A Human Factors Curriculum for Surgical Clerkship Students

Mitchell A. Cahan, MD; Anne C. Larkin, MD; Susan Starr, MEd; Scott Wellman, MD; Heather-Lyn Haley, PhD; Kate Sullivan, BA; Shimul Shah, MD; Michael Hirsh, MD; Demetrius Litwin, MD; Mark Quirk, EdD

Hypothesis: Early introduction of a full-day human factors training experience into the surgical clerkship curriculum will teach effective communication skills and strategies to gain professional satisfaction from a career in surgery.

Design: In pilot 1, which took place between July 1, 2007, and December 31, 2008, 50 students received training and 50 did not; all received testing at the end of the rotation for comparison of control vs intervention group performance. In pilot 2, a total of 50 students were trained and received testing before and after rotation to examine individual change over time.

Setting: University of Massachusetts Medical School.

Participants: A total of 148 third-year medical students in required 12-week surgical clerkship rotations.

Interventions: Full-day training with lecture and small-group exercises, cotaught by surgeons and educators, with focus on empathetic communication, time management, and teamwork skills.

Main Outcome Measures: Empathetic communication skill, teamwork, and patient safety attitudes and self-reported use of time management strategies.

Results: Empathy scores were not higher for trained vs untrained groups in pilot 1 but improved from 2.32 to 3.45 on a 5-point scale (P < .001) in pilot 2. Students also were more likely to ask for the nurse’s perspective and to seek agreement on an action plan after team communication training (pilot 1, f = 7.52, P = .007; pilot 2, t = 2.65, P = .01). Results were mixed for work-life balance, with some trained groups scoring significantly lower than untrained groups in pilot 1 and no significant improvement shown in pilot 2.

Conclusions: The significant increase in student-patient communication scores suggests that a brief focused presentation followed by simulation of difficult patient encounters can be successful. A video demonstration can improve interdisciplinary teamwork.

Arch Surg. 2010;145(12):1151-1157

Mounting data indicate that the field of surgery has an increasing shortage of general surgeons.1 There is also evidence that practicing surgeons need to communicate more effectively with their patients and team members, learn skills to prevent medical errors, and increase satisfaction with their own professional work.2-5 Students undergoing surgical clerkships become aware of the frustration of patients, residents, and attending physicians via the “hidden” or informal curriculum, and many who might have selected surgery as a specialty choose other specialties instead.6

Methods

Curriculum

The curriculum was presented as a one-day “Introduction to Human Factors” at the beginning of each block. The final framework, designed to be cotaught by a surgeon and a medical educator, is given in Table 1. The “Pa-
PATIENT COMMUNICATION: EMPATHY AND CARING

The “Patient Communication: Empathy and Caring” session consisted of a 45-minute interactive presentation regarding eliciting the patient’s perspective, giving bad news, and reducing patients’ fear and anger. Students then interacted with 3 actors working as simulated patients and were required to (1) calm a patient with a wound infection, (2) reassure a patient with a postsurgery deep venous thrombosis, and (3) give news of a ruptured appendix that occurred while the patient waited in the emergency department. After each simulation, the trained surgeon and education faculty teams had a debriefing session focusing on the skills emphasized in the presentation.

During “Communicating With Team Members,” students watched a 7-minute video produced expressly for this project in which a nurse and senior resident had different degrees of anxiety about a postoperative patient. Their poor communication regarding the patient fueled the patient’s anxiety and led to the involvement of the chief of surgery. An extensive debriefing session followed.

The “Work-Life Balance” session focused on 3 skills: learning the most while in the hospital setting (“Stick to your resident like glue”), completing the required reading (“Prioritize what you need to know tomorrow”), and balancing work and life (“Sleep when you can sleep; eat when you can eat; call home when you can call home.”). Satisfaction scores for the “Communicating With Team Members” and “Work-Life Balance” curricula were evaluated together.

