REVIEW

Update in Medical Education

Carol K. Bates, MD¹, Shobhina G. Chheda, MD, MPH², Kathel Dunn, MSLS³, Linda Pinsky, MD⁴, and Reena Karani, MD⁵

¹Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, MA, USA; ²Department of Medicine, Section of General Internal Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA; ³National Library of Medicine, Bethesda, MD, USA; ⁴University of Washington School of Medicine, Seattle, WA, USA; ⁵Brookdale Department of Geriatrics, Samuel M. Bronfman Department of Medicine, Department of Medical Education, Mount Sinai School of Medicine, New York, NY, USA.

KEY WORDS: feedback; evaluation; knowledge assessment; retention; curricula; Minicard.

J Gen Intern Med 25(5):465–9

DOI: 10.1007/s11606-009-1244-x

© Society of General Internal Medicine 2010

In this article, we summarize 10 manuscripts that represent important advances in medical education in 2008 from a general internal medicine perspective.

METHODS

We limited our search to articles published between January 1, 2008 and December 31, 2008 in one of 12 publications: Academic Medicine, Annals of Internal Medicine, British Medical Journal, Journal of the American Medical Association, Journal of the American Geriatrics Society, Journal of Continuing Education in the Health Professions, Journal of General Internal Medicine, Lancet, Medical Education, Medical Teacher, New England Journal of Medicine, and Teaching and Learning in Medicine. Previous literature review for the best medical education articles consistently found that the above journals contained the medical education articles most relevant to academic general internists and those articles most likely to impact their teaching and practice. The authors' knowledge of trends and issues in medical education additionally shaped the review process. We defined "medical education" broadly and included all study designs and levels of learners. More than 1300 citations were examined. We reached consensus on our final articles based on the questions "Is this relevant to academic general internists?" and "Will this article change the practice of teaching and learning?"

RESULTS

We chose ten articles that fall into four themes: (1) feedback and evaluation, (2) knowledge assessment and retention, (3) curricula, and (4) continuity clinic.

FEEDBACK AND EVALUATION

Donato AA, Pangaro L, Smith C, Rencic J, Diaz Y, Mensinger J, Holmboe E. Evaluation of a novel assessment form for observing medical residents: A randomized, controlled trial Medical Education 2008;42:1234–42.

The ABIM Mini-CEX is a validated tool when compared with the ABIM monthly evaluation and in-training exam but had limited ability to predict unsatisfactory clinical performance in a study using standardized videotapes¹. The format does not allow for recording of specific comments and action plans. This randomized controlled study of 80 faculty from four different internal medicine (IM) residency programs examined if a redesign of the ABIM Mini-CEX (the Minicard) would more accurately distinguish satisfactory from unsatisfactory clinical performance when compared with the ABIM Mini-CEX and, if the redesigned card increased the quantity and quality of intended feedback and written observations. The Minicard preserved the Mini-CEX direct observation format but condensed seven ACGME domains into three ACGME competencies, reduced the score range to a 4-point scale adding behavioral and adjectival anchors, and allowed space for specific comments. Faculty were randomized to use of the Minicard or the Mini-CEX, and each received one-hour of training. In one hour of training, faculty practiced rating videotaped trainee standardized interactions with patients using their respective forms. Testing took place 2–3 weeks later when faculty viewed and rated 6 videotapes of satisfactory and unsatisfactory examples of a resident taking a history, conducting an exam and counseling a patient. Videotapes were identical to those used in prior studies of the Mini-CEX¹. Faculty were asked to document on the form and also provide feedback that they would have given the resident on a separate page (intended feedback). Accuracy was calculated and number of comments as intended feedback and comments on forms were counted. Results showed that participants using Minicards classified performances correctly 85% of time compared with 73% for Mini-CEX. Minicard users correctly identified 96% of unsatisfactory performances compared with 52% by Mini-CEX users. However, Minicard users correctly classified only 73% of satisfactory performances compared with 95% classified by Mini-CEX users. Minicard users recorded twice the total number of observations compared to Mini-CEX users.

While this study was limited by the lack of blinding of study subjects and faculty trainers, it demonstrates that with one hour

Received July 27, 2009

Revised December 2, 2009

Accepted December 28, 2009

Published online February 12, 2010

of faculty training, a redesigned Mini-CEX card produced a larger number of recorded observations and better overall accuracy by substantially increasing identification of unsatisfactory performances and by increasing reliability at the pass fail level (0.520 versus 0.299). As we consider the importance of establishing a minimum resident clinical competence, the trade-off of sensitivity for specificity is acceptable allowing earlier identification of unsatisfactory performance. The increased number of comments made on the Minicards should allow for clearer resident and program director understanding of resident performance and hence facilitate efforts to improve resident clinical skills.

