Journal of Surgical Research **145,** 327–335 (2008)
doi:10.1016/j.jss.2007.02.020

RESEARCH REVIEW

Navigating Towards Improved Surgical Safety Using Aviation-Based Strategies

Lillian S. Kao, M.D., M.S., F.A.C.S.,*,‡,[1] and Eric J. Thomas, M.D., M.P.H.†,‡

**Department of Surgery, †Department of General Medicine, and ‡Center for Clinical Research and Evidence Based Medicine, University of Texas Health Science Center at Houston, Houston, Texas*

Submitted for publication December 22, 2006

Safety practices in the aviation industry are being increasingly adapted to healthcare in an effort to reduce medical errors and patient harm. However, caution should be applied in embracing these practices because of limited experience in surgical disciplines, lack of rigorous research linking these practices to outcome, and fundamental differences between the two industries. Surgeons should have an in-depth understanding of the principles and data supporting aviation-based safety strategies before routinely adopting them. This paper serves as a review of strategies adapted to improve surgical safety, including the following: implementation of crew resource management in training operative teams; incorporation of simulation in training of technical and nontechnical skills; and analysis of contributory factors to errors using surveys, behavioral marker systems, human factors analysis, and incident reporting. Avenues and challenges for future research are also discussed.

Key Words: **safety management; aviation, health services research; computer simulation; quality assurance; surgery; attitude of health personnel.**

INTRODUCTION

Healthcare has often been compared to the aviation industry as both involve teams of highly trained professionals interacting with advanced technology in high-risk situations. In both professions, errors can impact significantly on outcome in terms of safety and public trust, and they can have financial, legal, and political ramifications. Efforts to improve patient safety have taken numerous cues from the aviation industry, which has a track record of implementing programs to prevent, identify, and manage errors.

Application of aviation-based strategies to surgical disciplines has focused on measuring and improving both technical and nontechnical skills. Although lack of technical competence is a major reason for errors, communication breakdown between personnel and fatigue or excessive workload are also contributory factors [1, 2]. While flight and surgical training are often compared in terms of teaching and measuring an individual's technical proficiency (e.g., simulator training and testing), aviation-based strategies for promoting and evaluating teamwork are also being increasingly incorporated in healthcare. The first half of this paper will focus on the attitudes and behaviors that contribute to a safe working environment, and potential methods for improving teamwork, communication, and other nontechnical skills postulated to reduce adverse events. The second half of this paper will compare surgical training to flight training, including the current and future role of simulators. Finally, limitations of the aviation-healthcare analogy will be discussed, recommendations regarding the adoption of aviation-based strategies to improve surgical safety will be made, and avenues for future research will be addressed.

Safety Culture and Safety Climate

The term safety culture encompasses all of the attitudes, perceptions, behaviors, competencies, and values that contribute to an organization's prioritization of and commitment to safety [3]. The term safety climate refers to the measurable aspect of an organization's safety culture and focuses on perceptions and attitudes, although the two terms are often used as synonyms. In high-risk industries, safety climate describes a group's perceptions of behaviors and procedures as they relate to the overall priority of safety in the workplace [4]. In the aviation industry, there are several tools that have been used to measure safety climate such as the Cockpit Management Attitudes

[1] To whom correspondence and reprint requests should be addressed at Department of Surgery, University of Texas Health Science Center at Houston, 5656 Kelley St., Suite 30S 62008, Houston, TX 77026. E-mail: Lillian.S.Kao@uth.tmc.edu.

0022-4804/08 $34.00

Questionnaire (CMAQ) [5] and the Flight Management Attitudes Questionnaire (FMAQ) [6]. The FMAQ was adapted by Helmreich *et al.* [6] from the CMAQ to be applicable across organizations and cultures. Scores on the CMAQ and FMAQ, used in aviation to assess the attitudes of crewmembers within and between organizations, have been correlated with teamwork quality and flying performance as assessed by trained peer evaluators [7].

The Safety Attitudes Questionnaire (SAQ) was adapted from the FMAQ for use in healthcare. The 23-item questionnaire requires 10 to 15 min for completion and utilizes a five-point Likert scale (ranging from disagree strongly to agree strongly) to evaluate six key factors: teamwork climate, job satisfaction, perceptions of management, safety climate, working conditions, and stress recognition. The SAQ has been formally evaluated with pilot testing and confirmatory factor analyses and has been modified for use in different healthcare arenas including intensive care units [8, 9], labor and delivery units [10], operating rooms [11, 12], general inpatient [13], and ambulatory settings. Based on extensive testing across three countries, 203 clinical areas, and over 10,000 health care providers, the SAQ has been psychometrically validated for use in the medical field [14, 15].

