Session CUD

Update in Medical Education
SGIM 31st Annual Meeting
Pittsburgh, PA

April 11, 2008

Carol K. Bates, MD
Judith L. Bowen, MD
Shobhina G. Chheda, MD, MPH
Kathel Dunn, MSLS
Reena Karani, MD
On Continuity


Funding: None indicated, though the lead author acknowledges an annual stipend for editing of online curriculum and participating programs pay a fee for the curriculum.

Summary: Residents in three institutions were surveyed on their perspectives on continuity clinic experiences. Resident ratings were highest on diversity of illness seen and on a variety of markers of preceptor quality. Residents’ value of clinic was not associated with planning a career in general internal medicine.

1. Background: Recent recommendations from APDIM and the ACP have raised concerns about continuity clinic, referring to practices as “dysfunctional” and “inadequate.”

2. Aims: To measure resident perspectives on the learning environment, patient mix, and teaching quality in continuity resident practice.

3. Methods:
   - Residents from 3 medicine training programs (Johns Hopkins Hospital, Christiana Care Health System, and Henry Ford Hospital) were surveyed electronically through the Hopkins Internet Learning Center. Each training program is a purchaser of the Hopkins online curriculum.

4. Results:
   - 218 of 260 residents (83.8%) completed the survey. All 3 study sites utilize electronic medical records. Residents rated their clinics as least valuable as compared to training in the wards and ICUs.

   - Preceptor characteristics had the greatest impact on resident value of clinic. Clinic operations, patient volume and characteristics, and resident career plans had a much smaller impact on resident value of
In multivariate analysis, the greatest associations with resident value of clinic are below.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relative risk</th>
<th>Confidence interval</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preceptor is good role model</td>
<td>3.44</td>
<td>2.00-5.29</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Resident highly values general medical wards</td>
<td>3.06</td>
<td>1.87-4.38</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Preceptor aware of clinic/social resources</td>
<td>2.35</td>
<td>2.17-2.52</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Patient flow smooth</td>
<td>1.81</td>
<td>1.27-2.30</td>
<td>.002</td>
</tr>
<tr>
<td>Patients have wide range of ages</td>
<td>1.71</td>
<td>1.03-2.38</td>
<td>.039</td>
</tr>
</tbody>
</table>

5. Limitations

• None of these residencies have primary care programs. As few as 10-25% of residents in one program intended to pursue a generalist career; this may have skewed responses.
• The survey did not distinguish generalist career intentions between primary care and hospitalist careers; it was therefore impossible to select for responses amongst those planning primary care careers.
• The survey measured only resident perceptions of clinic and did not objectively measure variables of clinic throughput, preceptor teaching skills, etc.
• The survey cannot assess the importance of an EMR in the resident’s experience since all sites used EMR’s.

6. Implications

• Satisfaction in residency clinics may not affect resident career choices.
• Maximal resident satisfaction may depend more upon preceptor characteristics than upon clinic or patient characteristics. Programs may therefore wish to focus on preceptor selection and faculty development as key elements in improving resident clinic satisfaction.

Article #2. **Use of a registry-generated audit, feedback, and patient reminder intervention in an internal medicine resident clinic – a randomized trial.** Thomas KG, Thomas MR, Stroebel RJ, McDonald FS, Hanson GH, Naessens JM, Huschka TR, Kolars JC. JGIM 2007;22:1740-44.

**Funding:** Education Innovation Award from Mayo Clinic College of Medicine

Summary: A diabetes registry with audit and feedback improved processes of care but did not affect intermediate clinical outcomes.

1. **Background:** Previous work has demonstrated some benefit of disease registry audit and feedback in improving guideline adherence in faculty practice.

2. **Aims:** To assess impact of disease registry audit and feedback upon resident practice.
3. Methods
- Categorical residents in a single program (Mayo Clinic) were randomized to receive this intervention; control residents were assigned to usual clinic education on diabetes care.
- Study residents were intended to attend two one-hour sessions introducing registries in general and the specific diabetes registry functions. Study residents received quarterly feedback comparing their performance with aggregate resident performance. They also received lists of patients who were not in compliance with guideline recommendations. Quarterly letters were automatically generated to intervention residents’ patients who had not had HgbA1c measurements in 6 months and/or LDL measurements within 12 months.

4. Results
- 39 residents were randomized to each of the intervention and control groups. 87% of residents attended at least one of the didactic sessions on registry use. A total of 483 diabetic patients were cared for by the residents. Baseline measures of HgbA1c, LDL, blood pressure, and adherence to guideline monitoring were no different between study and control residents. Patients cared for by residents in the intervention group were statistically more likely to have had HgbA1c and LDL levels measured within target frames, but the intermediate clinical outcomes of HgbA1c, LDL, and blood pressure measures were no different between patients cared for by intervention versus control residents.

<table>
<thead>
<tr>
<th></th>
<th>Intervention Group</th>
<th>Control group</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HgbA1c within 6 mo</td>
<td>155 or 252 (61.5%)</td>
<td>111 of 231 (48.1%)</td>
<td>.01</td>
</tr>
<tr>
<td>LDL within 1 yr</td>
<td>191 of 252 (75.8%)</td>
<td>148 of 231 (64.1%)</td>
<td>.02</td>
</tr>
<tr>
<td>HgbA1c &lt;7.0%</td>
<td>156 of 252 (62%)</td>
<td>135 of 231 (58%)</td>
<td>.42</td>
</tr>
<tr>
<td>LDL &lt;100 mg/dl</td>
<td>152 of 252 (60%)</td>
<td>141 of 231 (61%)</td>
<td>.76</td>
</tr>
<tr>
<td>BP &lt;130/85</td>
<td>126 of 252 (50%)</td>
<td>116 of 231 (50%)</td>
<td>.96</td>
</tr>
</tbody>
</table>

5. Limitations
a. This study was done in a single institution
b. It is unclear whether information sessions had an impact upon results.
c. Since letters to patients were sent in an automated fashion and since letters alone may have prompted improved compliance with monitoring guidelines, it is unclear whether this intervention had any impact on resident knowledge, behaviors, or attitudes.

6. Implications
a. Registry audit and feedback can improve processes of care, but intermediate and meaningful differences in patients outcomes remain illusive and will likely require increasingly complex interventions.

Funding: Swiss Research Foundation on Alcohol

Summary: Residents trained in brief alcohol interventions (BAI) were more likely to use three of 12 recommended counseling strategies in patients seen in outpatient practice than were control residents. No differences were seen in the use of the remaining nine strategies or in patient outcomes.

1. Background
   - Prior work has demonstrated a reduction of alcohol consumption in non-dependant drinkers using the BAI model. Prior studies have shown BAI training can change physician behavior, but these studies were not blinded or were not controlled.
   - BAI steps are as follows:
     1. Address consumption
     2. Explain safe drinking limits
     3. Provide feedback on patient alcohol use
     4. Ask patients to express opinion on limits
     5. Ask patients to their opinion on feedback
     6. Ask patients about positive aspects of alcohol use
     7. Ask patients about negative aspects of alcohol use
     8. Ask patients about opinion on importance of change
     9. Ask patients about readiness to change
    10. Help patients set goals
    11. Support self-efficacy
    12. Provide an information leaflet.

2. Aims:
   - To increase BAI component use in trained residents
   - To demonstrate reduced hazardous drinking amongst patients of trained residents

3. Methods:
   - Primary care residents were assigned to the interventional BAI training or to a traditional didactic program on lipid management. Five residents with prior training in alcohol treatment were excluded. BAI training occurred in two sessions delivered in two half days that were two days apart.
     i. Session 1:
        1. Definition and prevalence of low-risk, hazardous, and dependent drinking; BAI efficacy studies
        2. Theoretical model of patient-centered BAI
        3. Videotape demonstration and discussion
        4. Role-play
        5. Summary checklist of BAI components, textbook on alcohol, and patient educational materials
     ii. Session 2:
        1. Practice with trained standardized patients
One week after training, consecutive patients received a self administered questionnaire on alcohol, tobacco, drug use, cholesterol, immunizations, depression and accidents. Patients were selected for study inclusion if age 19 or older, had a scheduled appointment, and were hazardous drinkers in the past 12 months. Patients were blinded to the aims of the study. Hazardous drinking was defined as:

i. >14 drinks/week and/or >4 drinks/occasion for men <65
ii. >7 drinks/week and/or >3 drinks/occasion women and for men >65

Patients received a feedback form summarizing results of their questionnaire and were instructed to give the form to the resident during the visit. Hazardous drinkers were then interviewed after the visit to determine the type and number of BAI components residents conducted, had an AUDIT assessment, and then interviewed by telephone three months after the visit to assess alcohol intake.

4. Results

- There were ultimately 13 residents in each of the intervention and control groups. They had a median age of 32 vs. 31 years, were predominantly female (58% vs. 64%), and had a median of 5 vs. 4 years of clinical practice experience respectively in this European program.
- 2438 patients were approached. 506 patients were current hazardous drinkers of whom 381 met eligibility. Patients were excluded if they ineligible, if they left before the interview, or they declined participation. 260 patients were enrolled and interviewed after the visit. Enrolled patients were significantly younger than un-enrolled patients (44 vs. 48 years) but were otherwise similar.
- There was no statistical difference in whether or not residents addressed alcohol consumption (BAI 54% vs. control residents 46%) BAI residents performed more BAI components than controls (2.4 vs. 1.5 \( p=.001 \)). BAI residents were more likely to explain safe drinking limits (27% vs. 10% \( p=.001 \)), provide feedback to patients on alcohol use (33 vs. 21% \( p=.03 \)), and to ask patient opinions on safe drinking limits (19% vs. 6% \( p=.02 \)). There were no differences in residents on other BAI components which were performed between 4 and 22% of encounters.
- 219 (84%) patients completed 3 month follow-up. Patients with interval visits had no difference in BAI components between intervention and control residents. In both groups, 37% of hazardous drinkers had become low risk drinkers with no significant difference between groups.

5. Limitations

- Single institution
- Appointment length, number of visits per provider session, and patient and provider visit agendas beyond alcohol use are not described.

6. Implications
• Although the authors don’t explicitly state this, the patient questionnaire and feedback instrument followed by post visit interviews that occurred for all study patients may have been more powerful than the specific BAI behaviors presented and measured. Perhaps the take home lesson is that patient questionnaires and their analysis may affect patient outcomes more than detailed educational interventions with residents.


