

2007

**THE EIGHTH ANNUAL
MASTER OF HEALTH
PROFESSIONS EDUCATION
SUMMER CONFERENCE**

August 2-3, 2007

.....

**PERFORMANCE ASSESSMENT:
SCHOLARLY STANDARDS AND
SOCIAL ACCOUNTABILITY**

ARA TEKIAN, PhD, MHPE - CONFERENCE ORGANIZER
UNIVERSITY OF ILLINOIS AT CHICAGO

.....

JOHN NORCINI, PhD - CONFERENCE KEYNOTE SPEAKER
*FOUNDATION FOR ADVANCEMENT OF
INTERNATIONAL MEDICAL EDUCATION
AND RESEARCH*

.....

JEANNE K. HEARD, MD, PhD - CONFERENCE KEYNOTE SPEAKER
*ACCREDITATION COMMITTEES AT THE ACCREDITATION
COUNCIL FOR GRADUATE MEDICAL EDUCATION*

The Eighth Annual MHPE Summer Conference

August 2 – 3, 2007

**Room 106 & Faculty & Alumni Lounge
(College of Medicine West)**

Sponsored by:
Department of Medical Education
Ara Tekian, PhD, MHPE
Conference Organizer



MHPE Vision Statement

**Better Healthcare through Effective and Innovative Health Professions
Education Leaders and Scholars**

Purpose: The purpose of this activity is to provide a venue for Master of Health Professions Education students, alumni, and faculty to present and discuss innovations in health professions education, and for physicians in the audience to learn about ongoing educational research and program development topics.

Intended Audience: Physicians in any discipline and other health professions educators.

Program Objectives:

At the conclusion of this program, participants should be able to:

1. Describe new methods of performance assessment in clinical settings.
2. Evaluate the effectiveness of different assessment methods.
3. Apply newer concepts of test validation to performance assessment.
4. Apply current concepts of performance assessment to the ACGME competencies.
5. Identify and evaluate different types of scholarship presented during the conference.

The University of Illinois at Chicago (UIC) College of Medicine is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

The University of Illinois at Chicago (UIC) College of Medicine designates this educational activity for a maximum of 10.5 AMA PRA Category 1 Credits™. Physicians should only claim credit commensurate with the extent of their participation in the activity.

The Eighth Annual MHPE Summer Conference

August 2-3, 2007

Ara Tekian, PhD, MHPE

Conference Chairperson

Department of Medical Education, College of Medicine/University of Illinois at Chicago

PERFORMANCE ASSESSMENT: SCHOLARLY STANDARDS AND SOCIAL ACCOUNTABILITY

Meals, Breaks, and Reception will be in the Faculty Alumni Lounge (College of Medicine West).

All Paper and Works in Progress Sessions will be in Room 106 (College of Medicine West).

THURSDAY, AUGUST 2, 2007

8:30 – 9:00

CONTINENTAL BREAKFAST

9:00 – 9:05

WELCOME

Ara Tekian, PhD, MHPE

Conference Chairperson

Associate Professor, Department of Medical Education

9:05 – 9:10

MESSAGE FROM DEPARTMENT HEAD

Leslie J. Sandlow, MD

Head, Department of Medical Education

*Senior Associate Dean, Educational Affairs/College of Medicine,
University of Illinois at Chicago/Chicago, Illinois*

9:10 – 10:10

KEYNOTE ADDRESS

Future of Performance Assessment

John Norcini, PhD

*President and Chief Executive Officer, Foundation for Advancement of
International Medical Education and Research/Philadelphia,
Pennsylvania*

10:10 – 10:30

BREAK

10:30 – 12:00

SESSION 1: PAPERS

RESEARCH & EVIDENCE-BASED MEDICINE

CHAIR: JULIE GOLDBERG, PhD; DISCUSSANT: ALAN SCHWARTZ, PhD

**Assessing the Evidence for Evidence-Based Medicine Skills:
A Literature Review and Tool Box**

Deirdre Jenkins, MD

Department of Hematology and Hematologic

Malignancies/University of Calgary/Calgary, Alberta,

Canada

**Assessment of an Evidence-Based Medicine Course in a
Mexican Medical School: A Randomized Controlled Trial**

Melchor Sánchez-Mendiola, MD

Universidad Nacional Autónoma de México/Mexico City,

Mexico

Scholarly Research During Surgical Residencies:

A Competency-Based Approach

Dimitri Azar, MD

Department of Ophthalmology and Visual Sciences, Illinois

Eye and Ear Infirmary/University of Illinois at

Chicago/Chicago, Illinois

CHAIR: JULIE GOLDBERG, PhD; DISCUSSANT: STEVE DOWNING, PhD

**Which Software Program for Generalizability Studies is Best?
Comparing G_String II and EduG**

Pamela Derstine, PhD

Accreditation Council for Graduate Medical

Education/Chicago, Illinois

**Are Nurse, Peer, and Faculty Performance Evaluations in
Surgical Residents Reliable? A Generalizability Study**

Pamela A. Lipsett, MD, FACS, FCCM

Department of Surgery, School of Medicine /Johns Hopkins

University/Baltimore, Maryland

12:00 – 12:10

POSTER SESSION INTRODUCTION

DISCUSSANT: JANET RIDDLE, MD

**What Qualities Do Medical Students Like to See in Their
Clinical Teachers as Role Models?**

Zaka Khan, MD

Family Medicine/Baqai Medical University/Pakistan

Teacher's Training Workshop on Teaching-Learning Methodology and Evaluation in Four Medical Colleges of Nepal

Nirmal Baral, MD

Department of Health Professions Education, Department of Biochemistry and Molecular Biology/B.P. Koirala Institute of Health Sciences/Ghopa, Dharan, Nepal

Improving Clinical Skills in the Emergency Department: Impact of Curricular Change at the Ankara University Medical School

Ayşe Gürel, MD

Ankara University Medical School/Ankara, Turkey

12:10 – 1:15

LUNCH AND POSTER SESSION

1:15 – 2:15

SESSION 2: PANEL

ROLE OF COLLEGIALLY: DEVELOPING YOUR SCHOLARSHIP AGENDA

MODERATOR: ARA TEKIAN, PhD, MHPE

Melchor Sánchez-Mendiola, MD

Universidad Nacional Autónoma de México/Mexico City, Mexico

Liselotte Dyrbye, MD

Department of Internal Medicine/Mayo Clinic/Rochester, Minnesota

Barbara Barzansky, PhD, MHPE

American Medical Association/Chicago, Illinois

2:15 – 3:15

SESSION 3: PAPERS

GRADUATE MEDICAL EDUCATION

CHAIR: JANET RIDDLE, MD; DISCUSSANT: ILENE HARRIS, PhD

What Instructional Practices are Used by Emergency Medicine Physicians Viewed as Excellent Teachers?