The SAQ has been used as a tool for assessing the effectiveness of quality improvement programs on provider attitudes. For example, Thomas *et al.* [13] performed a randomized controlled trial of executive walk rounds on inpatient units; they demonstrated a higher mean SAQ score among units randomized to the intervention, but did not report an effect on medical errors or patient outcome. Other nonrandomized trials have suggested that a safety intervention program improves both safety climate and clinical outcomes [8, 16, 17]. In surgery, the SAQ was used in a pre- and poststudy of a preoperative briefing at Kaiser Permanente [17]. The preoperative briefings resulted in improved perceptions about the overall safety climate and specifically about teamwork, responsibility for patient safety, and handling of medical errors. The authors attributed the elimination of wrong-site surgery (from 3 to 0 per year) to the preoperative briefings. The number of near misses and reports of faulty or missing equipment increased, suggesting that the safety climate was more amenable to the reporting of problems. Thus, interventions that have resulted in a measurable change in safety attitudes have also demonstrated improved clinical outcomes. However, these studies are limited by methodological flaws associated with before and after studies. Additionally, these studies have not proven causation nor have they elucidated the mechanisms by which improved attitudes translate into increased patient safety.

Safety climate questionnaires may also have other applications in health care such as evaluation of reasons for failure when quality intervention programs are not successful, identification of potential safety improvement goals, and comparison of caregivers, units, or hospitals for benchmarking. Other safety climate and culture surveys exist, but have not been studied in the surgical setting [18], nor have they been studied as extensively [19].

One caveat in using safety climate questionnaires in surgery is that several differences exist between pilots and surgeons. Using cross sectional surveys modified from the CMAQ, Sexton *et al.* [2] compared perceptions about error, stress, and teamwork between healthcare professionals and cockpit crew members. Surgeons tended to downplay the effects of stress and fatigue on decision-making compared with pilots. Additionally, surgeons' perceptions of communication and teamwork were different from those of their team members; surgeons rated their teamwork with anesthesia higher than vice versa. Furthermore, senior surgeons preferred hierarchies whereby junior surgeons tended to accept decisions without question. These perspectives should be considered when adopting aviation-based strategies to surgery. Despite these interdisciplinary differences, safety climate questionnaires may provide insight into perceptions about communication and teamwork among surgical teams and suggest areas for future improvement.

Behavioral Markers Systems

Teamwork is essential to ensuring safety in both aviation and healthcare. Based on confidential interviews with surgeons, most surgical errors are not attributable to an individual but involve multiple personnel and steps; approximately 43% of errors are due to poor communication [1]. Given that positive attitudes about teamwork translate into behaviors that reduce error [7], behavioral markers are important to measure and quantify. Behavioral markers are observed behaviors that have been correlated highly with team performance in context-specific, critical situations in high-risk industries [20]. These behaviors have been measured in aviation using tools such as the Line Operations Safety Audit (LOSA) [21]. Markers associated with superior performance include assertiveness, leadership, and situational awareness [22]. Similar efforts have been made to characterize desired behaviors during medical crises, even using modifications of the LOSA, in the fields of neonatology [23, 24], anesthesia [25], emergency medicine [26], and surgery [27, 28]. However, the behaviors that are advantageous in a cockpit may not be the same as those required during surgery, and validity of these marker systems in the healthcare arena has not completely been demonstrated.

Yule *et al.* [29] reviewed previous research on the nontechnical skills required by surgeons in the operating room and evaluated five behavioral marker systems for rating these skills. The authors identified behaviors relating to communication, leadership, and decision-making as important; however, they found a paucity of data linking these behaviors to outcome in healthcare.

Furthermore, the existing behavioral marker systems were early in development, incomplete, and not psychometrically validated [29]. Future research on the application of behavioral marker systems to healthcare will need to focus on validating these rating tools and on linking specific behavioral markers to errors or adverse events. Ultimately, modifiable behavioral markers should be identified, interventions developed to optimize behavior, and rigorous assessment performed regarding the effects of behavior modification on patient safety.