Participating in the Study

One hundred students rotate through four 12-week blocks during each academic year. A visual summary of the research design can be seen in the Figure. Students in the intervention group were asked to complete a “Commitment to Change” form and an evaluation of the program immediately after the training session. The control and intervention groups completed a series of human factors evaluation measures on the last day of the clerkship. Data for the 2 intervention blocks were combined for the purpose of analyses; the same was done for data from the 2 control blocks. The data were later split again to examine differences by block.

Pilot 1

In 2007, a total of 148 third-year medical students received training. The full-day human factors curriculum was presented to students in block 2 and block 4. Students in blocks 1 and 3 received the standard curriculum without human factors training. Students in the intervention group were asked to complete a “Commitment to Change” form and an evaluation of the program immediately after the training session. The control and intervention groups completed a series of human factors evaluation measures on the last day of the clerkship. Data for the 2 intervention blocks were combined for the purpose of analyses; the same was done for data from the 2 control blocks. The data were later split again to examine differences by block.

Pilot 2

In 2008, 50 students received training. Students in block 1 and block 2 received the training and completed the same series of evaluation measures before training and at the end of the clerkship.

EVALUATION MEASURES

PATIENT COMMUNICATION

Skills in communicating with patients were evaluated using a series of 4 brief video vignettes, each presenting a communication challenge included in the human factors curriculum (eg, delivering bad news). For example, students are presented with a mother awaiting news of her son, who had been in an auto-

<table>
<thead>
<tr>
<th>Table 1. Surgery Human Factors Clerkship Curriculum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Session</strong></td>
</tr>
<tr>
<td>Communicating with patients: empathy and caring</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Communicating with team members in the surgical setting</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Balancing work and life</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Figure.* Research design and methods for 2 pilot years of the human factors curriculum.
mobile collision. The students were given the information that the son has been diagnosed with brain death, then shown a video of the anxious mother asking, “How’s Luis?” Students were given 90 seconds to respond in writing to each vignette.

**Communicating With Team Members/Work-Life Balance**

Students were asked how they would handle a situation in which their resident and the nurse gave conflicting directions. Next, they were asked to provide descriptions of strategies they found useful for (1) managing time spent in the hospital, (2) finding time to read and study, and (3) balancing their professional and personal lives.

**STATISTICAL ANALYSES**

**PATIENT COMMUNICATION**

Responses to the vignettes were randomly sorted and masked as to participation group, then coded using a previously developed 5-point scale (Table 2). Two of us (H.-L.H. and K.S.) coded 60.2% of the responses to determine interrater reliability. One of us (K.S.) coded all responses used in the analyses. Mean scores were calculated for each student’s performance across all 4 cases.

**COMMUNICATING WITH TEAM MEMBERS**

Four main points were identified as desirable in a response to the teamwork scenario (Table 2). Three of us (S.S., H.-L.H., and K.S.) jointly used this scoring system to rate a random sample of 10.0% of the cases. An agreement rate of 90.0% between raters was reached on the first trial of this system, so the system was finalized and used by 1 of us (K.S.) to rate the remaining responses.

**WORK-LIFE BALANCE**

Key elements of an excellent response were identified for each of 3 questions: assessing students’ strategies for managing time in the hospital (4 possible points), finding time for reading and studying (4 possible points), and balancing work and personal/family responsibilities (3 possible points). Three of us (S.S., H.-L.H., and K.S.) tested the rubric on a random sample of 10.0% of the responses from pilot 1, masked as to whether each response was given by a student from the control or intervention group. After discussion, alterations were made to the rubric and another sample scored; this evaluation loop was conducted 3 times until the 3 scorers reached 80.0% agreement on scores, at which time 2 of us (K.S. and H.-L.H.) completed the scoring of all responses. Table 2 contains the scoring rubric for these items.

Perceptions of the program were calculated by obtaining means and standard deviations to questions regarding the effectiveness of the program, clarity of objectives, and effectiveness of the instructors. The rating scale ranged from 1 (extremely ineffective) to 5 (extremely effective). Commitments to change were qualitatively assessed and combined into categories (see Table 3 for representative commitments).