Kristi Grall, MD

Department of Emergency Medicine/University of Arizona/Tucson, Arizona

Resident Assistants Can Successfully Reduce Duty Hour Violations While Improving Quality of House Officer Inpatient Educational Experiences

Hilary Haftel, MD, MHPE

Pediatrics and Communicable Diseases/University of Michigan/Ann Arbor, Michigan

To Call or Not To Call? Phone Communication Between General Surgery Residents and Attending Surgeons

Roy Phitayakorn, MD

Department of Surgery/Case Western Reserve University Medical Center/Cleveland, Ohio

3:15 – 3:35

BREAK

3:35 – 4:50

SESSION 4: WORKS IN PROGRESS

CHAIR: STEVE DOWNING, PhD; MODERATOR: RACHEL YUDKOWSKY, MD, MHPE

Fellowship in Secondary Hospital Medicine – Curriculum Design and Program Evaluation

Rashmi Vyas, MD

Department of Physiology/Christian Medical College/Tamil Nadu, India

Medical Student Self-Assessment vs. Standardized Patient Ratings

Gail Beeman, MD, MHPE

Robert J. Kaplan Clinical Skills Center/University of Tennessee, Health Science Center/Memphis, Tennessee

Carrying Portfolios Forward – Can Portfolios Be Strengthened for Use as Summative Assessment?

Kevin Van Kanegan, DDS

College of Dentistry/University of Illinois at Chicago/Chicago, Illinois

4:50 – 5:15

BREAK

5:15 – 6:30

MHPE GRADUATE RECOGNITION CEREMONY, 2007 BEST MHPE THESIS AWARD, AND RECEPTION

6:30

CONFERENCE ADJOURNS FOR THE EVENING

FRIDAY, AUGUST 3, 2007

8:00 – 8:30 **CONTINENTAL BREAKFAST**

8:30 – 9:30 **KEYNOTE ADDRESS**

Assessing Resident Performance: ACGME Current and Future Directions

Jeanne K. Heard, MD, PhD

Director of the Department of Accreditation Committees at the Accreditation Council for Graduate Medical Education/Chicago, Illinois

9:30 – 11:05

SESSION 5: PAPERS

EVALUATING TANGIBLES & INTANGIBLES

CHAIR: TIMOTHY MURPHY, PhD; DISCUSSANT: GEORGES BORDAGE, MD, PhD

Description and Impact of Using a Standard-Setting Method for Determining Pass/Fail Scores in a Surgery Clerkship

Nancy Schindler, MD

Feinberg School of Medicine/Northwestern University/Chicago, Illinois

A Policy Capturing Study Examining the Factors that Influence Image Quality in Plastic Surgery Education

Fred Watkins, MD, MHPE

Division of Plastic Surgery/Hotel Dieu Hospital/Kingston, Ontario, Canada

Improving the Rating Scale for the Communication Skills OSCE

Cherdsak Iramaneerat, MD, MHPE

Department of Educational Psychology, College of Education/University of Illinois at Chicago/Chicago, Illinois

Professionalism: Opinions of Key Stakeholders

Steve Crossman, MD

Fairfax Family Medicine Residency Program/Virginia Commonwealth University/Fairfax, Virginia

**The Impact of Race on Student Well-Being:
A Multi-Institutional Study**

Liselotte Dyrbye, MD

*Department of Internal Medicine/Mayo Clinic/Rochester,
Minnesota*

11:05 – 11:20

BREAK

11:20 – 12:05

**SESSION 6: YEAR IN REVIEW IN HEALTH PROFESSIONS EDUCATION
LITERATURE**

Melchor Sánchez-Mendiola, MD

Universidad Nacional Autónoma de México/Mexico City, Mexico

12:05 – 12:15

CLOSING REMARKS

John Norcini, PhD

*President and Chief Executive Officer, Foundation for
Advancement of International Medical Education and
Research/Philadelphia, Pennsylvania*

Jeanne K. Heard, MD, PhD

*Director of the Department of Accreditation Committees at the
Accreditation Council for Graduate Medical Education/Chicago,
Illinois*

12:15 – 12:20

RECOGNITION OF EXCELLENCE
Best Conference Presentation Award

12:20

CONFERENCE ENDS

1:30

IRB TRAINING SESSION (*optional-Room 106*)

∞ MHPE Graduate Recognition Ceremony ∞

Thursday, August 2, 2007 at 5:15 pm

Welcome

Ilene B. Harris, PhD; Professor and Director of Graduate Studies, Department of Medical Education, University of Illinois at Chicago

Remarks

Ara Tekian, PhD, MHPE; Conference Chairperson, Department of Medical Education, University of Illinois at Chicago

Leslie J. Sandlow, MD; Head, Department of Medical Education; Senior Associate Dean, Educational Affairs, College of Medicine, University of Illinois at Chicago

Joseph A. Flaherty, MD; Dean, College of Medicine, University of Illinois at Chicago

Clark Hulse, PhD; Dean, Graduate College, University of Illinois at Chicago

Presentation of the Graduates and their Advisors

Ilene B. Harris, PhD

Presentation of the 2007 Best MHPE Thesis Award

Georges Bordage, MD, PhD; Professor, Department of Medical Education, University of Illinois at Chicago

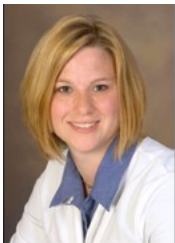
Closing Remarks

Ilene B. Harris, PhD

Graduates and Advisors



R J Gillespie (Summer 2007): Validity and Reliability of Patient Surveys in Assessing Pediatric Residents in an Outpatient Clinic
✧Advisors: P. Bashook, R. Yudkowsky, J. Daren



Kristi J. H. Grall (Spring 2007): What Instructional Practices Are Used by Emergency Medicine Physicians Viewed as Excellent Teachers?
✧Advisors: I. Harris, M. Gelula, D. Simpson



Roy Phitayakorn (Spring 2007): To Call or Not To Call? Phone Communication between General Surgery Residents and Attending Physicians
✧Advisors: R. Yudkowsky, I. Harris, R. Williams



Robert Sedlack (Spring 2007): Validation of a Colonoscopy Simulation Model for Skills Assessment
✧Advisors: A. Schwartz, S. Downing, J. Koler



Gordon Woods (Spring 2007): Development and Implementation of an Integrated Local Faculty Development Fellowship
✧Advisors: M. Gelula, R. Yudkowsky, I. Harris

A Note on Academic Regalia and the Presentation of Stoles

Academic regalia originated in the twelfth century medieval European universities of Bologna, Oxford, Cambridge, and Paris. The academic costume that we have today developed from the long robe and hood garments worn by scholars who were primarily monks and friars. Their dress met practical needs and incorporated church and state ceremonial traditions. Later, the beautiful robes of Roman Popes and the garments of church prelates set the tradition followed by bishops and vice chancellors as they became heads of universities. Universities developed regulations dictating costume styles to distinguish their officials from doctors, from lesser clerics, and from townspeople.

The use of academic costumes in this country was limited and sporadic before the Civil War. Subsequently, a renewed interest was spurred by the growth of American universities and their graduate programs and by increased contact with European universities. Also, students wanted to wear garments that would distinguish them as graduates at their graduation ceremonies.

Academic regalia include a gown, a cap, a hood, and a stole worn over the hood. Each of these components of the regalia contributes to the identification of the degree, the institution, and the field studied. At this ceremony, we are presenting each graduate with a stole. The sky blue identifies the specialty training area of the graduate as Health Professions Education.

ABSTRACTS

Assessing the Evidence for Evidence-Based Medicine Skills: A Literature Review and Tool Box

Deirdre Jenkins, MD

*Department of Hematology and Hematologic Malignancies/University of
Calgary/Calgary, Alberta, Canada*

[Co-author: Ara Tekian, PhD, MHPE
*College of Medicine, Department of Medical Education/
University of Illinois at Chicago/Chicago, Illinois*]

Background: Despite the wide spread implementation of evidenced-based medicine (EBM) into undergraduate medical education, there is a lack of consensus on the best method(s) to assess the success of this curricular change. While several reviews have been published, none have focused on the assessment of this complex skill in undergraduate medical students. Our aim is to summarize the current literature, to critique the various methods, and to make recommendations for further areas of development and research.

