active learning in a Year 2 pathology curriculum. *Medical Education.* 2005;39:1045-1055.
8. Nieder, G.L., Parmelee, D.X, Stolfi, A., Hudes, P.D. Team-based learning in a medical gross anatomy and embryology course. *Clinical Anatomy*.2005;18:56-63.
9. Albanese, M.A, Mitchell, S. Problem-based learning: a review of literature on its outcomes and implementation issues. *Academic Medicine*.1993;68:52-81.
10. Norman, G.R., Schmidt, H.G. The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine.* 1992;67:557-565.
11. Wood, W.B. Clickers: A Teaching Gimmick that Works. *Developmental Cell* 2004;7:796-798.
12. Bayne, R. *The Myers-Briggs type indicator : a critical review and practical guide*. *London*:*Chapman and Hall*.1995.
13. Gladwell, M. *The tipping point : how little things can make a big difference*. 1st ed. *Boston*:*Little, Brown*.2000.
14. Homme, J., Asay, G., Morgenstern, B. Utilisation of an audience response system. *Medical Education.* 2004;38:575.

The Pre-Entry Program at UTMSH: Effect on Academic Performance of First-year Medical Students

William E. Seifert, Jr., Ph.D.,[1] R. Andrew Harper, M.D.[2]

[1]Department of Biochemistry and Molecular Biology
[2]Department of Psychiatry and Behavioral Sciences and
Office of Educational Programs
The University of Texas Medical School at Houston
6431 Fannin Street
Houston, TX 77030 USA

Phone: (+)1-713-500-6283 Fax: (+)1-714-500-0652 Email: William.E.Seifert@uth.tmc.edu

ABSTRACT

The Pre-Entry Program at The University of Texas Medical School at Houston is established to assist entering students who are judged to be at risk for academic difficulty. It requires a significant commitment of time on the part of faculty, staff and students. The effectiveness of this program needs to be evaluated.

This was a causal-comparative study of students invited to the Pre-Entry Program between 1999 and 2005. Students were self-selected into two groups, attendees (n = 174) and decliners (n = 81). The proportion of students with unsatisfactory performance and the rate of attrition from the first year class were compared for each group by a Pearson chi-square test. An analysis of covariance was used to compare the academic achievement as measured by the National Board of Medical Examiners (NBME) Subject Examinations in Biochemistry, Gross Anatomy and Physiology using Medical College Admission Test scores as the covariate.

There were no statistically significant differences in the incidence of unsatisfactory performance or in the rates of attrition from the first-year class between accepters and decliners or in the mean performances of the two groups on the NBME Examinations in Biochemistry and Physiology. The decliners had a statistically significant higher mean performance on the NBME Gross Anatomy Examination ($p = .04$), although the effect size ($d = 0.29$) was not educationally significant.
The effect of the Pre-Entry Program on the academic performance and attrition rate of at-risk first-year medical students is minimal. The program should remain voluntary and further studies should be performed to determine the non-academic effects of the program.

INTRODUCTION

The University of Texas Medical School at Houston (UTMSH), as an educational institution of The State of Texas, endeavors to attract students who are the best and brightest among the applicant pool while at the same time recruiting a student body of diverse ethnic background that reflects the population of the State of Texas. As a state medical school, a primary goal of UTMSH is to produce physicians that will meet the health care needs of that population. Thus, UTMSH is committed to attracting and retaining students from under-represented populations within the State of Texas including rural, disadvantaged, and minority students (including African-American, Hispanic-American and Native American).