**RESULTS**

**STUDY PARTICIPANTS**

**Pilot 1**

Ninety-seven students participated in the first year of this study. Forty-eight students attended the human factors training session in alternating rotation blocks: 25 in block 2 and 23 in block 4. Control blocks were composed of 23 students in block 1 and 26 students in block 3 for a total of 49 clerks in the control group.

---

**Table 2. Scoring Rubric**

<table>
<thead>
<tr>
<th>Possible Points</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Patient communication (5 items)</td>
</tr>
<tr>
<td>1</td>
<td>Fire the warning shot = bad news; offer</td>
</tr>
<tr>
<td></td>
<td>hope = palliative care; offer appropriate</td>
</tr>
<tr>
<td></td>
<td>apology = medical error</td>
</tr>
<tr>
<td>1</td>
<td>Patient perspective or naming emotion (physician or patient perspective)</td>
</tr>
<tr>
<td>1</td>
<td>Ask an open-ended question (not just yes or no)</td>
</tr>
<tr>
<td>1</td>
<td>Action plan, follow-up, addresses need for more information</td>
</tr>
<tr>
<td>1</td>
<td>Effective communication: no poor phrasing, weird or awkward statements, jargon, objectifying body parts (“the stomach” vs “your stomach”)</td>
</tr>
<tr>
<td>4</td>
<td>Working with a team (4 items)</td>
</tr>
<tr>
<td>1</td>
<td>Move conversation out of patient hearing</td>
</tr>
<tr>
<td>1</td>
<td>Ask nurse’s perspective</td>
</tr>
<tr>
<td>1</td>
<td>Explain student perspective (resident gave directions)</td>
</tr>
<tr>
<td>1</td>
<td>Show respect, negotiate a compromise, and ask whether nurse is okay with final plan</td>
</tr>
<tr>
<td>11</td>
<td>Work-life balance (3 subscales)</td>
</tr>
<tr>
<td>4</td>
<td>Subscale 1: managing time in the hospital (4 items)</td>
</tr>
<tr>
<td>1</td>
<td>Look for learning opportunities: learn from nurses and other staff; stick to your resident, especially at night, be available</td>
</tr>
<tr>
<td>1</td>
<td>Prioritize: Make to-do lists, multitask, read what you need to know first</td>
</tr>
<tr>
<td>1</td>
<td>Solicit direction: ask what’s coming up, understand role, seek advice</td>
</tr>
<tr>
<td>1</td>
<td>Anticipate: carry readings with you, read while you wait, know what the day’s activities are, think ahead, plan</td>
</tr>
<tr>
<td>4</td>
<td>Subscale 2: finding time to read and study (4 items)</td>
</tr>
<tr>
<td>1</td>
<td>Recognize opportunities: use downtime appropriately, ask questions</td>
</tr>
<tr>
<td>1</td>
<td>Set and work on short-term goals</td>
</tr>
<tr>
<td>1</td>
<td>Plan for and work toward long-term goals</td>
</tr>
<tr>
<td>1</td>
<td>Study skills: study weaknesses, read actively, compare and contrast illnesses, organize information in your head</td>
</tr>
<tr>
<td>3</td>
<td>Subscale 3: balancing professional and personal life (3 items)</td>
</tr>
<tr>
<td>1</td>
<td>Identify factors that are urgent or important that go into decisions</td>
</tr>
<tr>
<td>1</td>
<td>Active decision making: plan and schedule based on prioritizing, make active decisions not passive decisions</td>
</tr>
<tr>
<td>1</td>
<td>Tell others (eg, family and friends) about your priorities, tell them to be realistic about their expectations of you</td>
</tr>
</tbody>
</table>

*Subscale 1 is worth 4 points; subscale 2, 4 points; and subscale 3, 3 points.*
Fifty students participated in the intervention in the second year of the pilot. Forty-four of the students, 24 from block 1 and 20 from block 2, completed all the evaluation measures at pretraining and postcourse assessments. Only students with complete data at both time points are included in the results.

**Evaluation Measures**

In pilot 1, trained and untrained students were tested on the last day of the clerkship. In pilot 2, because all students were trained, a pretest was administered before the first module on the day of the training session and a posttest was given at the end of the block. Improvement was measured within each block and by comparing results across blocks.