Funding: Partnerships for Quality Education (PQE) funded by The Robert Wood Johnson Foundation.

Summary: A chronic care residency training intervention focused upon asthma increased chronic care model implementation and reduced emergency room visits.

1. Background
   • Chronic disease management consumes a substantial proportion of health care expenditure; many physicians feel inadequately trained to treat chronic illness.
   • The chronic care model (CCM) components include:
     i. Partnering patients local organizations with chronic illness resources
     ii. Instilling a culture of quality improvement in the health care organization;
     iii. Providing patient support for self management
     iv. Provide structured and planned care designed to maintain health
     v. Embed evidence-based guidelines into clinical practice
     vi. Utilize IT to send patient reminders, identify patients needing outreach, and monitor provider performance.

2. Aims: Assess impact of CCM resident training upon resident use of CCM and on their continuity clinic patients’ resource use.

3. Methods
   • The Maine Medical Center CCM was implemented in resident continuity clinics between July 2002 and December 2003.
   • The clinics developed partnerships with local and state asthma education programs, scheduled individual patients with asthma educators, embedded guidelines into the EMR, planned patient follow-up visits, and developed an electronic asthma registry.
   • Residents visited the outside asthma education programs, attended the visits with asthma educators, did chart reviews looking at guidelines adherence, did QI projects using registry data, and had didactic sessions. The 59 participating residents included 13 in pediatrics, 27 in internal medicine, 2 in medicine/pediatrics, and 17 in family medicine.
   • Residents completed a pre and post intervention survey on components of the CCM. Control residents were drawn from eight other institutions recruited
from other CCM sites where CCM implementation did not extend to all clinic sites.

- Patient outcomes for the intervention institution were assessed by measuring visits to the emergency department comparing 7/01-6/02 as baseline with 7/02-6/03 post intervention.

4. Results

- Differences in resident access to CCM elements are shown below (expressed as % of residents reporting favorable response):

<table>
<thead>
<tr>
<th>Chronic care model element</th>
<th>Intervention</th>
<th>Control Group</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access detailed info on community programs</td>
<td>27.5</td>
<td>51.2</td>
<td>26.0</td>
</tr>
<tr>
<td>Patients in self-management course or support group</td>
<td>0.0</td>
<td>75</td>
<td>5.8</td>
</tr>
<tr>
<td>Patients’ self care routinely observed and assessed</td>
<td>2.5</td>
<td>17.1</td>
<td>7.8</td>
</tr>
<tr>
<td>All or most pts receive individual asthma plans</td>
<td>26.8</td>
<td>56.1</td>
<td>11.8</td>
</tr>
<tr>
<td>Easy to consult asthma guidelines</td>
<td>61.0</td>
<td>95.1</td>
<td>61.0</td>
</tr>
<tr>
<td>Patients called regularly at home to monitor</td>
<td>0.0</td>
<td>5.0</td>
<td>5.8</td>
</tr>
<tr>
<td>Easy report on no well visit in 6 months</td>
<td>19.5</td>
<td>36.6</td>
<td>20.3</td>
</tr>
<tr>
<td>Medical record access during phone consultation</td>
<td>75.6</td>
<td>85.4</td>
<td>62.3</td>
</tr>
<tr>
<td>Individual or team report on guideline performance</td>
<td>15.8</td>
<td>30.8</td>
<td>14.5</td>
</tr>
<tr>
<td>Easy specialist access</td>
<td>53.7</td>
<td>75.6</td>
<td>62.3</td>
</tr>
<tr>
<td>Get timely pharmacist input on complex regimens</td>
<td>46.3</td>
<td>58.5</td>
<td>63.6</td>
</tr>
<tr>
<td>Have someone advocate to solve access barriers</td>
<td>56.1</td>
<td>65.9</td>
<td>37.7</td>
</tr>
</tbody>
</table>

* Not reported, not statistically significant

- Patients cared for by intervention residents had substantial decreased ER utilization compared with all other patients seen at Maine Medical center.

<table>
<thead>
<tr>
<th>Continuity Intervention</th>
<th>All other MMC patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td><strong>Peds ED</strong></td>
<td></td>
</tr>
<tr>
<td>Asthma</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>193</td>
</tr>
<tr>
<td><strong>Adult ED</strong></td>
<td></td>
</tr>
<tr>
<td>Asthma</td>
<td>64</td>
</tr>
<tr>
<td>Total</td>
<td>567</td>
</tr>
</tbody>
</table>

(Intervention patients: 184 pediatric, 257 adult)

- This was estimated to reduce charges from $9293 to $4823 for pediatric patients and $22,890 to $14,442 for adult patients

5. Limitations

- Data was collected only during the intervention period; there is no information on sustainability.
- The impact and importance of specific components of the chronic disease model cannot be assessed.
6. Implications
   - A chronic disease training program can improve health outcomes and reduce costs of care.

Article #5: “Continuity” as an organizing principle for clinical education reform.”

Summary: The authors present a philosophical construct on the importance of continuity in student education. They assert benefits in skill development (i.e. chronic disease management), professionalism, patient-centered care, core doctoring skills, enhanced lifelong learning, inter-professional understanding and collaboration. They acknowledge barriers to continuity including:

| Care                              | Underinvestment in ambulatory care infrastructure |
|                                   | Underinvestment in informational technology infrastructure |
|                                   | Traditional academic medical center organization and culture |
| Curriculum                        | Departmental boundaries and culture |
|                                   | Lack of agreement on educational and assessment strategies |
|                                   | Administrative costs |
|                                   | Inflexible accreditation and other regulatory standards |
| Supervision                       | Incremental faculty teaching effort |
|                                   | Lack of recognition and academic advancement for supervisors |
|                                   | Narrowness of faculty expertise |
|                                   | Lack of interdisciplinary teaching models |
|                                   | Insufficient competency-based evaluation instruments |

Finally, the authors acknowledge a wide range of clerkship organizational structure with seven different organizational structures possible between the extreme ends of entirely sequential discipline-specific clerkships with no continuity component to a fully integrated longitudinal experience. For a full description of the authors work in a novel fully integrated longitudinal third year clerkships, see Article #6.


Summary: A novel clerkship ingrates the core clinical disciplines across the entire third year of medical school rather than sequential individual clerkships. Outcomes showed better or equivalent scores on subject exams and on end of year OSCEs. Students expressed greater proficiencies in a variety of competencies than matched peers in conventional clerkships.

1. Background: Current clerkship models provide fractured clinical experiences with little continuity of patient care or mentoring relationships.
2. Aims: To revise core clerkships into an integrated longitudinal model

3. Methods
   - Of the 189 rising 3rd year students at Harvard, 8 of 18 volunteers were enrolled in this year long program. Each student was paired with faculty members in internal medicine, neurology, ob/gyn, pediatrics, and psychiatry and then worked with these preceptors in longitudinal ambulatory care for the year. Each student developed a panel of patients and followed those patients through the continuum of care. They followed at least 10 pregnant patients through pregnancy and delivery. Students also spent 6 weeks working with a surgeon. In evaluation, intervention students were compared with control volunteer students and with the body of remaining students in their class. There were no statistically significant differences between volunteers and controls on MCAT, USMLE I, or second year OSCE scores; on career plans; or on attitudes on patient-centered care.

4. Results
   - The clerkship achieved the goal of continuity. 100% of intervention students reported that they had very often or often seen patients before admission as compared to only 20% of the comparison students (p < .001). 100% of intervention students had seen patients after discharge compared with 10% of control students (p < .001). Intervention patient cohort sizes ranged from 46-115 patients per student. Intervention students received the majority of their feedback from faculty vs. residents (88.1% vs. 31.5%) and the majority of their mentoring from faculty vs. residents (77.55 vs. 37%).
   - Intervention students performed as well or better than controls in NBME shelf exams and a summative multi-station OSCE, though there were no statistically significant differences.

5. Limitations
   - This program was implemented in a small community hospital; it is not clear that it can be replicated in larger programs.
   - The authors don’t explicitly state program costs, though they acknowledge support.
   - Participating students volunteered for the program; the experience might have been different with unselected students.

6. Implications: This intriguing model provides a template for marked innovation in student teaching.

   Educational Interventions

**Article # 1: A Randomized Controlled Trial Using Insinuated Standardized Patients to Assess Residents’ Domestic Violence Skills Following a Two Hour – Workshop**

Funding: University of Kentucky Center for Research on Violence against Women

Summary: In this randomized controlled trial, 27 volunteer internal medicine residents from a single institution were randomized to a two hour interactive domestic violence (DV) workshop or a control (chronic pain) workshop. Outcomes were measured by thirteen standardized patients that were recruited and were trained to two DV cases. Reliability of their evaluations was established. SP’s portrayed either a 27 year old female with a shoulder injury or a 27 year old female with six weeks of tiredness and sadness. These SP’s were then insinuated into resident continuity clinic at 1-3 months post workshop and/or 4-7 months post workshop. Clinical care given by the residents was assessed using a checklist completed immediately after the clinic visit by the SP. Residents were assessed on whether or not DV was identified; performance on 12 (injury case)-14 (depression case) checklist items; and performance on 8 DV safety plan counseling items. The DV workshop residents did not identify DV as an issue in SP’s any more frequently than did control residents (64% vs. 56%, p=.86) However, the DV trained residents performed better (greater than 75% of items) on the DV checklist items compared to control residents (36% vs. 9%, p=.04) and on safety plan counseling (40% vs. 13%, 9=.04). After a DV patient was identified the DV workshop trained residents provided better clinical care to insinuated SP’s than control residents. Limitations include that this was a single site study with a voluntary and paid group of residents. Also residents were aware that SP’s would be used to assess workshop skills in their clinic settings.