Efforts to increase the percentage of underrepresented minority students in medical schools in the United States started during the period of social activism and the civil rights movement of the late 1960s. In 1968, at the annual meeting of the Association of American Medical Colleges (AAMC), a task force was established to set goals for minority-group enrollment and the strategies by which these goals could be achieved.[1] Nickens, *et al*.[1] describe the increase in underrepresented minority groups since that meeting as having three phases. In the first phase, there was a rapid increase in minority enrollment from about three percent in 1968 to almost ten percent in 1974. This was followed by a period of stagnation where the absolute number of underrepresented minorities increased due to the increase in the number of accredited medical schools[2] but their proportion of the total medical school enrollment remained relatively constant at between 8-10 percent. In 1990, the AAMC initiated Project 3000 by 2000, an effort to increase the number of underrepresented minority entering U.S. medical schools each year to 3000 by the year 2000. These renewed efforts to increase minority enrollment resulted in a sharp 27 percent increase over the three-year period from 1990 to 1993.[1,2] In response to these efforts to increase minority enrollment, many medical schools have initiated supplemental programs to recruit and retain minority and disadvantaged students.

Programs for underrepresented minority or educationally disadvantaged students sponsored by medical schools can be generally classified into three types. One type is designed for undergraduate college students to enhance their academic skills and increase their exposure to the practice of medicine. The goals of these **undergraduate enrichment programs** are to help the students be better candidates for medical school and to improve their chances of admission. They are also frequently recruitment tools for the medical schools. These summer programs usually emphasize academic preparation as well as allow the students to observe physicians in their practice of medicine. A second type of program is a **post-baccalaureate program** that is designed to help students who have been rejected by medical schools improve their academic credentials so that they can be more competitive in the application process. Most of these programs are usually year-long programs designed to improve the participant's academic skills, especially in the sciences. The third type of program is designed for students who have already been accepted into medical school but are judged to be at-risk to encounter academic difficulty in medical school, particularly in the basic science years. These **pre-entry programs** are designed to prepare the students for the rigors of medical school and help ensure that they successfully complete their training. Most pre-entry programs are presented as five-to-six week sessions in the summer prior to the start of the first year.[3-8]

The results of studies to determine the effectiveness of pre-entry programs suggest that students who participate in these pre-entry programs encounter less academic difficulty than students who were invited to the programs but did not participate. Richardson and Saffran[3] report that students who participated in the pre-entry program at the Medical College of Ohio performed better in all subjects except behavioral science than invitees that did not participate. A report describing the pre-entry program at the Boston University School of Medicine[4] showed similar results. Despite having lower MCAT scores, participants in the pre-entry program had significantly higher proportions of acceptable grades in two courses than did minority non-participants and slightly higher rates of acceptable grades in six of the nine courses presented in the first year. Kornitzer, *et al.*[8] showed that students who participated in the pre-entry program at Mount Sinai School of Medicine had less academic difficulty in their first year of medical school. They propose that these pre-matriculation programs play an important role in the academic success of certain educationally disadvantaged medical students because of the academic preparation they provide for some first-year courses, the opportunity to develop study-skills and confidence, and the relationships that are built with fellow attendees and faculty members.

In 1990, the Pre-Entry Program (PEP) at UTMSH was implemented to assist entering students who were judged likely to encounter academic difficulty in Medical School, particularly in the first year. The PEP is a five-week curriculum presented in the summer prior to matriculation that is designed to prepare the students for the rigors of medical school and help ensure that they successfully complete their training. During these five weeks, the students are presented a review of basic concepts of biomedical sciences (biomolecules, pH, thermodynamics, and kinetics) in addition to instruction in four areas of the basic sciences: human anatomy, biochemistry, neuroscience and physiology. In the biochemistry classes, students are introduced to protein structure, enzymatic catalysis, and intermediary metabolism, focusing on the generation of energy by catabolism of carbohydrates, lipids and proteins. In anatomy, the objectives are to cover basic aspects of musculoskeletal anatomy and cardiopulmonary systems. After introduction to the basic organization of the body, students dissect the muscles of the back and observe the spinal cord, spinal nerves and meningeal coverings. They then study the thorax, focusing on the lungs, heart, and great vessels and sympathetic chain. The physiology sessions introduce the students to the concept of homeostasis and the role of membranes in maintaining the composition of the various body fluid compartments. Students also study the concepts of forces and flows in the body using the cardiovascular system as an example. The neuroscience presentations cover the basic biophysical aspects of neuronal signaling, including the ionic basis of resting potentials, production of action potentials and their propagation along axons, synaptic organization, mechanisms of neurotransmitter release and postsynaptic response, and the integration of responses from multiple presynaptic cells. The students also participate in workshops on ethics and professionalism, improving study skills, managing the stress of medical school, and taking patient histories (from a standardized patient). These introductory courses and workshops are designed to help the student adjust to the rigors of the academic program of medical school.