Crew Resource Management

Crew resource management (CRM) was conceptualized at a 1979 conference focusing on human factors contributing to airplane crashes. Since then, CRM has evolved to become an integral part of flight training; in particular, CRM aims to improve flight safety by focusing on teamwork to prevent and manage crew-based errors. Helmreich *et al.* [30] proposed a schema for classifying the three mechanisms by which CRM could defend against errors. The three tiers of this error *troika* are: avoidance of error, "trapping" of errors before they are committed, and mitigation of the consequences of those errors that are not trapped. The tenets of CRM have been applied to medical fields including anesthesiology [25] and emergency medicine [26, 31, 32]. Additionally, CRM is being taught at the American College of Surgeons meetings, and proponents have advocated mandatory CRM training for surgeons [33].

Several studies in healthcare have demonstrated positive attitudinal and behavioral changes as a result of CRM [16, 26, 31], but few have evaluated its effects on clinical outcome. Morey *et al.* [26] assessed the impact of a CRM-based curriculum in an emergency department: the Emergency Team Coordination Course (ETCC). Using a quasi-experimental design (one pre- and two post-test periods), the authors evaluated team behavior, emergency department performance as measured by observed errors and an admission evaluation survey, and attitudes. They noted improved team attitudes and behaviors as well as decreased clinical errors as a result of the team training.

Although there is good evidence that CRM results in positive attitudinal reactions across high-risk disciplines, a review of the current literature demonstrated that more rigorous studies linking CRM to learning, behavioral modifications, organizational differences, and ultimately safety are needed [32]. However, one of the challenges in validating CRM in the aviation industry has been that a reduction in errors or airplane accidents is difficult to demonstrate due to the rarity of the event. Another challenge is that standard trial designs (e.g., randomized controlled trials) evaluating CRM may be difficult to perform, expensive, and unlikely to be feasible [34]. Furthermore, results from such trials may be difficult to generalize or interpret in that CRM training may be institution or procedure dependent.

In a 2002 editorial in the Journal of the American Medical Association, Leape *et al.* [35] argued that lack of controlled trials evaluating aviation-based safety practices should not preclude adoption of common sense changes with evidence of efficacy in other industries. They suggested that multiple small changes at all levels are responsible for aviation safety, and that failure to demonstrate a reduction in airplane crashes does not diminish the effectiveness of safety improvement efforts. They also recommended that while targeting all potential changes in an organization, whether aviation or healthcare, is impractical, prioritization of practice changes is necessary for effective allocation of time and resources. While CRM as a whole has not been rigorously evaluated in healthcare, there is evidence, as described below, that individual components may be applicable to surgical safety issues.

CRM encompasses team building, briefing strategies, situation awareness (described below), and stress management [30]. Application of several of these CRM-based components has been individually assessed as related to surgery. For example, lack of communication between operating room personnel is common [36], and failure to communicate has been correlated with medical errors [1, 37]. Therefore, operating room communication between team members has been the focus of several quality improvement studies aimed at preventing errors. Team training exercises [16] and preoperative briefing strategies [16, 38], both components of CRM, have been studied in the operative setting. In a pre- and poststudy design, Awad *et al.* [16] documented that operative team training resulted in an increase in communication, as measured using a survey with a 7-point Likert scale, and an increase in compliance with antibiotic prophylaxis and deep venous thrombosis prophylaxis guidelines. Preliminary studies also suggest that preoperative checklists and debriefings are feasible, transferable, and efficient, although further investigations regarding durability and impact on outcome are necessary [17, 38, 39].

Situation awareness (SA) refers to the comprehension of what is going on and what is likely to happen next [40]. Studies of SA in aviation, using the Situation Awareness Global Assessment Technique (SAGAT) have been correlated to performance [22]. Three levels of SA have been identified in aviation: Level 1 or perception of information from the current situation, Level 2 or comprehension of the current situation, and

Level 3 or projection of future events based on understanding of the current situation [22]. In an analysis of the Aviation Safety Reporting System database, almost 97% of errors were due to a lack of Level 1 or Level 2 SA [41]. Lack of SA in surgery can also lead to medical errors; Way *et al.* [42] performed an analysis of 252 laparoscopic bile duct injuries; the majority of the injuries involved misidentification of the common bile duct as the cystic duct and failure to recognize this misperception, resulting in transection or resection of the incorrect duct(s). Further studies on whether measures to improve SA reduce medical errors are necessary.