1. Background
   • Fifteen percent of women report abuse by an intimate partner each year
   • Less than 50% of residents report feeling very or somewhat prepared to provide DV counseling.
   • Prior research on impact of resident education on DV has not randomized individual residents and looked at the outcome of clinical care

2. Aim: Assess the impact of a two hour domestic violence workshop on clinical practice of residents through use of insinuated SP’s

3. Methods
   • IRB-approved; Conducted Fall 2003-04
   • Participants: 27 internal medicine volunteer residents from a single institution
     Paid $200
   • Residents received dinner and general orientation and were consented (received information that within six months 1-4 SP’s would be insinuated into their continuity clinic practice, but unaware of issues SP would assess, then randomized to a DV workshop or a control (chronic pain) workshop
   • Workshop –two-hours; interactive
     o Introduction to topic
- One resident interviewed a SP in front of group (10-15 minutes)
- Faculty discussed case and led topic discussion
  - Background information on DV; signs of abuse; methods of screening for DV; elements of safety plan; information on state law
- Insinuated SP’s used to measure clinical care
- Two cases developed
  - 27 year old with shoulder injury
  - 27 year old with feeling tired and sad for 6 weeks
- 13 SP’s trained and underwent reliability testing to assure at least 90% agreement with trainer
- All participants encountered either injury or depressed SP; and most (21/27 residents) encountered both; with 3 residents in each the control and intervention workshop encountering one or the other SP case.
- Outcomes
  - Short term (1-3 months post workshop)
  - Long term (4-7 months post workshop)
  - SP had appointment in clinic and immediately after visit completed checklist
- Logistic regression with encounter with insinuated DV patient as unit of analysis comparing 25 insinuated patients seeing residents who received DV workshop compared with 23 insinuated DV patients seeing residents who received control workshop
- Three dichotomous dependent variables
  - DV identification
  - Resident score 75% or higher on DV relevant items from SP checklist (> 11/14 on injury case or > 9/12 on depression case) considered successful
  - Resident score 75% or higher on 8 DV safety plan counseling items considered successful

### 4. Results
- 27/64 residents volunteered
- 44% were women
- Female residents were no more likely than male residents to identify DV victim (62% vs. 59%) but were more likely to provide adequate safety counseling (43% vs. 15%, p=.04)
- The length of time between the workshop and the SP encounter did not affect results.

Summary of Results (Counted as successful for clinical encounter or successful for safety plan counseling if scored ≥ 75%)

<table>
<thead>
<tr>
<th></th>
<th>DV workshop resident encounter with SP (n=25 encounters)</th>
<th>Control workshop resident encounter with SP (n=23 encounters)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification of SP as DV victim</td>
<td>16/25 (64%)</td>
<td>13/23 (57%)</td>
<td>.86</td>
</tr>
</tbody>
</table>
Successful clinical encounter | 9/25 (36%) | 2/23 (9%) | .04
Successful safety plan counseling | 10/25 (40%) | 3/23 (13%) | .04

5. Limitations
- Single institution
- Volunteer residents who received financial incentive
- Residents aware they would be assessed on workshop skills by SP’s in their clinic setting

6. Implications
- A two hour workshop on DV can improve residents’ clinical skills in caring for DV victims once they are identified.
- Other interventions will need to be assessed to enhance identification of DV victims and to further improve clinical care delivered to these women
- Insinuated SP’s can be successfully used to assess clinical care outcomes of educational interventions.


Funding: None indicated

Summary: This study from a single institution described and evaluated the implementation and outcome of a redesign from a traditional monthly morbidity and mortality conference (MMC) to a new monthly, patient safety morbidity and mortality conference (PSMMC). The redesign process was conducted by an eight person multidisciplinary team. Key goals for the PSMMC were: 1) To teach system thinking to residents, fellows and faculty in support of the ACGME core competencies of system based practice and problem based learning and improvement; 2) Provide a forum for discussion of adverse events and reasons contributing to occurrence; 3) Assist in the transformation of departmental culture to one which values patient safety and QI; and lastly 4) To expand knowledge and skills through modified root cause analysis process. Outcomes assessed include attitudes of residents and fellows regarding patient safety as measured on pre-post 20-item; 5-point Likert scale survey administered prior to the new conference and following the 8th month of the new conference; system improvements generated from the patient safety PSMMC and attendance at PSMMC. 58 of 111 residents and fellows completed the pre-post attitudinal survey. 6/20 items showed substantial change with 4 occurring in the desired direction. 11/14 remaining items trended in the desirable direction. Average attendance increased from 41 ± 8 to 50 ± 10 participants (p<0.03). Through the cases discussed in the 11 months, participants made 121 system improvement recommendations. Based on likelihood of achieving high impact changes, facilitators determined 39 (32%) of the recommendations should be pursued. Of these targeted changes, 23 (59%) were fully implemented; 11 (28%) were partially implemented and 5 (13%) were abandoned due to impracticality or redundancy.
This study demonstrated that an educational intervention could bring about some modest change in attitudes, participation of residents while enhancing system performance in hopes of leading to better patient outcomes.

1. Background:
   - Only 50% of programs use MMC to examine cause and possible remediation for medical errors
     - Focus on individual actions may interfere with identification of systems issues
   - MMC could be used to develop ACGME competencies of practice based learning and improvement and systems based practice

2. Aims: Describe and evaluate the implementation and outcome of a redesign from a traditional morbidity and mortality conference (MMC) to a new monthly, patient safety morbidity and mortality conference (PSMMC)

3. Methods:
   - Setting: University-based Department of Medicine residency training program
   - Redesign process
     - Conducted by an eight person multidisciplinary team
     - Key goals for the PSMMC
       - 1) To teach system thinking to residents, fellows and faculty in support of the ACGME core competencies of system based practice and problem based learning and improvement;
       - 2) Provide a forum for discussion of adverse events and reasons contributing to occurrence
       - 3) Assist in the transformation of departmental culture to one which values patient safety and QI
       - 4) To expand knowledge and skills through modified root cause analysis process
   - Conference description and timeline table
Outcomes

- Attitudes of residents and fellows regarding patient safety as measured on pre-post 20-item
- 5-point Likert scale survey administered prior to the new conference and following the 8th month of the new conference;
- System improvements generated from the patient safety PSMMC and attendance at PSMMC

4. Results:

- 58 of 111 residents and fellows completed the pre-post attitudinal survey
  - 6/20 items showed substantial change with 4 occurring in the desired direction.

<table>
<thead>
<tr>
<th>Item</th>
<th>Sample size</th>
<th>Initial survey</th>
<th>Follow-up survey</th>
<th>Mean change (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in desired direction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting systems do little to reduce future errors</td>
<td>56</td>
<td>3.3</td>
<td>3.0</td>
<td>-0.30 (-0.57 to -0.03)</td>
</tr>
<tr>
<td>Analyses of medical errors and follow-up improvement actions have led to positive changes within our department</td>
<td>57</td>
<td>2.7</td>
<td>2.4</td>
<td>-0.37 (-0.73 to -0.01)</td>
</tr>
<tr>
<td>When an MMC case is presented, I feel the doctor is blamed for the outcome even though the provider's identity is anonymous</td>
<td>55</td>
<td>3.6</td>
<td>3.9</td>
<td>0.27 (0.03 to 0.51)</td>
</tr>
<tr>
<td>Change in undesired direction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competent doctors do not make medical errors that lead to patient harm</td>
<td>58</td>
<td>1.9</td>
<td>2.2</td>
<td>0.26 (0.01 to 0.50)</td>
</tr>
<tr>
<td>MMCs enhance my overall learning</td>
<td>58</td>
<td>4.2</td>
<td>3.9</td>
<td>-0.28 (-0.49 to -0.06)</td>
</tr>
</tbody>
</table>

- 11/14 remaining items trended in the desirable direction.
- Average attendance increased from 41 ± 8 to 50 ± 10 participants (p<0.03)
- 7 residents voluntarily joined performance improvement teams
- Through the cases discussed in the 11 months participants made 121 system improvement recommendations
  - Based on likelihood of achieving high impact changes, facilitators determined 39 (32%) of the recommendations should be pursued.
    - 23 (59%) were fully implemented (TABLE)
- 11 (28%) were partially implemented
- 5 (13%) were abandoned due to impracticality or redundancy.

**TABLE: Implemented systems changes**

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information technology modifications</td>
<td>5</td>
</tr>
<tr>
<td>Education and training</td>
<td>5</td>
</tr>
<tr>
<td>Policy development/modification</td>
<td>4</td>
</tr>
<tr>
<td>Development of standardized reference materials</td>
<td>3</td>
</tr>
<tr>
<td>Modification of committee structure</td>
<td>2</td>
</tr>
<tr>
<td>Modification of medical documentation form</td>
<td>2</td>
</tr>
<tr>
<td>Development of hospital wide interdisciplinary team to investigate similar cases</td>
<td>1</td>
</tr>
</tbody>
</table>

5. Limitations: Single site study and no measurement of direct impact on patient safety

6. Implications: A patient safety conference can bring about some modest change in attitudes, prompt resident participation in performance improvement, and facilitate system improvements.


Funding: none

Summary: This study is a multi-site cross-sectional survey of internal medicine residents conducted to assess residents’ understanding of biostatistics and interpretation of research results. All fifteen internal medicine programs in Connecticut were contacted and eleven were able to participate including 7 university affiliated programs and 4 community based programs. The survey instrument contained 4 sets of questions including: 1) demographics; 2) attitudinal items regarding biostatistics; 3) confidence questions about interpreting statistics and 4) a 20 question biostatistics knowledge test which reflected the statistical methods and results most commonly represented in contemporary research studies. The knowledge test component was pilot tested and modified. The survey was administered in a 25 minute slot at the beginning of respective noon conferences. Of the total 367 residents available to attend in the 11 targeted residency programs, 277 (75.5%) completed the survey. The response rate from individual programs ranged from 28.1 to 80%. The overall mean percentage correct was 41.4% vs. 71.5% for fellows and general internal medicine faculty with research training. Seventy five percent indicated that they did not understand all the statistics they encountered in articles, but 95% felt that understanding these concepts was important. Most residents lacked knowledge in biostatistics needed to interpret results published in clinical research. Higher scores were associated with additional advanced degrees and prior biostatistical training suggesting that further curricula for residents without these experiences could be beneficial.
1. Background:
   • ACGME includes locating, appraising and assimilating evidence from scientific studies as part of it’s competency on practice based learning and improvement
   • Journal clubs tend to emphasize appraisal of studies and study conduct but fewer address selection and interpretation of statistical tests

2. Aim: Develop a new instrument and perform a multiprogram assessment of residents’ biostatistical knowledge and interpretation of student results

3. Methods:
   • Participants: Fifteen internal medicine programs in Connecticut were contacted and eleven were able to participate and included
     o 7 university affiliated programs
     o 4 community based programs.
   • Survey:
     o 1) 11 demographic questions
     o 2) 5 attitudinal items regarding biostatistics
     o 3) 4 confidence questions about interpreting statistics
     o 4) 20 question biostatistics knowledge test
       ▪ reflected the statistical methods and results most commonly represented in contemporary research studies
       ▪ reviewed 239 original articles published from January to March 2005 in each issue of 6 medical journals
       ▪ summarized frequency statistical methods described
       ▪ knowledge test component was pilot tested and modified
       ▪ also administered to fellows and faculty
     o The survey was administered in a 25 minute slot at the beginning of respective noon conferences.
   • Analysis
     o Psychometric properties of knowledge test
       ▪ Internal consistency using Cronbach alpha
     o Inter-rater reliability was determined by calculating interclass coefficients
     o Internal consistency reliability was determined by calculating Cronbach’s coefficient alpha

4. Results:
   • 277/ 367 (75.5%) residents completed the survey
     o Response rate from individual programs ranged from 28.1 to 80%
   • The overall mean percentage correct was 41.4% vs. 71.5% for fellows and general internal medicine faculty with research training.
   • 75% of residents indicated that they did not understand all the statistics they encountered in articles
   • 95% of residents felt that understanding these concepts was important.
   • 77% indicated they would like to learn more about statistics
88% of residents felt they had a fair to complete understanding of P values, but only 55% understood factors affecting statistical power and 38% could assess if the correct statistical procedure was used. Higher scores were associated with additional advanced degrees and prior biostatistical training.