### Patient Communication

#### Pilot 1

Three hundred eighty-two responses to vignettes were analyzed, 190 from control students and 192 from intervention students. Interrater reliability was determined by comparison of 288 vignette scores by 2 raters (H.-L.H. and K.S.), with a Cronbach α of .808. Scores were averaged across 4 vignettes for each student for a total of 97 mean scores, with 1 per student. The mean score for control responses was 2.84 (SD, 0.55) (on a scale of 1-5, with 5 being most favorable). The mean score for intervention responses was 2.85 (SD, 0.53).

Differences among the 4 blocks were evident when we performed a 1-way analysis of variance (ANOVA) with multiple comparisons tested using Bonferroni post hoc analysis. The first intervention group, block 2, received the highest score (mean, 3.11; SD, 0.42). The 2 control blocks, blocks 1 and 3, had similar scores. The block 1 mean (SD) of 2.85 (0.50) and the block 3 mean (SD) of 2.84 (0.59) indicate performance below the trained block 2, although not significantly so. Block 4, the final trained block, received the lowest scores, significantly lower than those of the other trained block (mean [SD], 2.58 [0.51]; mean difference, 0.53; P = .003). These results are summarized in Table 4.

#### Pilot 2

Each block rotation received the training and completed the empathy vignettes before training and at the end of the quarter, with 44 students completing premeasures and postmeasures. Because of the shift in design, we can look at individual improvement using paired t tests rather than the
ANOVAs required for the pilot 1 data. The mean pretraining score was 2.28 (SD, 0.61) at block 1 and 2.38 at block 2 (0.78), with a mean pretest score of combined blocks 1 and 2 of 2.32 (0.69). Posttest scores showed a significant improvement in performance in each block and when combined; the combined posttest score was 3.45 (SD, 0.44). These results are summarized in Table 4.

COMMUNICATING WITH TEAM MEMBERS

Pilot 1

The trained clerks received significantly higher scores as determined by a 1-way ANOVA on the portion of the evaluation measuring teamwork, scoring a mean of 1.72 vs 1.34 for the control group (t=7.5, P=.007). Trained clerks were more likely to ask for the nurse's perspective and to seek agreement on an action plan.

Pilot 2

Posttraining scores (mean [SD], 1.80 [0.93]) were significantly higher than pretest scores (mean [SD], 1.33 [0.89]) when evaluated using the paired t test (t=2.65, P=.01). Again, trained clerks were more likely to ask for the nurse's perspective and to seek agreement on an action plan.

WORK-LIFE BALANCE

Pilot 1

The control group received higher scores for 2 of the 3 measures of work-life balance. However, the difference as tested with a 1-way ANOVA was significant only for the item measuring effective use of time in the hospital (1.67 vs 1.13, f=12.39, P=.001).

Pilot 2

Work-life balance scores did not improve significantly from before to after testing in paired t tests for either block individually or combined. The only increase in scores was seen in the item measuring use of time in the hospital (1.22 vs 1.33 after testing, on a scale with 4.00 possible points), although the change was not significant. There was a significant decrease in scores on the item specifically measuring work-life balance when tested with a paired t test (0.97 before vs 0.74 after testing, t=2.76, P=.009, on a scale with 3.00 possible points). There was a significant correlation between pretest and posttest scores at the item level and combined (combined score correlation of 0.46, P=.001), indicating consistency over time of self-perceived ability to manage time. It was later determined that the wording of these items was less than ideal; measurement of this construct will be revisited.

Perception of the Program

The overall mean (SD) scores for patient communication and communicating with team members combined with work-life balance workshops across both years were 3.77 (0.82) (N=91) and 3.96 (0.81) (N=95), respectively. For patient communication, the workshop means ranged from 3.52 to 3.92. For communicating with team members and work-life balance, the workshop means ranged from 3.80 to 4.17.