In aviation, stress is recognized to be a risk factor for error, and stress management is a component of CRM. In comparison, in cross-sectional surveys of pilots and medical personnel, the latter were more likely to deny the effects of stress on performance [2]. In fact, surgical consultants agreed more often than pilots to statements such as "Even when fatigued, I perform effectively during critical times" (70% *versus* 26%) and " My decision-making ability is as good in medical emergencies as in routine situations" (76% *versus* 64%). Despite surgeons' lack of acknowledgment of stressors, they can have a negative effect on operative performance. Moorthy *et al.* demonstrated using motion analysis that operating room stress in the form of a competing task, noise, or need for speed all resulted in decreased dexterity and increased errors [43].

Perhaps as a consequence of denial, there has been relatively little rigorous evaluation of the effects of stress on surgical performance and on effective strategies for coping with stress in surgery. Wetzel *et al.* [44] performed interviews of 16 surgeons and determined that stress in the operating room affects communication, decision-making, and judgment. Coping strategies described in the interviews included using several of the behavioral markers discussed above: situational awareness, communication, and leadership [44].

Despite its widespread adoption, CRM has been recognized to have limitations, even in the aviation industry. The effects of CRM teaching may not be durable, with diminishing returns with repeated training [30]. Additionally, although CRM is generally well accepted, some trainees have rejected it or failed to incorporate the teachings in real-life situations [30]. Even when CRM resulted in attitudinal improvements toward safety, such as in the MedTeams study, healthcare providers rated the potential of CRM to change their behavior to be low [26]. Lastly, the training may not be easily transferable outside of specific organizations, as demonstrated by failure of CRM programs to export from the United States to other countries [30].

Future research in CRM training requires identification of the key components of CRM that are correlated to improved patient outcome and standardization of CRM training and measurement of these skills [32]. Novel trial designs and research methodologies may need to be developed, given the previously mentioned difficulties in the rigorous evaluation of these programs.

Human Factors Research

Human factors analysis is another tool used to determine the causes of errors, originally used for investigating military and commercial aviation accidents. The Human Factors Analysis and Classification System (HFACS) [45] serves as an error framework based on Reason's model of latent (at the system level or "blunt" end) and active (at the frontline or "sharp" end) errors [46]. This taxonomy consists of four levels of failure: unsafe acts (active error) and preconditions for unsafe acts, unsafe supervision, and organizational factors (latent errors). When failures align across multiple levels, an accident trajectory is created, resulting in an adverse event. In surgery, human factors analysis has been used to analyze errors and near misses during neonatal cardiac operations [47, 48]. Human factors analysis has also been used to examine an activity (e.g., an operation) in terms of its components such as technical and nontechnical skill demands [28], mental workload [49], interactions with technology [50] and the work environment [28], and team dynamics [51]. Using this framework for identifying and classifying errors may allow for a better understanding of why near misses and errors occur so as to develop preventive strategies.

Incident Reporting

In the aviation industry, reporting of near misses and adverse events is an important source of data regarding errors. Proponents of an incident-reporting system for surgery have suggested using the Aviation Safety Reporting System (ASRS) and the Aviation Safety Action Program (ASAP) as models. The ASRS is independent of the Federal Aviation Administration (FAA) and allows for voluntary, nonpunitive, anonymous (both the airline and the individual) reporting of incidents. Similarly, ASAP allows for airline-based incident reporting that allows for corrective actions. In surgery, errors but not near misses are reported in morbidity and mortality conference. However, reporting of adverse events is not anonymous and not always without recrimination. Furthermore, in surgical morbidity and mortality conference, errors are not always admitted (24% of cases) or explicitly discussed (40%) [52]. Even complications and deaths may be underreported when compared with other methods of surveillance [53]. Lastly, morbidity and mortality conference does not include a standardized mechanism for feedback to the healthcare system.

Incident reporting systems have been described in anesthesia, critical care, emergency medicine, and cardiac surgery [47]. Nonetheless, challenges and questions exist regarding its implementation in surgery. One problem is that incident reporting is often based on retrospective research resulting in potential recall bias [47]. Additionally, definitions and interpretations of a near miss may vary. Concerns about litigation and a culture of assigning blame in surgery may pose resistance to adoption of such an incident reporting system. Also, while incident reporting systems may potentially identify strategies to prevent or manage errors, they do not allow for quantification of their baseline incidence [54, 55]. Therefore, prioritization of safety issues may make resource allocation more difficult. Another question is whether organization of such an incident reporting system should be at a hospital, regional, or national level is unclear. Finally, no data exists as to whether incident reporting systems result in a reduction in adverse events or near misses. Future research requires standardization of definitions of errors and near misses, development of a methodology for accurately and easily collecting and reporting the data, and evaluation of the effectiveness of reporting in improving outcome.