Kogan JR, Shea JA. Implementing feedback cards in core clerkships. *Medical Education* 2008;42:1071-79.

This study sought to determine the feasibility of a cross-clerkship weekly feedback encounter card system, describe the content of feedback requested, and examine third year student satisfaction with the card system at one US medical center. The encounter card contained a section for students to request feedback on any of eight content areas; a section for faculty or residents to indicate those areas in which they provided feedback; and a section to document an action plan. One-hundred and twenty seven third year medical students were required to complete cards during six core clerkships (one per week in outpatient settings and two per week in inpatient settings). Additionally, a survey using a 5-point Likert-scale was used as part of an end-of-clerkship evaluation, to rate the usefulness of the cards; overall usefulness of feedback in the clerkship (not necessarily specific to the card); and the overall clerkship experience. 5369 feedback cards (78% of target) were completed. Students requested feedback in a median of three content areas per card. Most frequently requested content areas for feedback were presentation skills (57%, 3060 cards) and fund of knowledge (48%, 2577 cards). Least frequently requested content areas for feedback were physical examination skills (26%, 1359 cards) and counseling (13%, 698 cards). As indicated by completion of the faculty/resident section on cards, students received feedback in the areas they requested on 4295 or 80% of the cards. Despite evidence that students received feedback, overall ratings of usefulness of feedback (not related to cards) was statistically lower on a 1-5 scale (1=poor, 5=excellent) when compared with a historical control sample of 126 students from 2005, (mean 3.6, SD 1.1 vs. mean 3.7, SD 1.0; $p < 0.001$, effect size = -17).

Limitations are that this study did not measure feedback specificity, quality or type and did not assess student skill or performance following feedback. This study demonstrated the feasibility of a feedback card system, and highlighted that there were no differences in type of feedback requested in different clerkships or requested of faculty vs. residents. Additionally, it revealed the concerning finding that students generally did not request feedback requiring direct observation of patient encounters. As faculty we should be cautious in only providing feedback in areas in which students seek feedback as important skills such as physical examination and counseling skills may remain undeveloped in students. Additionally course directors should work to develop other measures of feedback effectiveness than student satisfaction.

Stark R, Korenstein D, Karani R. Impact of a 360-degree professionalism assessment on faculty comfort and skills in feedback delivery. *Journal of General Internal Medicine* 2008;23:969-72.

This single institution study evaluated whether implementation of an on-line 360-degree professional assessment

instrument from the National Board of Medical Examiners would impact frequency of professionalism feedback to internal medicine (IM) residents by 15 IM faculty advisors and whether these faculty would report any change in comfort and self-assessed skill in delivering feedback. The 360-degree professionalism assessment of residents was implemented during a 4-week ambulatory IM continuity clinic rotation. Faculty, receptionists, nurses and medical assistants evaluated residents with whom they worked. All evaluations were compiled into an "Individual Summary Report" which was provided to the faculty advisor with written instructions to use the report for formative feedback to the resident. Faculty advisors provided feedback at the end of the ambulatory rotation. Faculty advisors completed a survey prior to and 6 months post initiation of the comprehensive professionalism assessment. Completion rates were 82% from medical assistants, 76% from faculty, 73% from nurses and 48% from receptionists at six months. All 15 internal medicine faculty advisors completed pre-post self assessments. Results on a 7-point Likert scale showed increased faculty self-reported skill in giving both general (4.4 to 4.9 $p < 0.05$) and professionalism related (3.6 to 4.7 $p < 0.01$) feedback. Though they reported no increase in comfort in giving general feedback (5.50 at baseline), faculty did report increased comfort in giving professionalism feedback (4.3 to 5.1 $p < 0.05$).

This study was limited by the absence of controls and because skill in providing feedback was self assessed rather than externally validated. However, this study is important in its focus on the difficult assessment of professionalism and its demonstration of increased self-assessed faculty skill in addition to comfort in giving feedback. If both faculty skill and comfort can be enhanced it is more likely that residents will receive more meaningful feedback regarding professionalism.

KNOWLEDGE ASSESSMENT AND RETENTION

Kennedy TJT, Regehr G, Baker GR, Lingard L. Point-of-care assessment of medical trainee competence for independent clinical work. *Academic Medicine* 2008;83:S89-92.