5. Limitations:
   - Psychometric properties of the instrument were not known prior to use in the study
     - Instrument did demonstrate content validity, internal consistency and discriminative validity in this study
   - Surveyed residents who were present at time of their inpatient conference at this may have instituted selection bias
     - No demographic differences between responders and non responders
   - The comparison group of general medicine research faculty and fellows are the physicians with the greatest content expertise, yet even they had only 75% correct. Some might think this test is formulated at the level of an expert statistician rather than a reader of the medical literature.

6. Implications: Most residents lacked detailed knowledge in biostatistics. However, it remains unclear what level of biostatistical expertise is needed to interpret results published in clinical research and engage in practice of evidence based medicine. If there were consensus that all physicians should have the level of statistical knowledge that was assumed to be required in this study, the profession may need to do substantial curricular work amongst faculty and learners.


Funding: None

Summary: This study focuses on the development and validation of an instrument for assessing resident quality improvement (QI) projects. The content of the QI proposal assessment tool was decided by national and Mayo Clinic experts based on consensus findings of the 2004 Annual Achieving Competence Today Conference. From July-December 2005 the initial instrument was tested by faculty and residents on a monthly basis as part of the quality improvement curriculum for Mayo Clinic internal medicine residents. The instrument was modified through this iterative process, and the final 7-item Quality Improvement Proposal Assessment Tool (QIPAT-7) was developed. Items were structured on a 5 point rating scale. After IRB approval the instrument was pilot tested by 5 faculty and 2 chief medical residents on 3 randomly chosen resident QI projects from the prior year. Raters met and resolved all differences in assigned ratings. It was decided that in order to receive a score of 3 or more for an item all anchor descriptors should be achieved. The QIPAT-7 was then used to score the 45 consecutive
resident QI projects from July 2004 through July 2005. Principal factor analysis demonstrated that each of the 7 items represented 1 dimension. Item mean scores ranged from 1.9 to 3.4 on the 5 point scale. Excellent interrater reliability for each item (range 0.79-0.93) and internal consistency reliability among the items (Cronbach’s alpha=0.87) was demonstrated. QIPAT-7 is a useful tool, supported by content and internal structure validity evidence, for evaluation of resident QI proposals.

1. Background:
   - Residents are now required by the ACGME to demonstrate competency in systems based practice and problem based learning and improvement
   - Residencies are involving trainees in QI projects, beginning with proposals, as a way to address these requirements
   - There are no current published methods for assessing such QI proposals

2. Aim: Create a QI proposal assessment tool and assess the validity of its scores for evaluating resident QI proposals

3. Methods:
   - Participants: 5 faculty and 2 chief residents acted as raters using the instrument to evaluate 45 residents QI proposals
   - Instrument (QIPAT-7) development
     - Content was derived from national panel of QI experts that met in May 2004 and developed a consensus on the necessary components and assessment methodology of quality improvement proposals.
     - From July-December 2005 the initial instrument was tested by faculty and residents on a monthly basis as part of the quality improvement curriculum for Mayo Clinic internal medicine residents
     - The instrument was modified through this iterative process, and the final 7-item Quality Improvement Proposal Assessment Tool (QIPAT-7) was developed. Items were structured on a 5 point rating scale.
     - Three randomly chosen QI projects were then evaluated and the raters met to resolve all differences
       - Form was modified to indicate ALL anchor descriptors had to be satisfied to have the item rated as a 3 or higher
   - Data was collected through evaluation of 45 consecutive resident QI proposals completed from July 2004-July 2005
   - Analysis
     - Exploratory factor analysis was used to determine if items clustered into separate domains
     - Interrater reliability was determined by calculating interclass coefficients
     - Internal consistency reliability was determined by calculating Cronbach’s coefficient alpha

4. Results:
   - All seven items represented 1 dimension of QI proposal assessment
   - Item mean scores ranged from 1.9-3.4 on a five point scale
• Interrater reliability for each item ranged from 0.79-0.93
• Internal consistency reliability among items had a Cronbach’s alpha =0.87

5. Limitations:
• Developed at a single institution where residents receive an intensive (64 hour) QI curriculum
• Group of raters used to establish instrument reliability were very familiar to QI concepts and integral to scales development

6. Implications:
• The QIPAT-7 is the first validated tool for assessment of resident QI proposals
• Generalizability to other populations of residents and other educational settings is necessary
• It will be of further value to determine if QIPAT scores correlate with implementation of the project, improved patient outcomes and publication.


Funding: None indicated

Summary: This study sought to determine if multidisciplinary rounds (MDR) could enhance quality core measure performance and resident education while decreasing length of hospital stay. It was conducted at a university affiliated community teaching hospital with a total of 44 residents. The multidisciplinary group involved in rounds included members of the internal medicine service, case managers, nurse coordinators, dieticians, pharmacists and representatives from physical medicine and psychiatric services. MDR was led by the chief of medicine and a clinician educator and was conducted for 1 hour three times a week. Resident teams were only present for discussion of their patients and discussions were very focused on relevant issues. Outcomes on core measures for specific diagnosis (CHF, AMI and CAP) were chosen as focus for MDR and assessment. Length of stay data was sorted by diagnosis and medical attending from administrative databases. Pre-MDR looked at data from July 2002 to June 2003 and MDR was from July 2003 to June 2004. Additionally residents completed anonymous surveys regarding self-reported knowledge and attitudes about MDR after completing 2 blocks of MDR. Analysis of performance on core measures revealed statistically significant improvement in all 3 targeted diagnoses and in 5/10 of the individual core measures after the implementation of MDR. Also during this time period length of stay decreased 0.5 (95% CI 0.1-0.8) days for patients with targeted diagnosis and by 0.6 (0.5-0.7) days for all medicine DRG’s. 100% of medicine residents completed the survey and they reported increase knowledge regarding core measures, system based care and communication after the implementation of MDR. They also agreed that MDR improved efficiency, delivery of evidence based care and relationships with involved disciplines.
1. Background:
   - ACGME core competencies recognize the importance of training in systems based care and the delivery of evidence based care
   - Fiscal restraints are clearly a reality at most institutions

2. Aim: To determine whether resident centered multidisciplinary rounds lead to measurable improvement on core measures embraced by JCAHO and CMS while improving length of stay and enhancing resident education

3. Methods:
   - Setting
     - University-affiliated community teaching hospital in an urban setting-328 beds
     - 44 Internal medicine residents
     - 5 inpatient general medicine teams
       - Residents care for 80% of all medical admissions
   - Intervention/MDR rounds
     - Included members of the internal medicine service, case managers, nurse coordinators, dietician, pharmacists and representatives from physical medicine and psychiatric services.
     - Led by the chief of medicine and a clinician educator
     - 1 hour three times a week.
     - Resident teams were only present for discussion of their patients.
       - Focus on quality and core measure compliance
         - CHF
         - AMI
         - CAP
         - Tasks assigned to appropriate members of care team
     - Monthly performance was posted in the conference room
   - Outcomes
     - Performance of all core measures that could be impacted by MDR
       - CHF (discharge instructions, LVF assessment, ACE-inhibitor for LVSD, smoking cessation counseling)
       - AMI (aspirin at discharge, ACE-inhibitor for LVSD, smoking cessation counseling, beta blocker at discharge)
       - CAP (pneumococcal vaccination, smoking cessation counseling)
       - If core measure controlled by ED or ICU and not general medical unit, was not included
       - Core measure data abstracted and reported to Connecticut Hospital Association in accordance with JCAHO
     - Length of stay one year before and for first year of MDR from administrative data base
     - Resident attitudes and level of knowledge
       - Anonymous 2-section questionnaire
- Asked them to recall knowledge and attitudes for pre-MDR assessment and then to rate current attitudes and knowledge
  - Administered after completing at least 8-12 weeks of new MDR

4. Results:
- Analysis of performance on core measures revealed statistically significant improvement in all 3 targeted diagnoses and in 5/10 of the individual core measures after the implementation of MDR (TABLE).
- Length of stay decreased 0.5 (95% CI 0.1-0.8) days for patients with targeted diagnosis and by 0.6 (0.5-0.7) days for all medicine DRG’s.
- 100% of medicine residents completed the survey
  - Reported increase knowledge regarding core measures, system based care and communication after the implementation of MDR.
  - Agreed that MDR improved efficiency, delivery of evidence based care and relationships with involved disciplines.