Training and Competency Assessment

The healthcare field has not only modeled teamwork training and assessment tools after the aviation industry, but also training and competency assessment, namely with simulators. Comparisons have been made between healthcare and aviation in terms of need for selection of trainees, competency assessments, and training with simulators to learn and maintain technical skills.

Given the similarities between aviation and health care, trainee selection for both fields has often been compared. For example, future Air Force pilots must undergo a rigorous selection process that includes psychometric testing in addition to cognitive testing and academic prerequisites [56]. The Air Force Officer Qualifying Test includes sections such as electrical maze (measures path-choosing ability), block counting (measures ability to determine three-dimensional relationships of blocks), and rotated blocks (measures spatial aptitude). Psychometric testing has been demonstrated to have prognostic validity in the aviation industry; that is, a thorough selection process has been demonstrated to reduce the average pilot failure rate [57]. On the other hand, selection into surgical residency, an equally if not more technically challenging field, does not require demonstration of aptitudes such as visual-spatial skills.

Based on a study of a subset of European and North American surgeons, the current system of selecting surgical trainees is perceived as inadequate [58]. They identified cognitive factors, innate dexterity, and personality as the most important attributes for surgeon trainees. However, there is currently no single test for evaluating psychomotor skills, manual dexterity, or personality traits, and tests for these qualities have not been demonstrated to predict surgical performance [59]. Furthermore, visual-spatial evaluations of surgical trainees and experts have not had consistent results. Although studies have demonstrated positive correlations between visual-spatial ability and performance on a complex surgical procedure [60], other studies have found negative correlations between spatial ability scores and surgical ability scores [61]. Failure of visual-spatial test scores to discriminate novice from master surgeons also limits their utility [62]. Additionally, studies have demonstrated that practice and experience may allow trainees to overcome differences in visual-spatial ability, suggesting that psychomotor skills testing may not be the ideal method for identifying potential surgical trainees [60].

Further study is necessary to determine which psychomotor skills, if any, predict performance as a surgeon and the optimal method and time, before or during training, to measure these skills. Moreover, studies should be performed to assess whether selection of trainees based on psychomotor skills improves patient outcome.

The Accreditation Council for Graduate Medical Education (ACGME) requires that residents be provided education to obtain the specific knowledge, skills, and attitudes to attain competency in six areas: patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice. Once training has been completed, surgeons become certified by the American Board of Surgery after meeting standards in seven of nine operative areas during residency in general surgery and passing the qualifying and certifying exams [63]. Certification is then renewed once every ten years. Surgical subspecialization additionally often requires fellowship training and additional board certification. In aviation, pilots must undergo a much more rigorous evaluation [63] that includes a thorough medical examination biannually, yearly simulator checkouts, and unannounced airplane checkouts. To change aircrafts, pilots must undergo further specific training and testing, including training on the simulator. While some have advocated more rigorous and/or more frequent testing for maintenance of credentialing [33, 63], such practices have not been adopted by the American Board of Surgery. Further research is required regarding the optimal methods and timing for the assessment of surgical competency.

Simulation

As stated above, simulation has long been incorporated into training and competency assessments in the aviation industry. However, only recently has simulation in surgery been endorsed by professional associations as a method for potentially improving patient safety. In the surgical field, simulation has uses in the training and assessment of technical skills as well as of teamwork in critical situations, as described in the above section on CRM.

There are a variety of simulators, both physical and virtual, that have been developed to teach and assess basic technical skills. Multiple studies have been performed to assess the validity and reliability of the range of simulators, from physical box trainers such as the MISTELS simulator (McGill Inanimate System for Training and Evaluation of Laparoscopic Skills) [64, 65], to virtual reality simulators such as MIST-VR (Minimally Invasive Surgical Trainer, Virtual Reality) [66, 67]. While several trainers have face, content, and/or construct validity, few simulators have demonstrated predictive validity. There have been at least two randomized controlled trials demonstrating that virtual reality training with MIST-VR results in improved performance during laparoscopic cholecystectomy in the operating room [66, 67]. However, overall, current simulators have not been judged to be sufficient for determining surgical competency. In a review of the literature on simulators and competency assessment, Feldman *et al.* [68] determined that simulators are not ready for use as the sole means of determining surgical competency for a number of reasons: lack of a consensus regarding the definition of competency in terms of operating room performance, lack of standardization of psychomotor tasks for simulators, inability of current simulators to recreate an entire procedure, and a paucity of studies demonstrating predictive validity of simulator training.