This study explored attending physician's (AP) assessments of trainees' competence to provide independent clinical care and the process employed to make these assessments in the inpatient setting. General Internal Medicine and Emergency Medicine faculty, residents and students in three affiliate hospitals of an urban Canadian medical school participated in the study. In phase 1, investigators analyzed 216 hours of audio recordings of faculty/resident interactions. Audio recordings were transcribed and analyzed for emergent themes using grounded theory and a constant comparison method of analysis. Faculty were then interviewed one year later using 10 videotaped vignettes based upon events brought up in Phase 1 observations and crafted to present a dilemma relevant to decisions about supervision. Results revealed that AP consider four dimensions (knowledge and skills, discernment of limitations, truthfulness and conscientiousness) when making decisions about how much independence to allow trainees. Processes used by faculty to make these judgments included double checking of trainees information with data obtained by others, and language cues such as the structure and delivery of information and the presence of anticipated information in the trainee's case presentations.

Limitations of this study are that the behavior may have been influenced by the observer (Hawthorne effect) and that results may not be transferable to other venues of care. However, this study does suggest that AP consider more than clinical knowledge and skill when deciding how much supervision to provide. Concepts of discernment, conscientiousness, and truthfulness as well as clinical skill all go into the decision making process and AP double check and use language cues to inform their assessment of trainees. Information from this study could potentially form the basis for faculty development efforts to learn how to assess a trainee's readiness for independent practice.

Bell DS, Harless CE, Higa JK, Bjork EL, Bjork RA, Bazargan M, Mangione CM. Knowledge retention after an online tutorial: A randomized educational experiment among resident physicians. *Journal of General Internal Medicine* 2008;23:1164-71.

The study purpose was twofold: to assess time elapsed in retention of resident physicians' knowledge following an online tutorial on diabetes care and to determine if learner characteristics, including critical appraisal skill and self-efficacy, affect learning and retention. An online tutorial based on ADA guidelines for managing hypertension and lipids in patients with diabetes was delivered to 91 family medicine and internal medicine residents from two academic medical centers. Participating residents who completed the tutorial were randomly assigned to one of six follow-up intervals for taking the post-test: 0, 1, 3, 8, 21, or 55 days. 91 of 197 (46%) invited residents completed the tutorial, with 87 (96%) providing complete follow-up data. In the post-test, time delays were associated with lower scores: by days 3-8, retention was reduced by 50% and by day 55, performance was equivalent to the pre-test mean. There were no data to support the hypothesis that learning self-efficacy and critical appraisal skills had an impact on higher retention of material presented in the tutorial.

Study limitations are that the authors did not determine the service residents were on at the time that they participated in the study and that other confounders may have reduced long term retention of knowledge. Slightly less than half of the eligible resident physicians chose to participate by taking the tutorial, which may have skewed the pool; however the 96% follow-up provides strong evidence for the validity of the data. In summary, resident physicians' recall of knowledge after an online tutorial is largely lost after a short period of time. Therefore, reinforcement of knowledge through review or practice is recommended to facilitate knowledge retention.

CURRICULA

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. *JAMA* 2008;300:1181-96.

With Internet-based instruction, a learner can choose the time and location for learning; and with some customization, the instructional module—tutorial, video, email and so on—can be tailored to the individual learner. This meta-analysis of studies of Internet-based instruction with health professionals as learners included 201 (of 288) eligible articles with sufficient data to analyze. The articles eligible for analysis included studies that compared either Internet-based instruction against no educational intervention or Internet-based instruction compared with face-to-face instruction, paper, satellite-

mediated videoconferences, standardized patients and slide-tape self-study modules. Articles were excluded if there was no intervention or a non-Internet intervention; or if there were qualitative outcomes only, no relevant quantitative outcomes or duplicate publication. Study result means and standard deviations were converted to Hedges *g* effect sizes which were then pooled across each intervention. Pooled effect size favors the intervention if a positive number and is considered to be large if $> 0.8^2$. Pooled effect size comparing studies with no intervention versus an Internet-based intervention favored Internet-based interventions for knowledge outcomes (1.00), skills (0.85), and learner behaviors and patient effects: (0.82) with $p < 0.001$ for each of these three outcomes. In contrast, pooled effect size in comparison to other educational interventions showed no effect or non-effect compared to Internet-based interventions for knowledge (0.12 $p = 0.045$), skills (0.09 $p = 0.61$), and learner behaviors and patient effects: (0.51 $p = 0.18$).

Compared with no educational intervention, Internet-based instruction is thus an effective instructional method, but Internet-based instruction is neither more nor less effective than alternative educational instructional methods. This suggests that educators may wish to use Internet-based learning for reasons of efficiency, cost, time, and the ability to provide asynchronous learning, but that they should not be compelled to do so to optimize actual learning. It also suggests that future investigations of Internet-based learning need not compare outcomes against control of no intervention but that the focus should rather be on web versus conventional learning in specific content areas and on the particulars of the web methodology.