**TABLE: Significant changes in Outcomes of core measures between 1 year pre-MDR (July 2002-June 2003) and the First year of MDR (July 2003-June 2004)**

<table>
<thead>
<tr>
<th>Core measure</th>
<th>Pre-MDR</th>
<th>MDR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eligible #</td>
<td>Success rate %</td>
<td>95% CI %</td>
</tr>
<tr>
<td>Heart failure</td>
<td>780</td>
<td>65</td>
<td>62-69</td>
</tr>
<tr>
<td>Discharge instructions</td>
<td>261</td>
<td>21</td>
<td>17-27</td>
</tr>
<tr>
<td>AMI</td>
<td>201</td>
<td>89</td>
<td>83-93</td>
</tr>
<tr>
<td>Smoking cessation counseling</td>
<td>14</td>
<td>57</td>
<td>29-82</td>
</tr>
<tr>
<td>Beta blocker prescribed at discharge</td>
<td>81</td>
<td>90</td>
<td>81-96</td>
</tr>
<tr>
<td>CAP</td>
<td>351</td>
<td>27</td>
<td>23-32</td>
</tr>
<tr>
<td>Pneumococcal vaccination</td>
<td>261</td>
<td>10</td>
<td>7-14</td>
</tr>
<tr>
<td>Smoking cessation counseling</td>
<td>90</td>
<td>78</td>
<td>68-86</td>
</tr>
</tbody>
</table>

5. Limitations: This was conducted at a single institution around three specific core diagnoses and relevant core measures. The measure of resident knowledge was based on self-perception of knowledge and both pre-MDR attitudes and knowledge measures were based on recall.

6. Implications: This MDR model requiring no additional equipment, technology or FTE’s, but only three times a week hour long rounds presents an attractive model for achieving goals of resident education regarding systems based care, quality improvement on core measures and decreasing length of stay. Other institutions
should consider replication of this model to determine if similar outcomes can be achieved.

**Professionalism**


Funding: none indicated

Summary: This study sought to compare peer ratings of work habits (WH) and interpersonal attributes (IA) of medical students with later medical student performance evaluation (MSPE) rankings and ratings by residency program directors (PD’s). Peer assessment data collected during the second- and third-year on 240 students from three medical school classes at a single institution were compared against the MSPE rankings of those students. For 43 students in one class, peer assessment data were also compared to data obtained from internship PD’s. Both second- and third-year peer assessed WH were predictive of MSPE rankings but the third year assessment was found to be a stronger predictor. PD ratings were also significantly related to peer-assessed WH scores. IA scores were not predictive of either MSPE groups or program director ratings. This study shows that peer assessment of WH can predict MSPE rankings and internship performance. By identifying such issues early in medical training, medical schools can institute prompt remediation and education efforts.

1. Background: The dean’s letter or MSPE provides a summary of students’ performance during medical school and is, therefore, more a letter of evaluation rather than of recommendation. As recommended by the Association of American Medical Colleges, MSPE’s often include some ranking of students in comparison to their peers and MSPE rankings of graduates are closely related to PD’s later evaluations (1). Peer assessment is one method of measuring specific skills and attributes over the course of clinical training. If predictive of MSPE rankings or of residency program director’s ratings, data from these assessments could be useful in early identification efforts as well in the design of educational and remediation interventions.

2. Aims: To compare peer assessment ratings of WH and IA of second- and third-year medical students with later MSPE rankings and ratings by PD’s.

3. Methods: Medical students who graduated in 2004, 2005 or 2006 from the University of Rochester School of Medicine and Dentistry and who had participated in peer assessment exercises during second and third year.

- MSPE
  - Students were ranked in one of four groups (outstanding, excellent, very good or good) based upon weighted grades in required clinical clerkships.
Approximately 20% of students are designated as outstanding, 25% excellent, 50-55% very good and <5% good.

Peer assessment
- Students anonymously assessed 6-12 classmates once in second and third year. They assessed a different group of classmates each year.
- A standardized rating form included independent items assessing WH and IA. Cronbach’s $\alpha$ on WH and IA scales was >0.8 (2).
- Scores were standardized to yield z-scores with mean=0 and SD=1.0 for WH and IA.

PD survey
- 15-item survey sent to PD’s approximately 10 months after the class of 2004 had graduated
- Survey rated graduates on general clinical, interpersonal and professional qualities
- Factor analysis consistent with one-factor solution

4. Results: Of 281 medical students who graduated from the classes of 2004, 2005 and 2006, 240 (85.5%) had completed peer assessment forms and were included.
- Multivariate one-way ANOVA for second- and third-year WH and IA was significant ($F_{12,614}=9.93$, p<.001)
- Univariate ANOVA showed significant differences in peer assessed WH between the 4 MSPE groups in second ($F_{3,10.8}=44.90$, p<.001) and third year ($F_{3,9.65}=29.54$, p<.001)
- No differences in peer assessed IA between 4 MSPE groups.
- Quadratic discriminant function analysis to assess how peer assessed WH and IA were related to later membership in MSPE groups was conducted.
  - Model using WH scores did discriminate among students ($F_{6,468}=18.78$, p<.001, $R^2=.26$)
  - Model using IA scores was not statistically significant
- PD response received on 43 students from class of 2004 (44%)
  - Overall ratings significantly correlated with second ($r=.32$ [p=.015]) and third ($r=.43$ [p=.004]) year WH scores
  - No correlation with IA scores

5. Limitations: This study was conducted at a single academic medical center and thus results may not be applicable to other institutions or to situations where a different method is used to assign MSPE rankings. Further, this center has a well established infrastructure for conducting peer assessments. The importance of well trained students for the task of rating their peers cannot be underestimated. Finally, the low response rate (44%) from PD’s limits the reliability and validity inferences of the correlations between peer assessment and future ratings of graduates.

6. Implications: Among medical students, identification of attributes related to work habits can be assessed by peers and are predictive of later MSPE rankings and PD ratings. Such recognition, as early as during the second year of medical school, can be helpful to
trigger remedial and educational interventions sooner rather than later. In addition, this study provides evidence to support peer assessments during medical school training.

References

1. Lurie SJ, Lambert DR, Grady-Weliky TA. Relationship between dean’s letter groupings and later evaluations by residency program directors. Manuscript under review


Funding: Academic primary care grant and professionalism grant, both from Mayo Clinic College of Medicine

Summary: This study sought to determine the correlation between empathy and distress and well-being among a sample of US medical students. 545 students from three medical schools in Minnesota completed an electronic survey consisting of validated instruments to measure empathy, distress (i.e. burnout and symptoms of depression) and well being (quality of life). Domains of burnout inversely correlated with empathy independent of gender and symptoms of depression inversely correlated with empathy for women. A sense of personal accomplishment positively correlated with empathy across genders and, in multivariate analyses, both burnout (negative correlation) and well-being (positive correlation) independently correlated with empathy scores. This study is the first to explore such a relationship and the findings suggest that efforts to reduce student distress be included in professionalism education and training activities.

1. Background: Professionalism has been recognized as a core competency for physicians in training (1) and empathy- the ability to listen to, understand, sympathize and provide support- is one of its fundamental characteristics. Previous studies have reported a decline in empathy during the course of medical school (2,3) and researchers have hypothesized that the training curriculum itself may lead to or contribute to this finding. At the same time, physician burnout, depression and distress have also been recognized as important influences on practice habits. If these are found to impact medical student empathy, then efforts to reduce distress and improve well-being could be implemented.

2. Aims: Among a sample of US medical students, the aims of this study are to determine whether a) lower levels of empathy are associated with personal and professional distress and b) higher degree of personal well-being is associated with higher levels of empathy

3. Methods:
   • All 1098 medical students in the state of Minnesota were eligible.
   • Multi-institutional (Mayo Clinic College of Medicine, University of Minnesota-Minneapolis and University of Minnesota-Duluth), cross sectional study
• Survey
  - Electronically administered in April 2004, 118 questions
  - Empathy measurement: Interpersonal Reactivity Index (cognitive and emotive domain subscales)
  - Distress measurement: Maslach burnout inventory
    - Emotional exhaustion subscale
    - Depersonalization subscale
    - Personal accomplishment subscale
  - Depression: 2-item primary care evaluation of mental disorders tool
  - Quality of life (QOL)/well-being: Linear analog self-assessment questionnaire

4. Results: Of 1087 students with correct contact information, 545 (50%) responded
  - Medical student empathy scores were higher than normative samples of similarly aged college students
  - No significant difference in levels of cognitive and emotive empathy by year of training for men or women
  - Increasing depersonalization was associated with a decrease in cognitive and emotive empathy, independent of gender (all p<0.02)
  - Increasing emotional exhaustion correlated with lower emotive empathy scores for men (p=.009)
  - Symptoms of depression inversely correlated with empathy for women (all p<.01)
  - Increasing sense of personal accomplishment correlated with higher cognitive and emotive empathy, independent of gender (all p<.001)
  - Multivariate analysis:

<table>
<thead>
<tr>
<th></th>
<th>Emotive Empathy</th>
<th>Cognitive Empathy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Increase</td>
<td>Decrease</td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Personal accomplishment</td>
<td>Depersonalization</td>
</tr>
<tr>
<td>Overall QOL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Personal accomplishment</td>
<td>Depersonalization</td>
</tr>
</tbody>
</table>

All p<.05

5. Limitations: The response rate in this study was less than ideal (50%) and thus response bias is a possibility. In addition, the cross sectional nature of the study provides only evidence of association and not causality. Furthermore, despite the statistically significant correlation between measures, the programmatic and educational implications of such small changes in magnitude remain unclear.

6. Implications: This is one of the largest studies of medical student empathy and the first to explore the relationship between personal and professional distress and empathy. The findings of a relationship of both distress and well being with empathy, an integral
component of professionalism, among medical students suggests that efforts to promote empathy must consider these important influences.

References
1. ACGME outcome project. Available at http://www.acgme.org/outcome/


Funding: none stated

Summary: This study sought to examine changes in medical students’ perceptions of and participation in unprofessional behaviors before and five months into clinical clerkships. Students completed a survey to assess whether they observed, participated in and considered each of 27 behaviors as unprofessional. Student observation of (21 of 27) and participation in (17 of 27) unprofessional behaviors increased significantly over 5 months. Moreover, students perceived 6 unprofessional behaviors as increasingly appropriate and participation in 9 unprofessional behaviors was associated with diminished likelihood of considering those behaviors as unprofessional. This study provides evidence of concerning changes in perception and behavior of medical students after only a short exposure to clinical activities.

1. Background: Participation in unprofessional behaviors among medical students has been associated with subsequent disciplinary action by a medical board (1). Research suggests that non curricular elements are powerful contributors to the development of such unprofessional behaviors (2,3). Thus data from studies designed to assess the informal curriculum may facilitate development and timely implementation of curricula about professionalism.

2. Aims: To examine changes in student perceptions of and participation in unprofessional behaviors immediately before and midway through third year clinical clerkships.