Entering students are invited to participate in the PEP after they have been accepted into UTMSH. The criteria for selection include a low Medical College Admission Test (MCAT) score or grade point average (GPA) from their undergraduate institutions, a non-science major, a lapse in education since graduation from college, an older age, and attendance at a non-selective undergraduate institution.

The PEP requires a significant commitment on the part of the faculty and staff who present the program as well as on the part of the entering medical students, who sacrifice time that either could have been spent in summer employment or vacation. Demands on the time of medical school faculty are increasing and time that can be devoted to teaching is at a premium. Therefore, it is imperative to determine whether the resources devoted to the PEP are effective in decreasing the incidence of academic difficulty for those students who participate in it.

Following the first five years of its implementation, an initial evaluation of the PEP suggested that those who attended the PEP had better academic performance in their first-year classes and a lower attrition rate than those who declined the invitation to attend.[5] Since the time of that initial evaluation, the PEP has evolved as faculty members involved in the program have changed; thus a new and updated study was conducted to better understand the impact of the PEP on the academic performance and attrition of first-year medical students at UTMSH.

Purpose of the Study

The purpose of this study was to describe the effect of the PEP on the academic achievement of medical students during their first year of medical school. This study addressed two research questions: 1) How did the performance of first-year medical students who attended the PEP compare with those who were invited to the Program but declined to attend? and 2) How did the rates of attrition from the first-year class compare between those who attended the PEP and those who were invited but declined to attend?

The study investigated the following non-directional research hypotheses: 1) There is a statistically significant difference in the percentage of students with unsatisfactory performance in the first year of medical school between those who attended the PEP and those who were invited but declined to attend; 2) there is a statistically significant difference in the performance on the NBME Subject Examinations in Biochemistry, Gross Anatomy, and Physiology between first-year medical students who attended the PEP and those who were invited but declined to attend; and 3) there is a statistically significant difference in the rates of attrition from the first-year class between students who attended the PEP and those who were invited but declined to attend.

MATERIALS AND METHODS

The participants in this study were those students who were accepted into UTMSH and were invited to participate in the PEP between 1999 and 2005. These students were invited to the program because they were judged likely to encounter academic difficulty in medical school due to their low MCAT score, low undergraduate GPA, a non-science major, a lapse in education since graduation from college, an older age, and/or attendance at a non-selective undergraduate institution. The participants were divided into two groups, (1) those who accepted the invitation and attended the PEP (PEP-A; n = 174) and (2) those that declined the invitation and did not attend (PEP-D; n = 81).

Academic achievement in Biochemistry, Gross Anatomy and Physiology were measured by the National Board of Medical Examiners® Subject Examinations in Biochemistry, Gross Anatomy and Physiology. These examinations have been used by the course directors of the respective courses as an end-of-course examination and as an assessment of student achievement in the course. According to the NMBE Subject Examination Program Information Guide,[9] the Subject Examinations provide national norms and the items in the examination are selected only after extensive review and pretesting. The examinations are proprietary and confidential and were administered under specific procedures established by the NBME. Prior to December, 2004 the scores were reported on a 70/8 scale and the standard error of measurement was 2.5. From December, 2004 onward the scores are reported on a 500/100 scale and the standard error of measurement is 40. The reliability coefficients on the NBME Subject Examinations range between .85 and .89.[10] The Physiology Examination has only been administered in the Physiology course since 2002, and comparisons for the Physiology Examination were limited to the years 2002 through 2005. For statistical analyses, 70/8 scores were converted to 500/100 scores using tables provided by the NBME.