Kerfoot BP, Armstrong EG, O'Sullivan PN. Interactive spaced-education to teach the physical examination: a randomized controlled trial. *Journal of General Internal Medicine* 2008;23:973-78.

Learners have limited physical examination knowledge and skills. "Interactive spaced education" (ISE) repeatedly tests and educates learners over spaced intervals using e-mail delivery of learning points and questions. "Spaced" education may lead to greater learning retention than "bolus" education; immediate testing on learning may also bolster learning retention. This study examined ISE as a method to further enhance physical examination learning in second year students taking a course on physical diagnosis. 170 2nd year medical students in the Introduction to Clinical Medicine course at one academic institution were invited to participate. This course included weekly/biweekly 4-8-hour sessions on physical examination. Participant students additionally received three cycles of 36 emails which presented a physical diagnosis question. After answering the question, subjects were presented with the correct answer and associated curriculum. Of 120 students initially randomized into two cohorts, 85 completed all cycles. Scores increased from mean of 57.9% in cycle 1 to 74.4% in cycle 3. Of questions answered incorrectly in cycle 1, 64.7% and 73.4% were answered correctly in cycles 2 and 3. Although all participants received 3 cycles of questions via email, the second cohort started cycle 1 of 3 at the same time that the first group started cycle 3 allowing an interim use of the second group as a control. Cycle 3 scores for cohort A were 74% vs. 59% cycle 1 scores for cohort B when thus assessed simultaneously. Each email was estimated to require a mean of 2.7 minutes to complete. At the end of the course, students concluded that the optimal number of emails was 4.9 per week

with optimal number of cycles 2.7. 85% of learners recommended the same course be repeated next year for incoming 2nd year students; 83% wanted an ISE program for themselves next year as 3rd years.

Although this study assessed physical examination knowledge and did not assess examination skills, it appears that ISE is a well accepted method of curriculum delivery that can help to remediate knowledge deficits. These results also suggest that duplication of curricular material is accepted by learners and can consolidate learning.

Lane C, Hood K, Rollnick S. Teaching motivational interviewing: Using role play is as effective as using simulated patients. *Medical Education* 2008;42:637-44.

Optimal methods of training in motivational interviewing have been unclear. This study compared role play versus standardized patients in training practicing physicians in motivational interviewing. Seventy practicing physicians who had enrolled in a course on motivational interviewing were block randomized into standardized patient versus role-playing workshops for 2 days. Interviews with standardized patients were done at baseline and after training to assess skills before and after training. Skills were rated using the Behavioral Change Counseling Index (BECCI) a validated 11-item tool rating skills on a Likert scale where the maximal score is 44, but the ideal score has not been determined^{3,4}. Over the 2-day workshop, groups practiced skills in three sessions either with standardized patients (intervention) or controls (role playing). Scores on the BECCI improved in both the experimental (14.0 to 17.0; Wilcoxon's signed rank test $z = -4.39$ $p < 0.001$) and control (11.50 to 16.0; Wilcoxon's signed rank test $z = -2.27$ $p < 0.02$) groups, but there were no significant differences in the degree of change in either group (ANCOVA $F = 1.13$, $P = 0.29$). Results might have been influenced by the fact that all learners were seeking this training, and outcomes measured behaviors with standardized and not with real patients.

The study is limited by its small sample size and by statistically significant differences in skill level measured by BECCI prior to the intervention; it may be this difference that prompted the comparison of change in scores in each group and the choice to name this measure as the main outcome of the study. This study does suggest that role play might be as effective as the much more expensive cost of standardized patients in teaching motivational interviewing to practicing physicians. If replicated across other interviewing and counseling skills, this would significantly reduce educational program costs.

CONTINUITY CLINIC

Warm EJ, Schauer DP, Diers T, Mathis BR, Neirouz Y, Boex JR, Rouan G. The ambulatory long-block: an accreditation council for graduate medical education (ACGME) educational innovations project. *Journal of General Internal Medicine*. 2008;23:921-26.