Further work is necessary to develop increasingly sophisticated simulators for technical skills training that can recreate the actual operative experience, akin to flight simulators. However, human beings are more complicated and more diverse than airplanes, increasing the complexity and feasibility of such a task. Furthermore, the more difficult the planned procedure is, the greater the desire by healthcare providers to practice on a simulator prior to actually performing the procedure [69]. In addition to the development of high fidelity simulators, standardized methods of competency assessment must be developed and validated so as to be able to rigorously assess the impact of simulator training on operative performance. Lastly, studies are needed that demonstrate that technical skills training on a simulator actually improves patient safety.

In addition to its use for technical skills training, simulation has also been used for evaluation and training of nontechnical skills in crisis situations. Moorthy *et al.* [28] used a simulated operating theater to assess technical skills using a global ratings scale and checklist and nontechnical skills using a modified LOSA tool. Additionally, they simulated a crisis situation of hemorrhage [70]. Similarly, in aviation, Line Oriented Flight Training (LOFT) simulates high stress situations in which crews can practice the behaviors taught in CRM. The role of these simulated situations for training and competency assessment must be further evaluated.

Limitations of the Aviation-Healthcare Analogy

Despite the apparent similarities between surgery and aviation, there are several differences between the two fields that should be considered before universally adopting and instituting aviation-based strategies. First, the healthcare system is more complicated than aviation in terms of the regulatory structure. Unlike the FAA, which oversees airlines and pilots, there is no single regulatory agency in healthcare. While collaboratives such as the National Surgical Quality Improvement Project demonstrate an initiative by surgeons to self-regulate to improve patient safety, participation is far from universal [71]. Surgery and aviation also differ in terms of the interpersonal relationships between professionals [54]. As described previously, surgeons and healthcare teams differ from cockpit crews in their perceptions of the leadership hierarchy, which may affect team dynamics as well as the design and adoption of team building strategies [2, 55]. Furthermore, airline flights have a defined beginning and end, and the crew is centralized in location. In contrast, patient care involves team members at multiple locations becoming involved at different and sometimes overlapping time points, without necessarily a clear beginning or end [55]. Lastly, surgery differs from aviation in that new technologies are developed and incorporated more rapidly in healthcare than in aviation, which results in challenges in determining and assessing competency.

Not only is the healthcare system more complicated, but the human body is more complex and variable than an airplane—people differ from each other in terms of demographics, comorbidities, and disease processes [54]. Additionally, flights are cancelled when airplanes have malfunctioning systems, but emergency procedures are still performed despite organ failure or suboptimal conditions [72]. Ultimately, the true test of the strength of the aviation-healthcare analogy is whether or not aviation-based strategies are proven to be effective when evaluated by methodologically rigorous research.

Recommendations and Challenges for Future Research

The surgical community should develop and validate an evidence-based taxonomy of attitudes, skills, and behavioral markers that define surgical competency and excellence. To be able to rigorously evaluate interventions to modify aspects of the safety culture, these scales should be correlated with medical errors, and if possible, adverse event rates and patient outcome. As in aviation, however, low event rates may make validation difficult, and cultural differences may require that ratings systems be subspecialty or organization specific.

(1) The role of Crew Resource Management and team training in surgical education should be evaluated using methodologically strong study designs. The rationale for team training is sound; most medical care is delivered by teams rather than individuals, and as a result, most errors involve more than one healthcare provider. Individual technical skills are necessary but not sufficient to ensure competency and to prevent errors; cognitive and interpersonal skills are also required. Furthermore, communication breakdowns are one of the most commonly cited root causes for adverse events. Although CRM is integral in the aviation industry, further steps must be undertaken by healthcare researchers to determine how to translate CRM to medicine, the durability of the training, the impact on patient outcome, and cost-effectiveness. Particular challenges of incorporating team management strategies in healthcare lie in the practicality and feasibility of conducting rigorous research on the effectiveness of CRM and its components and in adaptation of these strategies across diverse institutions, subspecialties, and procedures. Simple before and after studies to assess the impact of team training should be avoided because of methodological flaws; alternative study designs such as cluster randomized controlled trials, controlled pre- and postintervention studies, or interrupted time series should be considered [73]. Additionally, novel and innovative research designs may need to be developed to address these issues.

(2) A nonpunitive incident-reporting system should be developed so that surgeons can better identify errors and develop strategies to prevent them. Challenges exist, however, regarding standardization of definitions, minimization of recall and selection bias, and development of feedback mechanisms on both a local and national level.