Comments included the following: “I hadn’t really even considered this as a facet of surgery coming in, and it was good to be reminded of the fact that delivering serious or even negative information will be a major part of the field” and “I found the scenario where the patients experience an appendix rupture after 24 hours to be quite helpful. It forced us to focus our interview on an acute problem to a degree that we had yet to do as medical students.” On the basis of the program evaluation of the first-year pilot, faculty merged the communicating with team members and work-life balance materials into 1 session under the heading of “Expectations of Students on the Clerkship.”

Commitment to Change

Students made an average of 2 commitment-to-change behaviors in each topic area. Table 3 presents a sample of the commitments made in each area.

PATIENT COMMUNICATION: CARING AND EMPATHY

Results suggest that a brief focused presentation followed by video demonstration and role-play simulation of patient encounters can be successful in changing students’ behavior. The present findings are consistent with the literature7-9 that suggests that physicians in training should develop their ability to listen to their emotions, reflect on patients' emotional clues, express empathy and caring, and de-escalate tense situations. Enabling students to reflect on their own, their fellow students', and the patient's points of view in a controlled situation provides valuable practice in “perspective-taking,” a structured approach to dealing with teammates.

This model mirrors simulation studies described in the aviation literature,10,11 which emphasize preparation, practice, and postsimulation debriefing. Aviation has natural comparisons to the surgical environment, including hypercomplexity, hierarchy, multiple decision makers, a compressed time frame, and a high degree of accountability.

An unpredicted finding was the importance of the collaborative effort involving medical educators and surgical faculty. During the pilot 1–block 4 intervention, the surgical faculty members were called away for emergencies, and so the sessions were taught by primary care physicians and educators. Block 4 was the only intervention block in which students scored lower than the control students. We posit that the lack of surgeons present suggested to the students that the sessions did not need to be taken seriously.
Multiple studies12–15 have proven that a team approach prevents errors in surgical settings. Leonard et al16 attribute communication failures as the leading cause of patient harm. Giddings17 recommends that we establish regular occasions for surgical teams (surgeons, nurses, anesthetists, assistants, and learners) to spend time together and with intensive care unit teams to debrief regarding cases so as to improve patient care. Frankel et al18 identify a key element of effective teamwork as the belief that even the “lowest” member of the team needs to be empowered to speak up when necessary.

In the test vignette, students demonstrated their respect for the nurse’s knowledge when a decision needed to be made regarding postoperative wound care. Comments regarding student commitments to change illustrated that this finding might indicate the beginning of a cultural change that recognizes the importance of a flattened hierarchy.

WORK-LIFE BALANCE

Data show that students are selecting career paths based largely on lifestyle.19 The present findings showed no effect in this area of the human factors curriculum, which may suggest that a more in-depth intervention is necessary to influence deep-seated beliefs that a career in surgery precludes a life outside medicine. Going forward, we will use discussion groups rather than a lecture format, during which students, residents, and faculty will share ideas on how to learn the most in the hospital, find time to study, and balance work and life. Pettitt’s study6 of students’ preconceived fears suggests that students need opportunities to share their concerns about burnout and work-life balance with experienced surgeons who are satisfied with the balance in their own lives.

IMPLICATIONS OF TRAINING FOR SYSTEM-BASED PRACTICE

The Institute of Medicine, in its 2000 seminal report “To Err Is Human,”20 underscored the significant role that human factors play in patient safety, leading to the advancement of quality improvement for systems of care. Research regarding patient safety and quality improvement has found that communication errors are among the most common reasons for poor outcomes in surgery.21–24 The knowledge and skills of human factors are now incorporated into residency training within the Accreditation Council for Graduate Medical Education competency of systems-based practice. The Department of Surgery, believing that teaching these skills is too critical to be relegated to 1 day, is following the lead of Mann,25 who crafted a competency-based curriculum for third-year clerks that presents surgical knowledge within the context of the Accreditation Council for Graduate Medical Education core competencies, including systems-based practice. Successful curriculum redesign using the Mann approach will integrate the human factors curriculum into the clerkship experience.