3. Methods:
   • All rising third year medical students at one academic medical center participated in the study.
   • Students completed a 27-item paper survey before the start of their clinical clerkships, and an electronic version of the same survey five months into their clinical clerkships.
   • Survey
     o Developed using literature and student suggestions. Edited by authors
4. Results: Response rate was 100% (N=104) for pre-survey and 62% (N=61) for post-survey

- Significant increase in student observation noted in 21 of the 27 behaviors (p<.05)
- Significant increase in student participation noted in 17 of 27 behaviors (p<.05)
  - For 15 of the 17 behaviors with increased student participation over time, student perception of the behavior as unprofessional declined by at least 5%
  - Magnitude of change was >14% for 6 behaviors (late to rounds, absent from mandatory lectures, use of workrooms for non-clinical activities, not correcting someone who mistakes you for a physician, taking food meant for patients, and having personal conversations in patient corridors)
- Students who acknowledged participation in following behaviors were significantly less likely than those who did not participate to consider these behaviors unprofessional: absence from mandatory lectures, not correcting someone who mistakes you for a physician, attending a drug-rep sponsored event, eating or drinking in patient corridors, taking food from lectures you are not attending, taking food meant for patients, inebriation at school events, discussing with patients information beyond your level of knowledge, and consenting a patient for minor procedures without supervision.

5. Limitations: The lack of pairing between pre- and post-behaviors prevents paired analyses of results. In addition, students may have underreported their true perceptions given the nature of the topic and social desirability. Finally, the survey used in this study was developed by the authors and has not been validated.

6. Implications: Survey results from this study show changes in medical students’ perceptions of and participation in unprofessional behaviors during a five month period of exposure to clinical medicine. Furthermore, participation in unprofessional behaviors was associated with an enhanced likelihood of viewing the behaviors as acceptable. One concerning explanation for such findings is that students witness, and subsequently emulate, unprofessional conduct by those who teach and supervise them. Thus further research into this important area is warranted.

References


Funding: Grant from the Arnold P. Gold Foundation and College of Medicine Chapman Education Center.

Summary: This study sought to validate a peer nomination form to identify outstanding students in clinical competency, caring and community service in three institutions. In addition, the authors attempted to find the simplest method to analyze data obtained from the forms. Factor analysis of data from the longer 12-item version of the forms was used to develop a shorter, 6-item version that displayed very similar factor characteristics across all three sites. Rankings based on peer nomination results analyzed by simply counting nominations distinguished at least the top 15% of students for each category. Moreover, this method resulted in very similar rankings to those obtained by the more complicated factor analysis method (85-90% of same students identified). Schools with ongoing peer assessments programs may wish to include this tool as a method of identifying outstanding students for commendation and recognition.

1. Background: Peer evaluations have been used to rate learner performance, guide feedback and influence behavior especially in the domains of cooperative learning and interpersonal skills (1,2,3). In addition to the more common peer evaluation method where colleagues rate one another on a variety of abilities, peer nomination involves learners nominating a limited number of classmates who best fit various situations (4). While less effective for providing feedback, this is a powerful method to discriminate extremes and thus may be appropriate for reliably identifying student exemplars for recognition and honors.

2. Aims: To assess the reliability of a revised peer nomination form across three institutions, to test the hypothesis that peer nomination could be used to identify students who demonstrated humanism and to compare the use of factor analysis with a less complex method of obtaining student rankings.

3. Methods: A convenience sample of 2 public and 1 private medical school participated in the study. Students in their third or fourth year were surveyed using one of two survey forms.
   - Forms: adapted from one developed and in use at two institutions, forms were pilot tested, and refined by a group of conference participants
     - Form A: 12-items, administered in 2003
     - Form B: 6-items, results from administration of Form A were used to develop Form B, administered in 2004
- Factor analysis of datasets: to analyze relationships that account for correlations among survey items
- Rankings based on counting nominations: scores calculated for each student as sums of number of nominations for survey subsets

4. Results:
- Form A: Response rate 46%-99%
  - 2 major factors (Eigen values>1) identified at all 3 schools. Items 1, 6 and 9 accounted for most of the variance (37%-61%) and contributed to the “clinical competence” factor. Items 2, 3, 7, 8, 10 and 12 accounted for 13-26% of the variance and contributed to the “caring” factor. Item 5 accounted for only 7%-9% of variance and contributed to “community service”
- Form B: Response rate 40-84%

<table>
<thead>
<tr>
<th>Form A</th>
<th>Form B</th>
<th>Tag</th>
<th>Peer nomination items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Emergency</td>
<td>The classmates you would like to have work at your side in a medical emergency.</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Likeable</td>
<td>The classmates who are the most likable.</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>Caring</td>
<td>The classmates who best personify the quote, “the secret of good patient care lies in caring for the patient.”</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Best docs</td>
<td>The classmates you think will make the best all-around doctors.</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>Community</td>
<td>The classmates who have shown exceptional interest in service to their communities.</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>Data</td>
<td>The classmates whose data you would unhesitatingly accept in the event that there are a number of conflicting observations about the same patient.</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>Discuss</td>
<td>The classmates you would like to call on to discuss a personally disturbing event.</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>Bad news</td>
<td>The classmates you would trust most to deliver bad news to a patient, for instance, a diagnosis of terminal cancer.</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>Residency</td>
<td>The classmates who would be the best choices for a highly desired residency. Residencies for future graduates will depend on these persons’ performance.</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Respect</td>
<td>The classmates who most demonstrate respect for colleagues and the health care team.</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>Own doc</td>
<td>The classmates you would want as the doctor for yourself or a loved one.</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>Listening</td>
<td>The classmates who have the best listening skills with patients.</td>
</tr>
</tbody>
</table>

*Third- or fourth-year medical students from three medical schools underwent peer nomination in 2003 and 2004; see text for details. Shorthand tags are used to refer to each item in Figures 1 and 2.
2 major factors (Eigen values>1) identified at all 3 institutions. Items 1 and 4 contributed to clinical competence, items 2 and 6 to caring. Item 3 was minor factor and contributed to community service.

- Counting nominations: number of nominations each student received for items relevant to characteristic counted.
  - Method identified top 15% of students for each characteristic at all 3 institutions
  - Ranking very similar to that obtained by factor analysis (85-90% of same students identified)

5. Limitations: The rates of survey completion by students from the three institutions varied considerably in this study (40%-90%). In addition, the study does not mention any training that students received before completing the forms.

6. Implications: This study provides evidence to support the use of a simple 6-item peer nomination form to identify junior or senior student exemplars in the areas of clinical competence, caring and community service. Rankings based on peer nomination results analyzed by simply counting nominations distinguished at least the top 15% of students for each area. Schools with ongoing peer assessments programs may wish to include this tool as a method of identifying students for commendation and recognition.

References

Clinical Reasoning


Funding: none

Summary: Precepting third-year medical students in the outpatient setting is challenging. The authors sought to develop specific strategies for students to use in diverse outpatient settings to more uniformly learn to obtain focused, appropriate history and physical examinations. They developed ‘focus scripts’ to guide students. Students and faculty at a medical school developed and pilot tested both ‘acute’ and ‘chronic’ focus scripts, both generic and specific to common problems presenting to clinicians in the ambulatory setting. Students at two other medical schools were randomized to receive instruction in use of focus scripts or no instruction. Scoring of write-ups (progress notes) revealed that
focus scripts facilitated a specific task of learning to document focused evaluations in acute and chronic office visits.

1. Background: Although third-year medical students are generally prepared to perform complete history and physical examinations when beginning clinical clerkships, they may not yet have any experience focusing the history and examination as it relates to the patient’s concern. In the time-pressured ambulatory setting, this skill set is particularly important. Based on medical education literature regarding illness scripts (1,2) and teaching scripts (3), the authors sought to develop a learning tool that would help bridge the gap between students with little or no experience thinking “clinically” and students with more experience. The goal was to develop a tool (framework) that would help students to gather and organize information similar to the information sought by preceptors using teaching scripts when listening to the students’ case presentations (symptoms, risk factors, physical signs associated with the patient’s diagnosis).

2. Aims: To test the hypothesis that students using the focus scripts would score higher in documentation in progress notes during their 1st week of outpatient medicine, especially in appropriateness of history and physical examination.

3. Methods: Design: randomized controlled trial; third-year medical students at 2 medical schools during their ambulatory-based internal medicine clerkship; unit of analysis was a written progress note at the end of the first week of the clerkship comparing students who were introduced to ‘focus scripts’ as a method of organizing the clinical encounter. Intervention students were introduced to the focus scripts by the clerkship director during orientation and were told that the scripts might improve organization in history and physical exam performance by use of the structured format. Sixty progress notes were coded using existing, published criteria for utility (4,5) consisting of 11 variables either present or absent. Analysis: independent two-group t-test, two tailed, alpha = 0.05.

4. Results:
   - Eleven progress note variables were assessed for the intervention and the control groups.
   - The following variables were significantly better in the intervention group compared with the control group(p< or =0.05): history is appropriate, physical exam is appropriate, analysis of lab values is appropriate, incorporation of lab values in diagnostic plan is appropriate, diagnosis is clear, and total score.
   - Other variables not significantly different between control and intervention group included: problem is clear, patient stability, differential diagnosis is appropriate, patient plan is clear, treatment appropriate for severity of illness, and cross-cover utility of the note.

5. Limitations: Data collected from only two schools and only 60 notes analyzed; some schools use template note formats and will not lend themselves to using this instructional method; the authors could not control for preceptor style of instruction and can only assume the variable influence to be random and equally distributed.
6. Implications: Proper documentation of an encounter is an important skill. Because of time constraints, focusing data gathering and presentations succinctly is critical for outpatient-based teaching experiences. Not only did the authors demonstrate improved documentation of appropriate history and physical examination in the intervention group, these students were also able to incorporate analysis of lab values and make clearer diagnoses than the control group students, suggesting that the ‘focus scripts’ may facilitate the development of patterns of clinical information, an important step in developing clinical reasoning skills.

References
5. Weed LL. Medical records that guide and teach. NEJM, 1968; 278:593-600.


Funding: none stated

Summary: The relation between knowledge structure and diagnostic performance is unclear. These authors sought to examine the relationships between expert-type concepts in the storage and retrieval of knowledge and diagnostic success, and teaching methods in building expert-type concepts in memory. After adjusting for clinical case presentation, the authors found an independent association between expert-type concepts in knowledge structure and diagnostic success. The authors also found an association between the use of diagnostic schemes by preceptors and the number of expert-type concepts in knowledge structure. The authors did not find an association between the teaching methods of the small group preceptors and diagnostic success, however.