Summary: Articles from MEDLINE and the Educational Resource Information Centre (ERIC) from 1966 to May 2007 were reviewed. The type of assessments can be broadly categorized into the following methods: questionnaires (satisfaction, self-assessment, and knowledge), test modules, OSCEs, peer assessment, literature review checklists, and multifaceted assessments. Most methods suffer from methodological flaws, and data on psychometric properties are infrequently reported.

Conclusions: Despite the numerous reported assessment methods for EBM skills in the undergraduate medical setting, there is a lack of consensus on the best method(s). Ironically there is only weak evidence to support any given intervention. Clearly further study with carefully designed instruments to assess the full range of EBM skills is needed.

Assessment of an Evidence-Based Medicine Course in a Mexican Medical School: A Randomized Controlled Trial

Melchor Sánchez-Mendiola, MD

Universidad Nacional Autónoma de México/Mexico City, Mexico

[Co-authors: Luis F. Kieffer-Escobar M.D., Salvador Marín-Beltrán M.D.]

Background: Evidence-Based Medicine (EBM) is considered one of the healthcare professional essential core competencies. Developing countries' medical schools have introduced EBM in their curriculum, following the recommendations of the World Federation for Medical Education (1). We report a controlled study of an EBM course in a public medical school in Mexico.

Objectives: To assess the educational effect of an EBM course on attitudes, knowledge and cognitive skills of medical students in a Mexican medical school.

Methods: The intervention was a one-semester EBM course, 14 weekly two-hour sessions in the 5th year of medical school. The study design was mixed: experimental randomized controlled trial for the 5th year class, and quasi-experimental with static-groups comparison for the 4th and 6th year groups. Outcomes were compared in the 5th year half-group that just finished the course (5th year A) with the half-group without the course (5th year B), with the 4th year students without the course and the 6th year group that took the course one year earlier. Assessment of EBM attitudes, knowledge and cognitive skills was done with two instruments: Taylor’s questionnaire, an instrument designed to assess the effectiveness of EBM teaching (2); and the summative test of our medical school’s EBM course, which was done following Downing’s 12 steps for objective test development (3). Attitude scores were compared with Kruskal-Wallis test, and knowledge scores with one-way ANOVA and Student-Newman-Keuls’ multiple-comparisons test. A *p* value of less than 0.05 was considered statistically significant.

Results:

GROUP	5 th year A (MBE+) n=48	5 th year B (MBE-) n=47	4 th year (MBE-) n=87	6 th year (MBE+) n=107	<i>P</i> values
Attitude score (Taylor’s questionnaire)	28.7 ± 2.2 *	23.9 ± 5.1	24.5 ± 5.2	26.7 ± 3.6	5 th A vs 5 th B <i>p</i> <0.001 5 th A vs 4 th <i>p</i> <0.001 5 th A vs 6 th <i>p</i> <0.05 5 th B vs 4 th NS 5 th B vs 6 th <i>p</i> <0.05 4 th vs 6 th <i>p</i> <0.05
Knowledge score (Taylor’s questionnaire)	4.2 ± 3.7	1.3 ± 3.3	1.06 ± 3.2	2.4 ± 3.8	5 th A vs 5 th B <i>p</i> <0.001 5 th A vs 4 th <i>p</i> <0.001 5 th A vs 6 th <i>p</i> <0.01 5 th B vs 4 th NS 5 th B vs 6 th <i>p</i> <0.05 4 th vs 6 th <i>p</i> <0.05
Knowledge score (summative end-of-course test)	58.5 ± 8.0	32.6 ± 6.7	30.6 ± 5.6	41.0 ± 10.9	5 th A vs 5 th B <i>p</i> <0.001 5 th A vs 4 th <i>p</i> <0.001 5 th A vs 6 th <i>p</i> <0.001 5 th B vs 4 th NS 5 th B vs 6 th <i>p</i> <0.001 4 th vs 6 th <i>p</i> <0.001

Discussion: Students that underwent the EBM course obtained higher scores than students without the course. The educational intervention had a positive effect on attitudes, knowledge and cognitive skills that persisted one year later.

Scholarly Research During Surgical Residencies: A Competency-Based Approach

Dimitri Azar, MD

*Department of Ophthalmology and Visual Sciences, Illinois Eye and Ear
Infirmary/University of Illinois at Chicago/Chicago, Illinois*

Background: Scholarly research during residency training provides several potential educational benefits to residents and residency programs and is a required activity by the Accreditation Council for Graduate Medical Education. Despite the known benefits, there is a paucity of published reports regarding models of resident research that emphasize the role of faculty leadership and mentorship, and their importance for strengthening scholarly research activities. In addition, there are no clear guidelines on how resident research committees should outline research goals, expectations, educational and assessment tools, timetables for residents, and help in defining the residency research requirements.

Objective: To develop a competency-based approach for scholarly research during residency programs through small-group practice-based learning.

Methods: Survey questions and analyses of resident publications were performed. The responses to the questionnaire were scored using a 5-point Likert scale. Input was also obtained from the Resident Education Committee, Resident Research committee, and from a committed member of the Medical Education Department to develop a proposal for competency-based instruction and assessment of scholarly research during residency.

Results: Sixty-five percent of the residents considered the scholarly research project an important aspect of residency training in Ophthalmology. Only 47% of the residents agreed that satisfying the graduation requirement is a desired benefit. The main barrier to performing scholarly research activities was dedicated time for research. Of the 49 reviewed scholarly research projects of the residents, 6 (12.24%) were published. A proposal was generated that includes small-group practice-based alternative to research, which would allow for choice, flexibility, and high academic standards to serve the purpose of familiarizing teams of residents with the procedures for preparing a substantive research proposal through a collective effort with faculty and mentors.

Discussion: In contrast to the numerous efforts aimed at restructuring graduate medical education and developing the competency-based model, there have been few attempts to modify the scholarly research requirements for surgical specialties and the methods of instruction and assessment in research during residency training. In order for residents to design, conduct, and present their research studies, they need to develop skills in literature searching, research design and methods, statistical analysis, data management, medical writing and presentation skills. Our findings suggest that there is a need to better define the scholarly research requirement for Ophthalmology and other surgical specialties. In addition, we have identified major obstacles that were encountered by Ophthalmology residents who matriculated in an academic residency program with an established tradition in research and scholarly activities. Instead of revamping the established system, we proposed a small group practice-based learning approach (incorporating workshops and dedicated research time) as a parallel alternative to the existing

system of having a rigorous research requirement. The novelty of this approach is the active involvement of highly motivated mentors in guiding the project selection, implementation, and assessment processes; the flexibility of project choice, the incorporation of outcome-based incentives; and the willingness to reassess the results of the new approach in 3-5 years.