Under the auspices of the Education Innovation Project of the Internal Medicine Residency Review Committee, this educational intervention redesigned the categorical Internal Medicine resident ambulatory practice at the University of Cincinnati Academic Health Center. Investigators created the ambulatory long block, a year-long continuous ambulatory group-practice experience separated from traditional inpatient

responsibilities, to improve the ambulatory experience both for the residents and the patients. The long block was associated with significant increases in resident satisfaction. On a 5-point Likert scale, pre- and post-assessments by learners improved as follows: time for learning (2.94 to 4.44 $p = 0.0004$); ability to focus in clinic without interruption (3.44 to 4.56 $p = 0.0057$); ability to balance ward/inpatient duties on clinic days (3.00 to 4.59 $p = 0.0018$); overall satisfaction with the learning environment (3.65 to 4.24 $p = 0.0075$); overall satisfaction with the clinical environment (3.44 to 4.33 $p = 0.156$); personal reward from work (3.33 to 4.44 $p = 0.0042$); and relationships with patients (4.06 to 4.72 $p = 0.0001$). Patient satisfaction improved as measured by Press-Ganey scores. Continuity of care was improved and the "no show" rate decreased from 28% to 18%. Improved quality was noted in chronic illness care such as diabetes: as measured by the rates of foot exam (35.7% to 59.7% $p < 0.001$), use of aspirin prophylactically (75.2% to 86% $p < 0.001$), the administration of pneumovax vaccine (70.9% to 78.5% $p = 0.003$), and achievement of goal BP < 130/80 in diabetics (37.9% to 47.1% $p = 0.002$). Preventive screening measures such as mammography (41.6% to 63.6% $p < 0.001$), colonoscopy (36.4% to 48.6% $p < 0.001$), Pap smears (7.7% to 61.7% $p < 0.001$), PSA (34.4% to 51.7% $p < 0.001$), and bone density screening (10.1% to 54.2% $p < 0.001$) and various immunization rates also increased.

It remains unclear whether it was the continuous nature of the long block, the separation of inpatient versus outpatient training, or the multidisciplinary team approach to care that led to these outcomes. Additional reports from other programs that have similarly separated inpatient from outpatient care should answer these questions in the future. Although the strength of this study is limited by the lack of controls, description and outcomes of innovative restructuring of the outpatient clinic experience informs the future of residency ambulatory training.

Oyler J, Vinci L, Arora V, Johnson J. Teaching internal medicine residents quality improvement techniques using the ABIM's practice improvement modules. *Journal of General Internal Medicine* 2008;23:927-30.

Standard curricula to teach Internal Medicine residents about quality assessment and improvement, important components of the Accreditation Council for Graduate Medical Education core competencies practiced-based learning and improvement and systems-based practice, have not been easily accessible. These educators incorporated a longitudinal quality assessment and improvement curriculum based on the American Board of Internal Medicine's (ABIM) Clinical Preventative Services Practice Improvement Module (CPS PIM), into the 2 required 1-month ambulatory rotations during the postgraduate year 2. The first rotation focused on PIM chart reviews, patient and system surveys; the second focused on resident initiated quality improvement (QI) projects in the resident continuity clinic. The percent of residents reporting that they were comfortable writing a clear aim statement improved from 71% (24 of 34 residents) to 96% (27 of 28 residents; $p < 0.01$, chi-square test). The percent of residents reporting that they were comfortable using a Plan-Do Study-Act (PDSA) cycle improved significantly from 9% (3 of 34 residents) to 89% (25 of 28 residents; $p < 0.001$, chi-squared test). Successful group QI projects and their outcomes include a focus on measurement of BMI where notation of BMI went from 4% to 79% of charts ($p < 0.001$), tobacco cessation where documentation of smoking status rose from 41% to 67% of charts ($p < 0.001$),

and a refill collaborative where inaccurate medication lists were reduced from 25% to 9% of charts ($p < 0.001$). Residents disseminated their projects as scholarly work in an internal resident research day, a hospital quality fair, a monthly QI in Progress seminar series, and at a regional Internal Medicine meeting.

Implementation of this curriculum requires faculty development, curricular space, and the payment to ABIM of \$25 per resident. This innovation represents an easily exportable model for QI initiatives in many domains.

Acknowledgements: *This work was presented as an Update in Medical Education oral session at the 2009 annual meeting of the Society of General Internal Medicine in Miami, FL. This work was presented at the 2009 meeting of the Society of General Internal Medicine*

Conflict of interest None.

Corresponding Author: Carol K. Bates, MD, Beth Israel Deaconess Medical Center, Healthcare Associates, Sh 607, 330 Brookline Avenue, Boston, MA 02215, USA (e-mail: cbates@bidmc.harvard.edu).

REFERENCES

1. Holmboe ES, Huot S, Chung J, Norcini J, Hawkins RE. Construct validity of the Mini Clinical Evaluation Exercise (MiniCEX). *Acad Med.* 2003;78:826–30.
2. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum; 1988.
3. Behaviour Change Counselling Index (BECCI) <http://www.motivationalinterview.org/library/BECCIForm.pdf> (accessed 9/5/09)
4. Personal correspondence with author 11/15/09 Claire Lane clairelane1978@gmail.com