1. Background: Diagnostic reasoning is a task that requires retrieval and application of stored knowledge in relation to a new clinical problem (1,2). How medical knowledge is stored (knowledge structure) is probably related to diagnostic success. Experts appear to store knowledge in a way that is consistent with how it will be retrieved during diagnostic reasoning. Diagnostic schemes are conceptual frameworks around which knowledge can be organized. It would make sense, then, that teaching medical students using diagnostic schemes in the learning experience might enhance the incorporation of expert-type concepts into the students’ knowledge structure.
2. Aims: The primary objective of this study was to examine the relationship between knowledge structure and diagnostic success in medical students, hypothesizing that students with expert-type concepts in knowledge structure would be more likely to achieve diagnostic success. The secondary objective was to examine the relationship between teaching strategy and expert-type knowledge structures, hypothesizing that use of diagnostic schemes as a conceptual framework for teaching would be associated with the presence of expert-type concepts in knowledge structure.

3. Methods: Design: cross-sectional observational study; Subjects: 30 first year medical students at one medical school; teaching setting included large group lectures on acid-base disorders using diagnostic schemes followed by small group sessions with preceptors, some using diagnostic schemes. Data gathered using a concept-sorting (3) computer program (ConSort®) that allows students to form a hierarchal network of diagnostic groupings and name the groupings, performance on problem-solving questions in four clinical presentations (acid-base disorders) extracted from a 69-item examination, and end-of-course questionnaire. Five nephrologists completed the ConSort® task, serving as a referent group to define expert-type knowledge structures. Concepts were included in the expert framework if they were identified in the knowledge structure of 3/5 nephrologists. Authors counted the number of student concepts shared with the expert framework to score student performance. Two raters each scored 10 students twice separated by 2 weeks to determine interrater and intrarater reliability.

Analysis: for study objective 1, multiple logistic regression was performed using diagnostic success as the dependent variable, controlling for clinical presentation and scheme utilization by preceptors during small group sessions. For student objective 2, multiple linear regression was performed with the number of expert-type concepts as the dependent variable. A two-sided t test was used to compare the number of expert-type concepts for students whose preceptors did or did not use schemes during small group learning.

4. Results:

- Concept sorting scoring interrater reliability was 0.9
- Concept sorting scoring intrarater reliability was 0.99
- Cronbach’s alpha for the entire 69-item examination was 0.85, and content validity was high (87.5% of students agreed that the exam evaluated material actually taught.)
- On the end-of-course questionnaire, 90% of students recalled schemes being used during lectures.
- No interaction between clinical presentation, scheme utilization by preceptors, and number of expert-type concepts in knowledge structure was detected.
- Number of expert-type concepts identified in students’ knowledge structure and diagnostic success had an odds ratio of 1.18 (1.03, 1.35, p = 0.016), meaning for each additional expert-type concept identified the odds of diagnostic success increased by 18%.
- Diagnostic success was not related to preceptors using a diagnostic scheme during small group teaching.
The number of expert-type concepts in knowledge structure was higher for students who had attended sessions where preceptors used diagnostic schemes for teaching (2.22 vs. 1.86, p = 0.01).

5. Limitations: Single institution, multiple small group instructors who may not have had a standardized approach to teaching the material (not described), observational design does not permit drawing causal conclusions, only analytic reasoning was studied, only cases related to acid-base disorders were used, all limiting our ability to generalize these findings to other settings, more advanced learner groups, and other case types.

6. Implications: The development of expert-type concepts in knowledge structures may improve subsequent diagnostic performance. If the work of the these authors prove reproducible in additional content domains and settings, how we teach content to medical students may impact knowledge structure or the way knowledge is stored and retrieved when needed for solving new clinical problems. Teaching using diagnostic schemes as a conceptual framework is unlikely to be harmful and may prepare students for their clinical experiences.

References


Funding: Operating grant from the Canadian Institutes of Health Research

Summary: The clinical case presentation is a natural “meeting place” for trainees and clinical teachers. The discourse serves several purposes (patient care, teaching, assessment, etc). This study explores the evaluative function of the exchange between attending physicians and trainees in an emergency department and provides a taxonomy of attending physician question types that serve to assess the performance of the trainee and guide the attending physician’s decision about appropriate supervision and autonomy on a case-by-case basis.

1. Background: A central premise of clinical training is that trainees must experience independent patient care in order to develop into autonomous, self-regulating physicians (1). Attending physicians (APs) must constantly assess trainees’ competence to act independently so they can provide a level of supervision that will allow appropriate independent practice while ensuring patient safety and quality of care. Interactions between trainees and APs are centered on case presentations. Through the case
presentation and subsequent discussion, APs make a determination about the degree of independence the trainee will be granted on that specific case (2,3).

2. Aims: The purpose of this study is to explore, through a discourse analysis of case presentations, the process of assessment of competence for case-specific clinical independence.

3. Methods: Observational study; Ten sessions of audio-recorded observations (30.5 hours) in the emergency department of one urban teaching hospital. Participants included 4 APs, 5 residents, 5 medical students. Purposeful sampling was employed to ensure inclusion of both genders and different levels of experience; recruitment continued until saturation of data was reached. Observations were conducted at different locations (e.g. minor versus major trauma) and different times of day to capture variation related to context (acuity, workload, etc). Audio-recordings were transcribed and divided into discrete case presentations. 26 case presentations were analyzed with a focus on questioning strategies.

4. Results: AP questioning strategies followed the basic framework proposed by Weinholtz (1): clarifying questions (asked to assure the AP’s understanding of the situation) and probing questions (asked to determine the extent of the trainee’s knowledge or understanding.) Probing questions were subdivided into two categories: case-related probing questions and knowledge-related probing questions. The coding framework also categorized challenging questions (asked to challenge trainees’ presuppositions).

Case-related probing questions tended to be asked toward the end of the case presentation, testing the trainee’s formulation for congruence with the AP’s formulation. Knowledge-related probing questions tended to be triggered by some inconsistency in the clinical information provided and seemingly intended to end with a teaching point. Challenging questions are asked for proof that a trainee actually does have the knowledge/understanding assumed to be shared between them (e.g. “is it hematemesis or hemoptysis?”), which sometimes results in additional unsolicited information from the trainee.

5. Limitations: AP questioning strategies may have been influenced by the observer (Hawthorne effect), although efforts were made to be as unobtrusive as possible. APs and trainees were not aware of the purpose of the study, however. This was a single site, single specialty study (emergency medicine) and may not generalize well.

6. Implications: Questioning strategies of APs are part of their tacit expertise as teachers. Clarifying questions and knowledge questions do not seem to serve an evaluative function. (Knowledge questions more often serve to determine the teaching need.) Case-related probing questions, which test the trainee’s assessment and plan against those of the AP, and challenging questions, which request evidence from the trainee to support assumptions of shared knowledge, both play a role in the AP’s assessment of the trainee’s competence. Specific discourse techniques can be successfully taught to clinical teachers. Information from this study could potentially form the basis of additional
faculty development for skills in assessing trainee’s readiness for independent practice. Trainees could also be taught to anticipate types of questions and to understand their importance, minimizing “dysfluency” that can arise in public questioning (and potentially have a positive effect on the learning environment).

References


Funding: none stated

1. Background: Though most medical schools administer comprehensive clinical skills assessments to identify students who have not achieved competence, the types of problems uncovered by these exams have not been characterized. Better understanding of these problems may facilitate more effective remediation and curricular improvements.

2. Aims: The purpose of this study was to characterize the problems students demonstrate in comprehensive standardized-patient (SP) assessments, the causes of these problems, and their amenability to remediation.

3. Methods: Purposeful sampling of key faculty in the nation’s medical schools responsible for remediation after the medical school’s comprehensive clinical skills assessment; consenting individuals were interviewed by phone and calls were recorded and transcribed for analysis. Grounded theory was used to analyze the data concurrently with data collection, using the constant comparative approach to determine when additional interviews were no longer yielding new information about themes.

4. Results: Thematic saturation was reached after 33 interviews; participants were MDs (73%), PhDs (15%), and other educational leaders (12%); all geographic regions of the AAMC were represented, as were both public and private medical schools. Three types of problems were identified: technique, cognitive, and non-cognitive. Technique problems included interviewing and physical exam techniques, and these problems were frequently blamed on poor techniques role modeled during clerkships. Participants thought technique deficits were straightforward to remediate.

Problems rooted in knowledge (e.g. when to examine the heart) or professionalism (e.g. did not care that the heart needed to be examined) were categorized as cognitive and non-cognitive. Cognitive problems related to underlying data-gathering,
clinical reasoning, or knowledge deficits and often manifest as an inability to focus the encounter appropriately, resulting in premature closure (failure to ask open-ended questions or adequately characterize the chief complaint). Non-cognitive problems related to communication or professionalism manifest as detachment, poor insight, lack of empathy, or resistance to the examination process. A small number of students denied performance problems, blaming external factors for their performance.

Problem etiologies included: testing environment, misinterpretation of expectations, difficulty investing in the artificial nature of an SP exam, faulty knowledge or clinical reasoning (e.g. generating an incorrect differential diagnosis). Poor clinical reasoning was in some cases attributed to inadequate knowledge (e.g. disorganized encounters because of failure to recognize the appropriate differential diagnosis for the problem). Participants perceived that clinical role models may contribute to the cognitive problems as well by role modeling use of pattern recognition rather than modeling reasoning through a differential diagnosis.

Remediation was perceived to be difficult for the cognitive and non-cognitive problems. Generally, if a knowledge problem was identified, participants agreed that correction would take time with extensive reading. Until knowledge was improved, it was impossible to know if the student could apply knowledge (reason) to the cases. Non-cognitive (professionalism) problems were the most challenging to remediate, requiring a change in the environment or the individual. The comprehensive assessment often verified what was known (problem behaviors) but had not yet been documented. Student insight was considered critical to the success of remediation.

5. Limitations: Authors relied on remediation officers to recall, summarize, and relay information from their experience rather than direct observation. This limitation is probably overcome by the large sample size and grounded theory qualitative methods.