The incidence of unsatisfactory performance and attrition rates from the first-year class were determined from data obtained from the official student records kept by the Registrar of The University of Texas Health Science Center at Houston. Data collected from the official student records included MCAT scores, undergraduate GPA, final grades in first-year courses, and whether the student withdrew or was dismissed, took a leave-of-absence, had to remediate a course, had to repeat the year or opted for the Alternate Pathway (the option to complete the first-year courses over a two-year span). To preserve confidentiality, each participant was assigned a number by the Office of Educational Programs, which collected and collated the archived student records prior to analysis.

The following statistical analyses were performed to compare the demographics of the two groups (PEP-A and PEP-D): *t*-test for unpaired samples for incoming age, undergraduate GPA and numeric MCAT score; Mann-Whitney *U* test for MCAT writing score; and Pearson χ^2 technique for gender and ethnicity. The rates of attrition and incidence of academic difficulty between those who participated in the PEP and those who declined the invitation were compared using the Pearson χ^2 technique. Academic achievement in Biochemistry, Gross Anatomy and Physiology as measured by the NBME Subject Examinations were compared by an analysis of covariance using MCAT score as the covariant. Only those students who took the respective NBME Subject Examination in their incoming year were included in the analysis. All *p* values < .05 were considered significant. Statistical analyses were performed using *Systat 12* software (Systat Software, Inc., Richmond, CA.

RESULTS

Overall, 255 students who matriculated at UTMSH were invited to the PEP between entering years 1999-2005. Of these, 174 students accepted the invitation and participated in the PEP; 81 students declined the invitation and did not participate. As shown in Table 1, a comparison of the two groups indicated that there were no statistically significant differences between the accepters and the decliners with respect to their ages at entry, undergraduate GPA, or MCAT scores. There also was no statistically significant difference in the ethnicities of PEP accepters and decliners (χ^2 = 7.97, d.f. = 7, *p* = .335) nor in their MCAT writing scores (*U* = 6721.50, d.f. = 1, *p* = .363). There was, however, a statistically significantly greater proportion of males who declined the invitation to the PEP (χ^2 = 5.83, d.f. = 1, *p* = .016) than those who accepted and participated in the PEP. Thus, the two groups compare well with regards to their demographics.

Table 1. Demographic comparison of those who participated in the PEP (PEP-A) and those who declined the invitation (PEP-D)

	Gender				
	Female	Male	Entering Age (years)	Undergrad GPA	MCAT
PEP-A	97	77	24.97 ± 4.41	3.49 ± 0.25	24.12 ± 3.28
PEP-D	32	49	24.04 ± 2.61	3.50 ± 0.23	23.88 ± 2.61
test	χ^2 = 5.83		*t* = 1.72	*t* = 0.14	*t* = 0.57
d.f.	1		253	253	253
p	.016		.086	.886	.566

Attrition from the first-year class was determined from the number of students who withdrew or were dismissed, took a leave-of-absence during the first year, had to repeat the first year due to academic difficulties, or opted for the Alternate Pathway. The rates of unsatisfactory performance were determined from the number of students who received a non-passing grade (marginal performance or failure) in one of the first-year classes. As shown in Table 2, a Pearson chi-square analysis showed no statistically significant difference in the rates of attrition from the first-year class or in the incidence of unsatisfactory performance between those who participated in the PEP and those who declined an invitation to the PEP. Therefore, the research hypotheses that there were statistically significant differences between PEP accepters and decliners in the proportion of students with unsatisfactory performance and in the proportion of students who left the entering first-year class were not accepted.

An analysis of covariance (ANCOVA) was performed on the student's performance on the NBME Subject Examinations that were given as final exams in the Biochemistry, Gross Anatomy, and Physiology courses taken by the first-year medical students with the student's MCAT scores used as the covariate. Only those students who took the examination during their entry year were included in the analysis. Tables 3 – 5 show the results of the analyses of covariance.