Which Software Program for Generalizability Studies is Best? Comparing G_String II and EduG

Pamela Derstine, PhD

Accreditation Council for Graduate Medical Education/Chicago, Illinois

Background: G-String II, developed by Ralph Bloch and Geoff Norman, provides a familiar “windows”-like interface with urGENOVA, one of three classic programs for G-studies developed by Robert Brennan. EduG, developed by the Canadian firm Educan with input from Canadian and Swiss specialists, is a fully functional stand-alone program for G-studies, also with a user-friendly interface. Both programs use ANOVA results to calculate Gcoefficients, presumably employing similar formulas, except that EduG applies “Whimbey’s correction” to the estimation of the variance of fixed facets. While urGENOVA was developed for unbalanced designs, it gives the same results as GENOVA (Brennan’s first G-study program) for balanced designs.

Objectives: Compare the usability and features of each program as well as results of G- and D-studies with different measurement designs to determine which program, if any, is ‘best.’

Methods: Two datasets containing several different measurement designs, including 1-, 2-, and 3-facet fully crossed and nested with variable assumptions for fixed and random facets, were analyzed using both programs and the results were directly compared. The first dataset consisted of Brennan’s four synthetic datasets. Published study results (variance component estimates and G-coefficients from G and D studies), available in Brennan’s textbook, *Generalizability Theory*, were used to verify results obtained using G-String II and EduG. The second dataset, derived from the competency-based outcomes assessment of pediatric residents via an online portfolio system over 2 ½ years, were used to verify findings obtained using the synthetic datasets and to test the unique features of each program.

Results: G-String II and EduG gave the same results for variance component estimates and G and D studies for all balanced designs except crossed designs with 2 or more facets in a mixed model. For such designs, the results obtained using EduG, but not G-String II, were the same as those published in the Brennan textbook. The formula used by G-String II for such designs contains an error, likely due to the fact that it is optimized for unbalanced designs, and the D study feature is an add-on (urGENOVA does not have this feature). When using G-String II for balanced nested designs, a similar anomaly occurs but can be corrected by adjusting the levels entered for D studies where the crossed facet was fixed. When the nested facet is fixed, the two programs give different results, because they each use a slightly different formula. However, because the GT textbook contained no results for such studies, it was not possible to independently verify the results. While only G-String II can be used for unbalanced designs, EduG has several extra features that were found to be very useful in exploring data to find

suitable subsets for further analysis. These include G-facets analysis (identify which level(s) of a facet might be eliminated to give higher G-coefficients), data reduction (allows easy removal of one or more facet levels and subsequent complete analysis), and optimization (allows up to 5 simultaneous D studies with results presented side-by-side). EduG also calculates both Grel and Gabs, as well as $\Phi(\lambda)$, while G-String II calculates only Gabs.

Discussion: Both programs are very easy to use and quite stable in the Windows XP environment used for this study. EduG is the recommended program for all balanced designs because of the many additional features it offers for manipulating the data; it cannot be used for unbalanced designs. The use of “Whimbey’s correction” for fixed facets does not affect the calculation of G coefficients because the estimates from the random model are used, not the corrected estimates. G-String II should be used for all unbalanced designs; it can be used for all balanced designs except for crossed, mixed model designs. It does not calculate Grel or $\Phi(\lambda)$. No recommendation can be made for which program to use for nested, mixed model designs where the nested facet is fixed.

Are Nurse, Peer, and Faculty Performance Evaluations of Surgical Residents Reliable? A Generalizability Study

Pamela A. Lipsett, MD, FACS, FCCM

Department of Surgery, School of Medicine /Johns Hopkins University/Baltimore, Maryland

Background: The reliability of ward evaluations (WE) has been questioned and variation in attending evaluations may be based on a number of factors. To address the concern about reliability of these evaluations, and to meet the requirements of a 360 degree evaluation, educators have explored the use of alternative raters. The additional contribution of these assessments has been questioned, as well as their reliability.

Objectives: Given that different raters may have different perspectives on resident performance and thus decrease reliability when combined, we sought to examine the reliability of nurse, peer, and attending WE by using generalizability methodology.

Methods: The ratings for all residents from July- December 2006 were obtained. Items common to all professional groups were in four domains: interpersonal skills, (interactions with colleagues and staff), professionalism (reliability), clinical judgment, and a global scale of overall competence were abstracted. This completed data set included eight PGY2 residents, (the object of measurement) and a total of 12 raters (nurses (2), peers (5) and attending physicians (5)) over four items. This allowed for a fully crossed and random design (subject (s) x items (i) x raters (r)). In order to look at the contributions of individual pair groups and because the number of facets was limited to 2 (item and rater), serial G and D studies were performed using different groups of raters; nurses and peers, peers and attending physicians, nurses and attendings, and all three (nurses, peers, and attendings) as raters.

Results: For the three groups combined, the greatest source of error was the residual error from SIR. The universe score variance representing the differences in resident performance resulted in

a 6.9% variance. Some items were more severely rated than others; accounting for a variance of 12.2%. Raters did not demonstrate a great deal of variation in their leniency (2.1%); 10.4% of the total variability was accounted for by the interaction terms si (4.5%) and ir (5.9%). For the G study design, where there were 8 subjects, 12 raters, and 4 items ($s \times i \times r$, random balanced) the Phi coefficient was 0.541. In order to achieve a Phi coefficient >0.8 , 10 items would be required by 40 raters.

Discussion: The combination of raters from different specialties adds complexity in determining the true universe score. These studies suggest that our WE should be changed so that the items asked in common on the evaluations are greater than the four used. The D studies suggest that this number should be increased to 10 and raters to 40 if the WE is to be considered to have optimal reliability. This recommendation assumes that the additional items and raters will perform in a similar fashion to those currently used. In addition, the studies are simplified in that they consider a single rotation on a surgical service, with a single level of resident training, a specific time, and does not consider the gender of the resident. More complex generalizability studies could be done to answer the contribution of these additional variables.

What Qualities Do Medical Students Like to See in Their Clinical Teachers as Role Models?

Zaka Khan, MD

Family Medicine/Baqai Medical University/Pakistan

Background: Medical students tend to incorporate a number of behaviors from observation of their teachers in clinical settings. These behaviors, attitudes, and beliefs are very important in subsequent transformation of the medical students to effective medical educators and/or health care professionals. Further, these are vital for their character building, interaction with their patients and colleagues, and for leadership.

Purpose: The purpose of this study was to identify the qualities of an effective teacher in clinical sciences, as indicated by the medical students, which could be later incorporated in their curriculum. This would in turn ensure that their training would be based on those qualities.

Method: A questionnaire, comprised of both the open and closed ended questions was developed by the Center for Medical Education in collaboration with the Department of Medicine at Baqai Medical College, Pakistan during the year 2004. It was administered to a group of 100 students in the 3rd, 4th, and 5th years of the MBBS course. This questionnaire had different questions seeking opinion from the students about what different qualities they would like to see in their teachers in clinical sciences, which they feel would make them role models in clinical medicine, in an order of preference.

Results: The response rate of the questionnaire was 90%. Analysis revealed that some of the important qualities indicated by the students, in order of preference were 'Effective Communicator' (62%), 'Competent' (40%), 'Punctual' (36%) and 'Respectful to the patients & colleagues' (19%).

Discussion: Different studies have been done in the past on incorporation of different attributes in the curriculum for training the medical students in important areas like doctor-patient communication. All have focused on building up these attributes during the training of the students in the medical school during different phases. However, the most important point to focus here is that students learn from their teachers. How the teachers think, act, and behave in their professional lives is vital, because their students adapt and behave accordingly. All these parts of ‘hidden curriculum’ are to be explored, in order to incorporate these experiences for the learners.