LIMITATIONS AND NEXT STEPS

This human factors curriculum was performed in 1 medical school with 2 medical school classes. Follow-up studies involving more students from other medical schools are required to generalize the findings.

CONCLUSION

The success of the pilot of an “Introduction to Human Factors” curriculum in the third-year surgical clerkship provides preliminary data that a brief intervention can improve students’ skills in communication with patients and team members. Students improved their skills by speaking with patients about difficult topics and with patients who presented as fearful and anxious. Students also communicated in a more collegial manner with a nurse in a simulated team interaction. A more student-centered educational approach may be required to change the perception of work-life balance in surgical careers. A human factors curriculum in the third-year surgical clerkship represents a culture change, and cultural changes take time. We believe that transparent discussion of communication and lifestyle in surgery must begin early in medical school and that it will lead to more students choosing a career in surgery and to surgeons who practice with better outcomes.

Accepted for Publication: September 15, 2010.
Correspondence: Mitchell A. Cahan, MD, Department of Surgery, University of Massachusetts Medical School, 55 Lake Ave N, Suite S3-725, Worcester, MA 01655 (Mitchell.cahan@umassmemorial.org).

Author Contributions: Study concept and design: Cahan, Larkin, Starr, Wellman, Litwin, and Quirk. Acquisition of data: Cahan, Haley, Sullivan, and Hirsh. Analysis and interpretation of data: Cahan, Starr, Haley, Shah, and Quirk. Drafting of the manuscript: Cahan, Starr, Haley, Sullivan, and Quirk. Critical revision of the manuscript for important intellectual content: Cahan, Larkin, Starr, Wellman, Haley, Shah, Hirsh, Litwin, and Quirk. Statistical analysis: Haley and Quirk. Obtained funding: Quirk. Administrative, technical, and material support: Cahan, Wellman, Haley, Sullivan, Shah, Hirsh, and Litwin. Study supervision: Cahan, Starr, and Quirk.

Financial Disclosure: None reported.

REFERENCES

5. Hassett JM, Zinnerstrom K, Nawotniak RH, Schimphausen F, Dayton MT. Utili-
What Are Human Factors?

Cahan et al define human factors as “the non-technical aspects of surgery,” which is a narrow interpretation of a broad, diverse, and well-established field. The International Ergonomics Association defines human factors as “the scientific discipline concerned with the understanding of interactions among humans and other elements of a system.” It “promotes a holistic approach,” encompassing “physical, cognitive, social, organizational, environmental, and other relevant factors.” Indeed, the surgeon’s technical performance is a human factor—fallible because of our human susceptibility to error—and only one of many. Human factors engineering is a profession devoted to the design (and continual redesign) of systems and processes to protect us from our own inherently imperfect and inconsistent behavior. Aiming to make work environments more effective and efficient, it has been responsible for safety advances in various high-risk domains for more than 50 years—the Human Factors and Ergonomics Society lists 23 technical groups, for fields including aerospace, surface transportation, and health care.

Surgery, which requires the coordination of information, interdisciplinary teams, and equipment across multiple locations and times, is naturally well suited for human factors analysis. Several sophisticated approaches to adapt human factors methods to the surgical domain have been described. By identifying the role that we, as humans working within complex systems, play in influencing surgical outcomes, major safety advances have been made within pediatric cardiac surgery, complex general surgery, and intensive care.

The intervention described in the study by Cahan et al teaches psychosocial skills that, although critical for the surgeon, should not be termed “a human factors curricu-

lum.” We agree that both kinds of skills are neglected in formal medical education and commend the authors’ efforts toward filling an important void in medical education. However, the first step in designing a human factors curriculum in surgery is to be sure we have an accurate and comprehensive understanding of what the term means.

Yue-Yung Hu, MD
Caprice C. Greenberg, MD, MPH

Author Affiliation: Center for Surgery and Public Health, Brigham and Women's Hospital, Boston, Massachusetts.

Correspondence: Dr Greenberg, Center for Surgery and Public Health, Brigham and Women's Hospital, 75 Francis St, Boston, MA 02115 (ccgreenberg@partners.org).

Financial Disclosure: None reported.