The results of the ANCOVA presented in Tables 3 and 4 indicate that there is no statistically significant difference between those who participated in the PEP and those who were invited but declined in their performance on the NBME Biochemistry (*F* = 0.001, *p* = .975) and Physiology Subject Examinations (*F* = 0.14, *p* = .705). Therefore, the research hypotheses that there were statistically significant differences between the PEP acceptors and decliners in their performance on these two examinations are not accepted. The results presented in Table 5, however, show that the performance on the NBME Gross Anatomy Subject Examination by those who declined the invitation to attend the PEP (adjusted mean = 496.02) was statistically significantly higher than those who participated in the PEP (adjusted mean = 474.71, *F* = 4.29, *p* = .040). Therefore, the research hypothesis that there is

Table 2. Comparison of rates of attrition and unsatisfactory performance

	Attrition		Unsatisfactory Performance	
	No	Yes	No	Yes
PEP-A	136	38	127	47
PEP-D	67	14	66	15
χ^2	0.71		2.17	
d.f.	1		1	
p	.401		.141	

Table 3. Results obtained from ANCOVA – NBME Biochemistry Examination

		MCAT	Biochemistry NBME			
Group	N	Mean ± SD	Mean ± SD	Adjusted Mean	*F*	*p*
PEP-A	162	24.20 ± 3.25	491.42 ± 81.76	491.11	0.001	.975
PEP-D	76	24.11 ± 2.42	490.79 ± 69.24	491.44		

Table 4. Results obtained from ANCOVA – NBME Physiology Examination

		MCAT	Physiology NBME			
Group	N	Mean ± SD	Mean ± SD	Adjusted Mean	*F*	*p*
PEP-A	92	24.63 ± 3.24	447.39 ± 77.33	446.31	0.144	.705
PEP-D	44	24.16 ± 2.27	449.09 ± 70.01	451.35		

a statistically significant difference between PEP accepters and decliners in their performance on the NBME Gross Anatomy Subject Examination is accepted. Since the effect size obtained, *d* = 0.29, was less than a third of a standard deviation, it can be argued that this difference is not educationally significant.

Table 5. Results obtained from ANCOVA – NBME Gross Anatomy Examination

		MCAT	Gross Anatomy NBME				Effect size
Group	N	Mean ± SD	Mean ± SD	Adjusted Mean	*F*	*p*	*d*
PEP-A	168	24.14 ± 3.29	474.67 ± 83.05	474.71	4.29	.040	0.29
PEP-D	73	24.15 ± 2.40	496.10 ± 61.05	496.02			

DISCUSSION

The purpose of this study was to determine the effect of the Pre-Entry Program (PEP) on the academic achievement of at-risk students in their first year of medical school. Students accepted into UTMSH are invited to the PEP if they have one or more of the following selection criteria: a low Medical College Admission Test (MCAT) score or grade point average (GPA) from their undergraduate institution, a non-science major, a lapse in education since graduation from college, an older age, or have attended a non-selective undergraduate institution. Students who are invited to the PEP may decline the invitation for several reasons including family commitments during the summer prior to matriculation, summer jobs, military service, vacation or travel plans, remediation of undergraduate requirements or a belief that the program would not benefit them. A comparison of the two groups revealed that male invitees to the PEP are more likely to decline the invitation to participate.

The results of this study do not support the hypotheses that the PEP has an effect on the incidence of unsatisfactory performance for at-risk first-year medical students or on the rates of attrition of these students from the first-year class due to dismissal, leave of absence, having to repeat the first-year, or entering the Alternate Pathway. This is contrary to what was reported in an earlier evaluation[5] that was done after the first five years of the PEP. Unfortunately, the data from that study were not available to the authors in order to make a valid comparison of the two cohorts.

In order to obtain a standardized measure of the academic achievement of the at-risk first year students, the performance on the National Board of Medical Examiners Subject Examinations in Biochemistry, Gross Anatomy and Physiology were compared between the students who participated in the PEP and those who declined the invitation. These NBME Subject Examinations are utilized as the final examinations in each of the respective first-year courses. This comparison did not reveal any statistically significant differences in the performance of the two groups on the Biochemistry or Physiology examinations. Students who declined the PEP did perform statistically significantly higher on the Gross Anatomy examination, although the low effect size indicated that the difference was not educationally meaningful.