Teacher’s Training Workshop on Teaching-Learning Methodology and Evaluation in Four Medical Colleges of Nepal

Nirmal Baral, MD

Department of Health Professions Education, Department of Biochemistry and Molecular Biology/B.P. Koirala Institute of Health Sciences/Ghopa, Dharan, Nepal

Background: Effective teaching is a concern of all teachers. Therefore, regular teachers’ training is emphasized globally. B. P. Koirala Institute of Health Sciences (BPKIHS), a health science university situated in the eastern region of Nepal has an established Medical Education unit which attempts to improve teaching-learning skills by training faculty members through organizing regular medical education training programs.

Objectives: The aim of the present study was to assess the effectiveness of a 3-day training workshop on “Teaching-learning methodology and Evaluation” held in four different medical colleges of Nepal.

Methods: The workshop was targeted at middle and entry levels of health profession teachers who had not been previously exposed to any teacher’s training program. The various components, such as teaching-learning principles, writing educational objectives, organizing and sequencing education materials, teaching-learning methods, microteaching and assessment techniques, were incorporated in the workshop. A team of resource persons from BPKIHS were involved in all the four medical institutions. The collection data had two categories of responses: (1) a questionnaire survey of participants at the beginning and end of the workshop to determine their gain in knowledge and (2) a semi-structured questionnaire survey of participants at the end of the workshop to evaluate their perception of the usefulness of the workshop. The latter category had items with a three-point likert scale (very useful, useful and not useful) and responses to open-ended questions/statements to document participants’ general views. The response was entered into a spreadsheet and analyzed using SPSS.

Result: The result showed that all participants (n = 92) improved their scores after attending the workshop (p < 0.001). The majority of respondents expressed that the teaching-learning methods, media, microteaching and evaluation techniques were useful in teaching-learning. The workshop was perceived as an acceptable way of acquiring teaching-learning skills, but 39.4% participants expressed that the duration of the workshop was too short. The overall impression of trainers was very positive.

Conclusion: Therefore, regular organization of such workshops with the addition of new advances in medical education would be highly beneficial to improve teaching-learning skills of medical teachers.

Improving Clinical Skills in the Emergency Department: Impact of Curricular Change at the Ankara University Medical School

Ayşe Gürel, MD

Ankara University Medical School/Ankara, Turkey

Background: To address the pressing need to improve the competencies of their graduates, Ankara University Medical School has initiated significant curricular revisions. Theory and available evidence suggest that implementing integrated courses in place of discipline-based courses and incorporating instructional features like problem-based learning (PBL) or clinical presentation model (scheme-based learning, SBL), promotes problem-solving skills. This study focuses on the cumulative effects of curriculum integration, PBL, and SBL on the diagnostic performance of students in an actual clinical setting, the Emergency Department.

Objectives: This project aims to evaluate problem-solving skills of students taught in two different curricular programs by assessing their diagnostic accuracy. The study also aims to assess the impact of structured skills education on students' procedural skills.

Methods:

Diagnostic Accuracy

The diagnostic accuracy of the interns from both groups will be assessed by using a data collection sheet developed by Sklar and colleagues⁵³ (Sklar, 1991) to compare the diagnostic processes of experienced emergency physicians with those of novices. The data sheet will be used to identify the diagnostic impressions and the students' degree of certainty. The students will be asked to fill in their initial leading diagnoses after examining the patient. Then they will be asked to estimate their level of certainty that each diagnosis would turn out to be the correct final diagnosis by marking a visual analog scale⁵⁴ (Aitken, 1969) between the extremes of "no chance" and "absolute certainty." The distance from the beginning of the line to the mark will be measured and the ratio of measured length to the total length of the line will be calculated to give a certainty score between 0 and 1. Students will then be asked to select three sets of tests that they think are appropriate. When the results of one test will become available, they will select another test. The term "tests" will include history and physical as well as ECG, radiography, and laboratory analysis. After the results of all tests will become available, the students will formulate a final diagnostic impression and estimate a level of certainty by marking along a visual analog scale as detailed above.

Level of Practical Experience and Self-confidence in Practical Skills

The Practical Skills Survey Form will be used to assess both variables. This survey form is developed by the modification and integration of two different forms utilized in two different studies (Board and Mercer, 1998) (Moercke, 2002). The students will be asked to rate their skill level and then to give an approximation of the number of times they have successfully performed

each practical skill. It is hoped that introducing two related concepts adjacently will help students rate their perceived level of competence in a realistic fashion as reflected by their level of practical experience.

Questions for Consultation:

How can I improve my instruments to test for the effects of PBL? SBL?

What analyses should I consider?

What Instructional Practices are Used by Emergency Medicine Physicians Viewed as Excellent Teachers?

Kristi Grall, MD

Department of Emergency Medicine/University of Arizona/Tucson, Arizona

Background: There is a pressing need to identify efficient and effective teaching methods in emergency medicine settings. Both overcrowding and demands for clinical productivity have been identified as causes of decreased time available for clinical teaching. (Atzema 2005) These causes are related, in that emergency department overcrowding leads to increasing pressures to be clinically productive. As more and more patients present to an emergency department, maintaining patient flow and taking care of patients takes up increasingly more of a faculty member's time, diminishing the time available for education. This situation is approaching critical levels in emergency departments, with respect to time available for clinical teaching. (Skeff 1997, Atzema 2005)

Objectives:

Two main questions were the focus of this research study:

1. To what extent are instructional methods described as efficient and effective in the literature on ambulatory care education also used by faculty in the emergency department, who are identified by peers and residents as excellent teachers?
2. What additional methods are used by these excellent emergency medicine teachers?

Methods: A study of instructional practices used by excellent emergency medicine teachers was carried out using a mixed method, but primarily qualitative, non-participant observational study followed by semi-structured interviews and a focus group. Four excellent emergency medicine teachers were identified by their peers and residents for participation. A total of 32 hours of observation of instructional practices was conducted. Data recorded included teaching methods used, contextual descriptions, learner level of training, number of emergency department beds occupied, patient's level of illness severity, and assessment of learner's stage in the Kolb learning cycle.

The observations were followed by a one-hour semi-structured interview session with each of the four participants to identify the extent to which they agreed with our observations, and also any additional strategies they use that they believe increases the effectiveness and efficiency of their

teaching. Each participant was also asked to describe the attributes of an excellent teacher and barriers to teaching in the emergency department.

Following the interviews, a two-hour focus group comprised of two of the four study participants was held, to continue the discussion regarding the teaching strategies they have used in the emergency department, as well as ideal teaching encounters, barriers to teaching in the emergency department, and alterations they make in their teaching to accommodate for resident experience level.

Results: The top four teaching methods observed were Questioning (used in 44% of teaching encounters), Advice-Giving (30%), Limited Teaching Points (27%) and Patient Updates (21%). In general, the least time-consuming teaching methods, such as Advice Giving and Patient Updates, were more often used with senior residents. The participants did not alter their teaching methods for increased crowding of the emergency department, but used less time-intensive methods with less severely ill patients. Barriers to teaching in the emergency department included lack of time to teach, distractions, patient care responsibilities, and lack of faculty support for teaching.