6. Implications: Problems in students’ comprehensive performance can be identified and categorized. Some problems can be addressed through correction of technique (proper examination technique). Some problems may be related to teaching history and/or examination in isolation from the actual clinical encounter and have broader curricular implications (1,2,3). It is possible that teaching clinical and communication skills in supervised clinical settings may improve clinical reasoning abilities. Some problems are related to clinical teachers role modeling inappropriate techniques (examining patients through gowns), reasoning (unable to explain how one obtained the correct diagnosis), communication (using a less patient-centered approach), and professionalism (4). Addressing these problems will require faculty development and feedback. Some problems are long standing but undocumented and require re-evaluation of multifaceted input in student performance assessment for early detection and intervention.

References


Fung CC, Relan A, Wilkerson L. Acad Med. 2007;82(10 Suppl):S97-S100

Funding: none stated

1. Background: No one will argue the importance of learning from experience in clinical education. The nature of that “learning” experience is less clear. Exposure, even when defined as immersive or comprehensive, may not provide the deliberate practice needed for medical students to develop expert-type skills required for more independent practice (1). Although the Liaison Committee on Medical Education required US medical schools to quantify students’ clinical encounters, optimum patient exposure needed to predict performance has been elusive. Studies on the relationship between process measures recorded by students and outcome-assessment measures have been sparse and inconclusive. Studies predicting performance from patient exposure have failed to show the expected outcomes.

2. Aims: The purpose of this study was to examine the relationship between the amount of comprehensive patient exposure experienced in three core clerkships in primary care and performance on a clinical performance examination (CPX).

3. Methods: Design: Retrospective study based on personal digital assistant (PDA) patient log data and CPX performance scores gathered on a cohort of 166 3rd year medical students at one medical school. From the raw data in the PDA records, authors used only those encounters coded as having “full” responsibility for patient care and those who presented with not more than 2 complaints for six clinical problems (chronic cough, hypertension, abdominal pain, back pain, chest pain, and diabetes) in the students’ internal medicine (hospital-based), primary care (ambulatory-based), and family medicine clerkships. The CPX is an 8 station observed structured clinical evaluation (OSCE), from which the authors selected for analysis the same six clinical problems named above. Analysis: history taking, physical examination, and patient-physician interaction components of the OSCE were used to generate CPX scores (Cronbach’s alpha = 0.34 to 0.65) used as the outcome variable. Univariate regression analyses were performed on each case using the number of patients will “full” responsibility on CPX case scores. Independent sample t tests were performed to examine whether there were
group differences in performance between the lowest and highest quartiles based on number of “full” responsibility cases seen.

4. Results: Wide variability in patient encounters by problem existed (5.82 for cough, 21.44 for hypertension). Immersive (“full”) patient encounters did not predict the students’ performance on the CPX. Examination of the lowest to highest quartiles of exposure to comprehensive encounters in relation to the CPX performance were not statistically significant for any of the 6 cases.

5. Limitations: single institution, all patient data were self-reported and a minimum number of listings was required. “Full” responsibility may have been interpreted variably.

6. Implications: These findings replicate others, which have consistently shown the lack of learning effects from patient exposure alone. Quantification of encounters as a process measure is an insufficient surrogate. Further research is needed to understand the quality of medical students’ learning experiences associated with patient encounters which result in learning transfer (2,3) and favorable performance. The complexity of “learning” in clinical settings needs further study.

References

Annotated Articles

This study examined the positive predictive value (PPV) of the In training exam (IM-ITE) examination of PGY-2 residents’ knowledge, on residents’ success in passing the Board certification exam (ABIM-CE). Though previous studies have demonstrated the PPV of the IM-ITE for the ABIM CE, this was done prior to changes made to the ABIM-CE. This study examined whether or not the IM-ITE continued to offer a predictive value for the ABIM-CE. The study examined four IM programs with IM-ITE pass rates less than 100% whose residents graduated in 2000, 2001, 2002, taking the IM-ITE and sitting for the ABIM-CE (n=158). The results were that residents who took the IM-ITE each year or took it only once during their residency, had a 90% or better PPV of passing the ABIM-CE. In comparison to their peer groups, the percentile cutoff ranged between the
The study concludes that the IM-ITE retains its PPV for the ABIM-CE.


The study is a retrospective assessment of the correlation of scores on the clinical skills examination as predictors of complaints to medical regulatory authorities. Physicians in Canada and the United States, (Medical Council of Canada, 1993; Education Commission on Foreign Medical Graduates, 1998; and the United States Medical Licensing Exam, 2004), are now required to have passed a clinical skills examination as a requirement for licensure (Canada) or certification (US). The study uses a cohort study of all 3,424 physicians who took the Medical Council of Canada’s (MCC) clinical skills examination between 1993 and 1996 who were licensed to practice in Ontario and/or Quebec, following them until 2005, examining complaints received between date of licensure and March 2005. Complaints are categorized into 55 (Quebec) or 57 (Ontario) “mutually exclusive categories” and subsequently categorized into two or three groupings including communication and attitude & quality of care. Of the 3,242 physicians, 21.5% had at least 1 complaint filed, and 17.1 % had complaint(s) retained in their file after investigation. The complaint rate was calculated as complaints (numerator) over years in practice (denominator). Low communication scores on the MCC’s clinical skills examination involved condescending, offensive, judgmental behaviors; ignoring patient responses during the encounter. The study found that of the clinical skills examination scores, “only the communication score was significantly associated with complaint rates.”


This study repeated a 1986 survey of procedures done by ACP members. Each survey (1986, 2004) asked internists about the number and variety of procedures done in practice, with 40 procedures named on each survey (34 in common with both surveys; 6 new on 2004 survey). 2500 questionnaires mailed; 2476 had valid addresses and 1389 (56%) were returned. Of the 1389 returned surveys, 96 excluded because physician retired or didn’t complete survey. Of 1293 left, 303 were in a subspecialty, leaving 990 surveys to be analyzed. The study found that the median number of hours spent in patient care remained the same at 50 hours; with the mean in 2004, 4.4 hours lower. Even with the similar number of hours in patient care, “the median number of different procedures done in practice decreased nearly by half between the 2 surveys, from 16 procedures to 7”; and there were significant differences in procedures done by geography (internists in smaller towns and smaller hospitals doing twice as many procedures as internists in
larger town or hospitals). Of the 6 procedures done in the survey, and required by ABIM for certification (lumbar puncture, thoracentesis, central venous catheter placement, arterial puncture for blood gases, abdominal paracentesis, joint aspiration) only one – joint aspiration – was done by a majority of internists. “Except for joint aspiration, the required procedures now were done by one fourth or fewer of the 2004 respondents.”


The study analyzed the submission of a single medical education research study to six different institutions’ IRBs (a proposed multi-institutional study). There were two stages in the review process, a review by an IRB administrator and/or IRB member followed by review by the whole IRB committee. The median time for preliminary review for five of the protocols was 6 days; the sixth protocol’s review took 101 days. Two protocols required full IRB committee review; three were expedited review (the 6th protocol ultimately was not reviewed in time to participate in the study). Time for IRB Committee review ranged from 6 to 115 days with the median number of requests for additional information or changes to the protocol was 13. The length of time in obtaining permission for the study resulted in a loss of statistical power (6th institution did not participate), loss of subjects (delayed the administration of the survey from fall to spring). The authors recommend a “standard electronic application that is appropriate for medical education research.”


The study had two purposes: “(1) develop an instrument to measure the methodological quality of education research studies and determine its reliability and validity, and (2) identify relationships between funding and study quality”. The authors developed a MERSQI (medical education research study quality instrument), a 10 item instrument reflecting study design, sampling, type of data, validity, data analysis and outcomes, each one with a maximum score of 3, total MERSQI score of 18. Using a sample of 210 medical education original research studies (excluding qualitative studies, meta-analyses and systematic reviews), published between 9/1/2002 and 12/31/2003 in 13 peer-reviewed journals, pairs of researchers assessed the articles using the MERSQI. The study established author experience and study funding by surveying the first authors’ of the studies (84% response rate) and the study cost by “multiplying the authors’ percentage of effort dedicated to the study by the national median salary for each author, according to specialty and academic rank, and then adding the cost of resources used”. Of the 210 studies, 149 (71%) did not have any funding, 30 (14%) had between $1 and $19,999, and 31 (15%) had $20,000 or more. The median funding was $20,000 with a median cost of $23,179. The study found that the higher quality medical education
studies, as determined by the MERSQI score, was associated with higher costs, and with experience level of first author and with $20,000 or more in funding.


Citing the Institute of Medicine report, *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*, as demonstrating the lower-quality care received by racial and ethnic minorities, the Society of General Internal Medicine Health Disparities Task Force proposed the development of guidelines for medical education on disparities in health and health care. The Task Force conducted a needs assessment of 60 randomly selected internal medicine program directors (37% response rate); 86% (22) reported “teaching some information about health disparities”. Respondents cited barriers to teaching health disparities were lack of qualified faculty and lack of standardized curricula. The Task Force subsequently conducted a literature review of MEDLINE, EMBASE and Cochrane Library of Systematic Reviews and found no documentation on published curricula. Recommendations were to establish “3 core teaching goals: 1) to help learners examine and understand attitudes, such as mistrust, conscious bias, and stereotyping, that practitioners and patients may bring to the clinical encounter; 2) to impart knowledge of the existence and magnitude of health disparities and the many solutions required to eliminate them; and 3) to provide the learner with the skills required to effectively communicate and negotiate across cultures, languages and literacy levels, including the use of key tools to improve communication.”


The objective of the study was to determine what effect, if any, the changing gender and race/ethnicity of the population of medical students has on the decreasing number of medical graduates planning full-time clinical practice. The study used anonymous American Medical Colleges Graduation Questionnaire (GQ) data from 108,408 U.S. allopathic medical graduates from 1997-2004. The study sample included 94,101 (86.8%) of GQ respondents answering eight items regarding graduates’ demographics, medical school characteristics, and specialty choice as well as answering a question regarding future career plans (full-time university faculty, other, undecided or full-time clinical practice). Overall, the authors confirmed the trend of graduates’ decreasing selection of full-time clinical practice as a career choice (from 51.3% in 1997 to 46.5% in 2004). Graduates reporting decreasing levels of debt, dual-degree graduation, being Hispanic, being Asian/Pacific Islander, not indicating race/ethnicity, and being female were more likely to plan a career not involving full-time clinical practice, or being undecided.