(3) The role of simulation in surgical education should also be rigorously investigated. Simulation, as in aviation, may be a useful adjunct to other teaching modalities for training and evaluating future surgeons in both technical and nontechnical skills. However, before widespread adoption of simulators can occur, they must be psychometrically validated. Moreover, simulator training should be correlated to improved patient safety.

CONCLUSIONS

Improvement of patient safety and reduction of medical errors require that healthcare providers incorporate strategies and lessons learned from other high-risk industries such as aviation. Strategies that have been used in the aviation industry to improve safety have included attitudinal and behavioral assessments of nontechnical skills such as communication and leadership, teamwork training, incident reporting and nonpunitive management of errors, highly selective requirements for recruitment into and maintenance in the industry, and use of simulation for skills training, However, challenges exist in terms of adapting these strategies to surgery such as demonstration of effectiveness using traditional research designs, adaptation and generalization of these strategies across diverse safety cultures, and prioritization of the safety strategies so as to most effectively allocate the time and resources necessary to fully develop, implement, and evaluate them.

ACKNOWLEDGMENTS

LSK is supported by the Robert Wood Johnson Foundation Physician Faculty Scholars Award. EJT is supported by the Agency for Healthcare Research and Quality (AHRQ) grant 1PO1HS1154401.

REFERENCES

1. Gawande AA, Zinner MJ, Studdert DM, et al. Analysis of errors reported by surgeons at three teaching hospitals. Surgery 2003; 133:614.
2. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. BMJ 2000;320:745.
3. Agency IAE Safety Culture. Vienna: International Atomic Energy Agency, 1991.
4. Zohar D. Safety climate in industrial organizations: Theoretical and applied implications. J Appl Psychol 1980;65:96.
5. Gregorich SE, Helmreich RL, Wilhelm JA. The structure of cockpit management attitudes. J Appl Psychol 1990;75:682.
6. Helmreich R, Merritt A, Sherman P, et al. The Flight Management Attitudes questionnaire. NASA/University of Texas/FAA 1993.
7. Helmreich RL, Foushee HC, Benson R, et al. Cockpit resource management: Exploring the attitude-performance linkage. Aviat Space Environ Med 1986;57:1198.
8. Provonost P, Weast B, Rosenstein B, et al. Implementing and validating a comprehensive unit-based safety program. J Patient Safety 2005;1:33.
9. Thomas EJ, Sexton JB, Helmreich RL. Discrepant attitudes about teamwork among critical care nurses and physicians. Crit Care Med 2003;31:956.
10. Sexton JB, Holzmueller CG, Pronovost PJ, et al. Variation in caregiver perceptions of teamwork climate in labor and delivery units. J Perinatol 2006;26:463.
11. Sexton JB, Makary MA, Tersigni AR, et al. Teamwork in the operating room: Frontline perspectives among hospitals and operating room personnel. Anesthesiology 2006;105:877.
12. Makary MA, Sexton JB, Freischlag JA, et al. Operating room teamwork among physicians and nurses: Teamwork in the eye of the beholder. J Am Coll Surg 2006;202:746.
13. Thomas EJ, Sexton JB, Neilands TB, et al. The effect of executive walk rounds on nurse safety climate attitudes: A randomized trial of clinical units [ISRCTN85147255] [corrected]. BMC Health Serv Res 2005;5:28.
14. Sexton JB, Helmreich RL, Neilands TB, et al. The Safety Attitudes Questionnaire: Psychometric properties, benchmarking data, and emerging research. BMC Health Serv Res 2006;6:44.