Discussion: This study was conducted to determine what teaching methods are used by emergency medicine faculty identified by their residents and teaching colleagues as excellent teachers. The teaching methods used most frequently by the four participants in this observation study and confirmed by the interviews and focus group included questioning, advice giving, limiting teaching points, and patient updates. Two of these have been documented in previous ambulatory care and emergency medicine teaching literature – questioning and limiting teaching points. (Heidenreich 2000, Bandiera 2005). The other two, advice giving and patient updates, were observed frequently in this observation study of excellent emergency medicine teachers, but have not been documented in the ambulatory care or previous emergency medicine literature. These four methods are also the most versatile, having been used by the majority of the participants under several emergency department conditions, including all circumstances of resident training level, department crowding level, and patient triage severity. All four of these methods require very little time to implement, and it is likely that this contributes to their popularity in the high-pressure emergency department environment.

Efficiency in teaching is needed in the face of an increasingly crowded, busy emergency department environment. Studying the instructional practices of emergency physicians viewed as excellent teachers gives insight into how these individuals continue to teach effectively despite the changes in the ER environment. Concentrating on the top four teaching methods – questioning, advice-giving, limited teaching points, and patient updates – using teaching methods geared toward the resident's level of experience, overcoming barriers to teaching, and providing faculty development to improve teaching skills of emergency medicine teachers may be the key to teaching in a highly demanding, high-pressure teaching environment such as the emergency department.

Resident Assistants Can Successfully Reduce Duty Hour Violations While Improving Quality of House Officer Inpatient Educational Experiences

Hilary Haftel, MD, MHPE

Pediatrics and Communicable Diseases/University of Michigan/Ann Arbor, Michigan

Background: Residency training requires a balance of education and service commitments. Duty Hour Regulations required significant changes to residency training to maintain the quality of patient care despite reduced resident workload, and still maintain an appropriate educational experience. We instituted a Resident Assistant (RA) program by hiring one clerical-level person per inpatient team responsible for all clerical tasks that residents had previously performed.

Objective: This study analyzed the effect of RAs on Duty Hour compliance and on the overall quality of inpatient rotations.

Design/Methods: Duty hour logs for Pediatric Residents on inpatient services prior to and following the initiation of the RA Program were analyzed for duty hour violations. Resident evaluations of inpatient rotations prior to and following the RA program were compared. Descriptive statistics and t-tests were employed for data analysis.

Results: There was a significant reduction in the number of duty hour violations subsequent to the RA program on all services, particularly regarding total work hours and violations of the 24+6 hours rule. Reductions ranged from 50-90%, compared to <15% decrease on non-RA services. t-tests demonstrated significantly improved quality of inpatient rotations on services that added an RA as assessed by rotation evaluations.

Conclusions: Resident Assistants can be successfully used to protect the time residents spend in the hospital and allow residents to conform to duty hour regulations. Inpatient rotations that include an RA were evaluated as better educational experiences than rotations without RAs. This model may be successfully applied to other inpatient services and other hospital programs.

To Call or Not To Call? Phone Communication Between General Surgery Residents and Attending Surgeons

Roy Phitayakorn, MD

Department of Surgery/Case Western Reserve University Medical Center/Cleveland, Ohio

[Co-authors: Williams RG PhD², Yudkowsky R MD MHPE³, Harris IB PhD³, Mellinger JD MD⁴, Hauge LS PhD⁵, Sullivan ME PhD MSN⁶, and Widmann WD MD⁷]

²*Department of Surgery, Southern Illinois University, Springfield, IL.*

³*Department of Medical Education, University of Illinois at Chicago, Chicago, IL.*

⁴*Department of Surgery, Medical College of Georgia, Augusta, GA.*

⁵*Department of Surgery, Rush University Medical Center, Chicago, IL.*

⁶*Department of Surgery, University of Southern California, Los Angeles, CA.*

⁷*Department of Surgery, Columbia University, New York, NY.*

This research was supported in part by a grant from the Center for Excellence in Surgical Education, Research, and Training (CESERT) of the Association for Surgical Education Foundation and from United States Surgical.

Background: Communication between general surgery residents and attending surgeons, in the context of residents' ongoing responsibilities for patient care is a critical, but hitherto underresearched component of patient care. Miscommunication between these groups of healthcare providers likely leads to medical errors and delays in medical treatment.

Objective: The purpose of this study was to identify and characterize various factors that may impact on a resident's decision whether to call the attending, or an attending's desire for a phone call, about hypothetical patient problems.

Methods: A research instrument consisting of 12 patient care scenarios was developed, of three types: issues related to professionalism, to resident errors, and to changes in patient status. Residents were asked to assume they were the chief resident and decide whether they would call the attending surgeon about each scenario and why. Attending surgeons were asked if they would want to be called by the chief resident and why. Scenarios occurred at either 1700 or 0200 hours. The survey was distributed at seven academic general surgery programs in the United States. Quantitative data was analyzed using Fisher's Exact Chi-square test. A grounded theory approach was used to analyze qualitative data.

Results: The survey return rate was 61% for residents (n=149) and 36% for attending surgeons (n=116). Residents and attending surgeons agreed about the need to call an attending in seven scenarios. There was disagreement in the other five scenarios: attendings desired a call in situations where a significant number of residents would not call. These differences were greatest in scenarios concerning resident errors and were independent of the scenario time of day. Residents would call an attending surgeon if they believed the patient might require an emergency operation, to provide the attending with information, or to ask for an attending's guidance. Residents would not call an attending if they thought the patient was medically stable or appropriately managed. Attending surgeons wanted a call from the resident because they are responsible for the patient's care, to direct the patient's care, and/or to correct the resident error. There were no differences between the phone call patterns of junior (PGY1-2) and senior (PGY3-5) general surgery residents. Only one scenario was significantly different between junior (<10 years) and senior (≥ 10 years) attendings. One-third of the scenarios were different between residents at different institutions, whereas there was no difference between attendings at different medical centers.

Discussion: Residents and attending surgeons generally agree about the need for a phone call when admitting a patient or creating a patient care plan, possibly due to socialization or a hidden curriculum early in medical training. However, residents and attendings disagreed about the need for a phone call in scenarios about resident errors or acute changes in patient status. Importantly, the time of day featured in the scenario is not an important factor for these differences. Successful initiation of communication requires agreement about a series of questions that can be loosely grouped into three different categories including patient factors, attending factors, and resident factors. Several solutions are proposed to begin creating a culture

of clear communication and detailed analyses of miscommunication near-misses and errors to develop guidelines for standards of practice and to resident education.

Fellowship in Secondary Hospital Medicine – Curriculum Design and Program Evaluation

Rashmi Vyas, MD

Department of Physiology/Christian Medical College/Tamil Nadu, India

Background: Christian Medical College (CMC), Vellore, India, is a national level training institution, owned by 64 churches which run 200 secondary hospitals. Secondary hospitals have between 20 and 200 beds, with a generalist practice of medicine, surgery and obstetrics; limited laboratory support such as complete blood count, routine urine and stool tests, clinical biochemistry and X-rays; and local community-based programs such as malaria and tuberculosis. CMC is a tertiary care teaching hospital where most of the CMC graduates have a service obligation following graduation for two years at the secondary hospitals. These hospitals are unable to retain the medical graduates once they finish their service-obligation. This is mainly because the graduates are not adequately prepared to practice in secondary hospitals and face academic isolation, as they do not receive adequate support and supervision for practice and learning during their service obligation. The secondary hospitals are facing a crisis, as many of them are closing down because of lack of staff. This situation is affecting the health care of the nation, especially in rural areas. A program, Fellowship in Secondary Hospital Medicine (FSHM), is being designed to prepare CMC graduates to practice effectively in secondary hospitals.