While this study attempted to determine the efficacy of the PEP on the first-year medical school academic performance and whether participation in the PEP affected the attrition rate from the first-year class, it should be noted that there are many factors that can affect academic achievement and whether a student chooses or is forced to alter their path towards a medical degree. These include personal characteristics such as motivation and self-discipline as well as personal situations such as illness, child rearing responsibilities or relationship difficulties that distract the student from his/her academic endeavors. These factors were not addressed in this study. An additional limitation of this study was that no attempt was made to compare subgroups (such as non-science majors, students with a lapse between graduation and entering medical school, socio-economic status, ethnicity) within the students who participated in the PEP and those who were invited but declined to attend. Additional limitations include: 1) the two groups are self-selected and 2) the reasons why those who declined the invitation to the PEP were not obtained to determine if there are common non-academic characteristics that explain their decision not to participate.

A reasonable conclusion of this study is that the results in academic performance and retention of students do not justify the resources put into the PEP. However, a review of the student evaluations of the PEP following the 2005 and 2006 presentations of the program indicate that the students almost unanimously feel that the program is valuable and would recommend it to an incoming student (data not shown). It is their perception that the PEP has better prepared them for the rigors of the first-year of medical school. The results of this study do suggest that the program should remain a voluntary program for incoming at-risk students; at this stage of their education, they can best judge if they need the program or not. It also suggests that further studies should be performed to determine in what ways other than academic the PEP is beneficial to the students.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the assistance of Elizabeth Green in the Office of Educational Programs in collecting and collating the student data. Without her help this study could not have been performed. The authors also acknowledge the assistance of Pat Caver in the Office of Student Affairs and the course directors and coordinators of the first-year courses in providing the data for this study. The authors also gratefully acknowledge Dr. Allison Ownby in the Office of Educational Programs for her review and helpful comments on this manuscript. This study was approved the UTHSC-Houston Committee for the Protection of Human Subjects (Protocol #HSC-MS-06-0606).

REFERENCES

1. Nickens, H. W., Ready, T. P., and Petersdorf, R. G. Project 3000 by 2000 – racial and ethnic diversity in U.S. medical schools. *New England Journal of Medicine. 1994;* 331(7):472-476.
2. Cohen, J. J., The consequences of premature abandonment of affirmative action in medical school admissions. *Journal of the American Medical Association.* 2003; 289(9):1143-1149.
3. Richardson, B. L., and Saffran, M. Effects of a summer preview program of study skills and basic science topics on the academic performance of minority students. *Journal of National Medical Association*. 1985; 77(6):465-71.
4. Ugbolue, A., Whitley, P. N., and Stevens, P. J. Evaluation of a pre-entrance enrichment program for minority students admitted to medical school. *Journal of Medical Education* 1987; 62(1):8-16.
5. Varagoor, G., Butler, P.M., Lund, K., Rosenfeld, G.C., and Schultz, S.G. Pre-Entry Program at UT-Houston Medical School: An Assessment, *Proceedings of the Eighth International Ottawa Conference on Medical Education and Assessment,* Philadelphia, PA, July 12-15, 1998.
6. Thompson, H. C., III, and Weiser, M. A. Support programs for minority students at Ohio University College of Osteopathic Medicine. *Academic Medicine.* 1999; 74(4):390-392.
7. Williams, M. T. Pre-matriculation program at the University of South Florida College of Medicine. *Academic Medicine* 1999; 74(4):397-399.
8. Kornitzer, B., Ronan, E., and Rifkin, M. R. Improving the adjustment of educationally disadvantaged students to medical school: the summer enrichment program. *Mt Sinai Journal of Medicine.* 2005; 72(5):317-21.
9. *National Board of Medical Examiners® Subject Examination Program Information Guide, 10th edition*, 2004.
10. National Board of Medical Examiners, personal communication, 2006.