Objectives: The purpose of this project is to develop, implement, and evaluate the processes and outcomes of a fellowship curriculum in secondary hospitals for CMC graduates. The aim of the program is to provide CMC graduates with the knowledge, skills, and attitudes to practice secondary hospital medicine and the ability to apply their learning into practice, following which they may choose to practice in secondary hospitals.

Methods: The curriculum design for the FSHM is based on a combination of traditional systems approach (with a focus on coherence among educational goals, instructional methods and performance assessment), deliberative inquiry (with a focus on group processes of deliberation in curriculum design, implementation and evaluation) and reconceptualist inquiry (with a focus on issues related to the “hidden curriculum”). Literature review, informal interviews, brainstorming sessions and surveys of graduates and faculty who have worked in secondary hospitals contribute to the needs assessment. Program evaluation will be done with objectives-oriented and participant-oriented approaches, using the Logic Model framework for evaluation.

Results of Needs Assessment: Based on the needs assessment, it is proposed that the curriculum should be of one year duration. The instructional methods will include: distance learning, development of skills in on-site sessions, and project work. Student assessment will be both formative (feedback for distance learning activities, project proposals) and summative (tutor-marked assignments, final exam, and project report). Faculty development workshops, networking and student support, and a partnership between CMC and secondary hospital faculty

are vital for the success of this program. The program must be evaluated to make ongoing changes for improvement and decisions about its value and effectiveness.

Discussion: The purpose of this program is to help CMC graduates make a transition from learning in a tertiary hospital environment to practicing at a secondary-level hospital. Hopefully this successful transition will stimulate graduates to view practicing at secondary hospitals as challenging and interesting. An anticipated long-term outcome is that graduates will make it their career choice, thus benefiting the hospitals and the community.

Medical Student Self-Assessment vs. Standardized Patient Ratings

Gail Beeman, MD, MHPE

Robert J. Kaplan Clinical Skills Center/University of Tennessee, Health Science Center/Memphis, Tennessee

Background: The ability to self-assess is important to maximize life-long learning that is key to excellent patient care. Studies of medical students show that self-assessment of performance based skills are relatively fixed over time and may vary with the type of task assessed. Studies also show that demographic and academic variables do not relate to the ability to accurately self-assess.

At our institution during the introduction to clinical medicine course, students receive written feedback regarding the patient's perceptions of the student during the encounter. Prior to receiving the feedback, the students are required to fill out a written self-assessment.

Objectives: To determine the relationship between the standardized patient's rating of each student on four perceptual items to the student's self-assessment of the same items.

Methods: Standardized patients rated the students on the following items: 1) engage with the patient, 2) technique of communication, 3) expression of empathy, and 4) completeness of tasks. For each item, the patient marked along a 15 centimeter line with "poor", "satisfactory", and "superlative" spaced evenly from left to right. The students were required to self-assess in the same manner prior to receiving the patient feedback. Ratings from both groups were measured and recorded as percentages. Differences in the percentages of patient and student ratings were calculated for each item.

Results: Ratings from standardized patients and self-assessments from students for 149 patient encounters were analyzed. For the item 1) engage with the patient, 30 (20.1%) students' ratings showed a $\geq 30\%$ difference between the standardized patient's rating and the student's self-assessment. The students' self-assessments were higher in all 30 (100%). For item 2) technique of communication, 31 (20.8%) sets of ratings showed a $\geq 30\%$ difference. The students' self-assessments were higher in 27 (87.1%). For item 3) expression of empathy, 33 (22.1%) sets of ratings showed a $\geq 30\%$ difference. The students' self-assessments were higher in 32 (97.0%). For item 4) completeness of tasks, 31 (20.8%) sets of ratings showed a $\geq 30\%$ difference. The student's ratings were higher in only 18 (58.1%). Seven students' (4.7%) self-assessments were

≥30% higher for all four items. Three students' (2.0%) self-assessments had differences that were ≥30% greater but mixed higher and lower.

Discussion: Methods? Given that patients in the “real world” rate physicians by subjective measures such as these items as well as performance-oriented skills, is it worthwhile revealing to the young medical learner that he/she is over-assessing certain aspects of the patient encounter? Does required self-assessment improve the feedback process from patients especially if there is discord between the self-assessment and the feedback? Longitudinal data analysis?

Carrying Portfolios Forward – Can Portfolios be Strengthened for Use as Summative Assessment?

Kevin Van Kanegan, DDS

College of Dentistry/University of Illinois at Chicago/Chicago, Illinois

Background: Portfolios are increasingly being implemented at health professions schools as a means to evaluate students' competency. The evidence to support their use for formative purposes, especially in regards to professionalism, with particular emphasis on reflection leading to self-evaluation, is quite strong and well-documented. The evidence available to endorse their use to make summative judgments is still not as strong. It has been suggested that the use of qualitative assessment methods, using credibility and dependability as proxies for validity and reliability, may be a better and more “organic” means to evaluate portfolios as summative assessment.

Objectives: Apply qualitative assessment methods to the portfolios compiled by the undergraduate dental students at the University of Illinois at Chicago, College of Dentistry to test their use as a credible and dependable method of summative assessment.

Methods: Evaluate the current use of portfolios using methods of qualitative assessment and statistically compare portfolios to other instruments currently being used at the college of dentistry that also measure the competencies of reflection/self-assessment.

Results: Pending

Discussion: The reality is that portfolios seem to be an assessment tool that is here to stay. This is certainly the case at the University of Illinois, College of Dentistry, where their value within a competency-based, outcome-measured curriculum is potentially rich. This, combined with strong administrative support for innovative methods, and we are left little choice but to keep polishing the diamond, even if it is still a bit lumpy and black. It is my hope that by applying more research pressure on this area, we may find out if portfolios really are a gem, or merely another fuel for which to seek alternatives.

Description and Impact of Using a Standard-Setting Method for Determining Pass/Fail Scores in a Surgery Clerkship

Nancy Schindler, MD

Feinberg School of Medicine/Northwestern University/Chicago, Illinois

Background: Research on performance evaluation highlights the importance of using multiple measures to develop an accurate profile of students, yet we found no literature describing the use of a standard-setting method for determining a pass/fail cutoff for a clerkship based on multiple assessment methods.

Method: Steps in setting an absolute standard for a pass/fail grade are described. The new cut-off score was used to compare what decisions would have been made had it been applied in previous clerkships.

Results: We successfully applied the Hofstee method to ascertain a new standard pass/fail cutoff for our total surgery clerkship score. Had this absolute score been used in 4 prior clerkships, 150 instead of 152 would have passed the clerkship, and 10 instead of 8 would have failed the clerkship.

Conclusions: A standard-setting method can be applied to a final clinical clerkship grade even when multiple performance measures are used.

A Policy-Capturing Study Examining the Factors that Influence Image Quality in Plastic Surgery Education

Fred Watkins, MD, MHPE, FRCSC

Division of Plastic Surgery/Hotel Dieu Hospital/Kingston, Ontario, Canada

Background: Clinical images are ubiquitous in Medical Education and it is axiomatic that only the highest quality images are used in teaching and assessment. A poor quality image, when used for instruction will only confuse the learner and when used for assessment may invalidate the test item. Despite the obvious importance of using only high quality images, there is very little written on this subject in the context of medical education.

Objectives: The purpose of this project was to investigate the factors that contribute to image quality in medical education.

Methods: The factors influencing image quality were explored using a policy-capturing design in the specific setting of a high stakes specialty qualifying examination. Seventy Board Certified Canadian Plastic Surgeons participated in the study.

Results: In the setting of this project:

- High quality images generated consensus amongst the participants.
- Lesser quality images resulted in widely divergent ratings amongst the participants

- Four factors studied were found to significantly influence image quality. These factors are (in descending order of influence):
 - 1) Accuracy of Reproduction
 - 2) Relevance (consistent with the objectives for the field of study)
 - 3) Degree of Prototypia
 - 4) Quality of Composition

Discussion: In the setting studied, peer review was not a reliable way to identify the lesser quality images and therefore a set of guidelines might be useful to assist image selection in medical education. The factors identified by this project as important in the image selection process have been incorporated into a preliminary set of guidelines.

Improving the Rating Scale for the Communication Skills OSCE

Cherdsak Iramaneerat, MD, MHPE

Department of Educational Psychology, College of Education/University of Illinois at Chicago/Chicago, Illinois

[Co-author: Rachel Yudkowsky, MD, MHPE

*College of Medicine, Department of Medical Education/
University of Illinois at Chicago/Chicago, Illinois]*

Background: Interpersonal and communication skills are one of the six core competencies that the ACGME mandates residency programs to assess in their residents. The University of Illinois at Chicago assessed communication skills of residents from many departments using an OSCE that employed an 18-item rating instrument. SPs rated residents' performance on this rating instrument using a 5-category rating scale. However, the quality of ratings that SPs provided on this rating scale has never been scrutinized.

Objectives: The purposes of this study were (1) to evaluate how well SPs used this rating scale, (2) to find ways to improve the quality of this rating scale, and (3) to determine whether the modifications of the rating scale help improve the quality of ratings.

Methods: We carried out a retrospective analysis of the ratings obtained from the communication skills OSCE for 68 internal medicine residents in 2003. Each resident encountered six SPs in six communication challenges. We analyzed the data using a many-faceted Rasch measurement (Facets) model – a probabilistic model that determines the likelihood of any given resident receiving a particular rating as a function of the communication competence of the resident, the difficulty of the item, the severity of the SP, and the difficulty of the case. We examined the distribution of ratings and category characteristic curves – plots of the probability that residents would receive various ratings as a function of residents' communication competence – of each item on the rating scale to determine the effectiveness of the scale. We modified the rating scale based on our findings. We then implemented the revised rating scale in the 2007 communication skills OSCE of 85 internal medicine residents, where each resident encountered six SPs in six communication challenges. We analyzed the data using a Facets model and checked the distribution of ratings and category characteristic curves of each

item on the revised rating scale to see whether our modifications help improve the quality of ratings.

Results and Discussion: The analysis of the 2003 communication skills OSCE ratings revealed that the rating scale functioned as a 4-category rating scale. The lowest rating category was rarely used. Eighty percent of ratings were confined to the two highest rating categories. In most items, residents who received higher ratings were not more competent than those with lower ratings. Six items had rating scales that did not fit the measurement model. Two items exhibited too unpredictable ratings. On the other hand, three items had redundant ratings.

We made several modifications of the rating instrument, including (1) combining redundant items; (2) separating an item that contained two aspects of communication into two items; (3) shortening the item description, but providing more detailed description on rating options; (4) adopting a 4-category rating scale; (5) providing unique behavior-specific descriptions of rating options for individual items; and (6) increasing the difficulty of items by demanding higher level of performance for high ratings. These modifications led to a new 13-item rating instrument called the *Revised UIC Communication and Interpersonal Skills Scale*.

The analysis of the 2007 communication skills OSCE ratings obtained from the revised scale revealed the improvement in rating scale functioning in several aspects. All items had rating scale structures that fit with the measurement model. Residents with higher ratings were more competent than those with lower ratings in all items. Only one item still had too unpredictable ratings and one item had too redundant ratings. The items appeared more difficult as intended. However, there were still a couple of unresolved problems. The lowest rating category was still underutilized in most items. The ratings in several items still had unbalanced rating distribution.

Professionalism: Opinions of Key Stakeholders

Steve Crossman, MD

*Fairfax Family Medicine Residency Program/Virginia Commonwealth
University/Fairfax, Virginia*

Background: In recent years, professionalism has become an increasingly important topic in health professions education and in the health professions. Much has been written, and many national and international standards have been established. A better understanding of the various perspectives of key stakeholders – nurses, resident physicians, faculty physicians, and patients – at a single institution would be helpful in determining what exactly to teach, how best to teach it, and how best to evaluate this important aspect of being a health professional. To our knowledge, this type of study has not yet been conducted.

Objectives: The objective of our work was to understand better how key stakeholders in a single institution define professionalism and how they believe it should be taught and assessed.

Methods: A survey with 5 open-ended questions was created and administered to a convenience sample of nurses, resident physicians, faculty physicians, and patients at a single family medicine residency ambulatory clinic. Qualitative methods were used to analyze the data.

Results: While there were some areas of disagreement, all four stakeholder groups did concur overall on a general definition of professionalism and this definition was consistent with published standards (except that the study groups did not include broader social issues in their definition). There was also general agreement on suggested methods for teaching and assessing professionalism.

Discussion: The results from this study could prove very helpful in designing and evaluating professionalism training programs for health care professionals.

The Impact of Race on Student Well-Being: A Multi-Institutional Study

Liselotte Dyrbye, MD

Department of Internal Medicine/Mayo Clinic/Rochester, Minnesota

[Co-authors: Matthew R. Thomas,¹ Anne Eacker,² William Harper,³
F. Stanford Massie Jr.,⁴ David V. Power,⁵ Mashele Huschka,¹ Paul J. Novotny,¹ Jeff A. Sloan,¹ and Tait D.
Shanafelt¹

¹Mayo Clinic Department of Internal Medicine; ²University of Washington School of Medicine; ³University of Chicago Pritzker School of Medicine; ⁴University of Alabama at Birmingham School of Medicine; ⁵University of Minnesota Medical School.

Purpose: Little is known about the training experience of minority medical students. In this study, the authors explore differences in the prevalence of burnout, symptoms of depression, and quality of life (QOL) among minority and non-minority medical students.

Method: Students (n= 3,080) at 5 medical schools were surveyed in the spring of 2006 with the Maslach's Burnout Inventory, PRIME MD, and SF-8 QOL. Students were also asked about the influence of their race on their training experience. Differences by race were tested using chi-square tests for proportions and Kruskal-Wallis tests for continuous variables. All tests were two-sided tests with type I error rates of .05.

Results: A total of 1,701 (response rate 55%) medical students completed the survey. Symptoms of burnout and depression were present in 46.5% and 48.5% of respondents, respectively. Mental QOL was lower among medical students than the general population (p<.0001). The prevalence of depression was similar regardless of minority status but a higher prevalence of burnout was observed among non-minority students (p<0.03). Minority students were more likely than non-minority students to report that their race/ethnicity had adversely affected their medical school experience (p<0.0001). Minority students reporting an adverse effect of race were more likely to be burned out (p=.001) and have symptoms of depression (p=.004) and low mental QOL (p=.0001).

Conclusions: Symptoms of distress are prevalent among medical students. Specific adverse experiences related to race may contribute to the distress of minority students. Additional studies are needed to define the causes of these perceptions and improve the learning climate for all students.