15. Modak I, Sexton J, Lux T, et al. Measuring safety related attitudes in the ambulatory setting: The Ambulatory Safety Attitudes Questionnaire. J Gen Intern Med 2007; in press.
16. Awad SS, Fagan SP, Bellows C, et al. Bridging the communication gap in the operating room with medical team training. Am J Surg 2005;190:770.
17. DeFontes J, Surbida S. Preoperative safety briefing project. Permanente J 2004;8:21.
18. Flin R, Burns C, Mearns K, et al. Measuring safety climate in health care. Qual Safety Health Care 2006;15:109.
19. Colla JB, Bracken AC, Kinney LM, et al. Measuring patient safety climate: A review of surveys. Qual Safety Health Care 2005;14:364.
20. Helmreich R, Merritt A. Culture at work in aviation and medicine: National, cultural, and professional influences. Aldershot, UK: Ashgate Publishers, 1998:136.
21. Helmreich R, Wilhelm J, Kello J, et al. Reinforcing and evaluating crew resource management: Evaluator/LOS instructor reference manual. NASA–University of Texas at Austin: 1990.
22. Endsley M. Toward a theory of situation awareness in dynamic systems. Hum Fact 1995;37:32.
23. Thomas EJ, Sexton JB, Helmreich RL. Translating teamwork behaviors from aviation to healthcare: Development of behavioral markers for neonatal resuscitation. Qual Safety Health Care 2004;13 Suppl 1:i57.
24. Thomas EJ, Sexton JB, Lasky RE, et al. Teamwork and quality during neonatal care in the delivery room. J Perinatol 2006;26:163.
25. Howard SK, Gaba DM, Fish KJ, et al. Anesthesia crisis resource management training: Teaching anesthesiologists to handle critical incidents. Aviat Space Environ Med 1992;63:763.
26. Morey JC, Simon R, Jay GD, et al. Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the MedTeams project. Health Serv Res 2002;37:1553.
27. Carthey J, Leval M, Wright D, et al. Behavioral markers of surgical excellence. Safety Sci 2003;41:409.
28. Moorthy K, Munz Y, Adams S, et al. A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre. Ann Surg 2005;242:631.
29. Yule S, Flin R, Paterson-Brown S, et al. Nontechnical skills for surgeons in the operating room: A review of the literature. Surgery 2006;139:140.
30. Helmreich R, Merritt A, Wilhelm J. The evolution of Crew Resource Management training in commercial aviation. Int J Aviat Psychol 1999;9:19.
31. Grogan EL, Stiles RA, France DJ, et al. The impact of aviation-based teamwork training on the attitudes of health-care professionals. J Am Coll Surg 2004;199:843.
32. Salas E, Wilson K, Burke C, et al. Does crew resource management training work? An update, an extension, and some critical needs. Hum Fact 2006;48:392.
33. McGreevy J, Otten T, Poggi M, et al. The challenge of changing roles and improving surgical care now: Crew Resource Management approach. Am Surg 2006;72:1082; Discussion 1126.
34. Musson D, Helmreich R. Team training and resource management in health care: Current issues and future directions. Harvard Health Policy Review 2004;5:25.
35. Leape LL, Berwick DM, Bates DW. What practices will most improve safety? Evidence-based medicine meets patient safety. JAMA 2002;288:501.
36. Lingard L, Espin S, Whyte S, et al. Communication failures in the operating room: An observational classification of recurrent types and effects. Qual Safety Health Care 2004;13:330.
37. Helmreich RL. On error management: Lessons from aviation. BMJ 2000;320:781.
38. Lingard L, Espin S, Rubin B, et al. Getting teams to talk: Development and pilot implementation of a checklist to promote interprofessional communication in the OR. Qual Safety Health Care 2005;14:340.
39. Makary MA, Holzmueller CG, Sexton JB, et al. Operating room debriefings. Jt Comm J Qual Patient Safety 2006;32:357.
40. Singh H, Petersen LA, Thomas EJ. Understanding diagnostic errors in medicine: A lesson from aviation. Qual Safety Health Care 2006;15:159.
41. Jones DG, Endsley MR. Sources of situation awareness errors in aviation. Aviat Space Environ Med 1996;67:507.
42. Way LW, Stewart L, Gantert W, et al. Causes and prevention of laparoscopic bile duct injuries: Analysis of 252 cases from a human factors and cognitive psychology perspective. Ann Surg 2003;237:460.
43. Moorthy K, Munz Y, Dosis A, et al. The effect of stress-inducing conditions on the performance of a laparoscopic task. Surg Endosc 2003;17:1481.
44. Wetzel CM, Kneebone RL, Woloshynowych M, et al. The effects of stress on surgical performance. Am J Surg 2006;191:5.
45. Shappell S, Wiegmann D. Applying reason: The human factors analysis and classification system. Hum Fact Aerospace Safety 2001;1:59.
46. Reason JT. Human Error. National, cultural, and professional influences. New York: Cambridge University Press 1990:173–174.
47. Carthey J, de Leval MR, Reason JT. The human factor in cardiac surgery: Errors and near misses in a high technology medical domain. Ann Thorac Surg 2001;72:300.
48. de Leval MR, Carthey J, Wright DJ, et al. Human factors and cardiac surgery: A multicenter study. J Thorac Cardiovasc Surg 2000;119:661.
49. Carswell CM, Clarke D, Seales WB. Assessing mental workload during laparoscopic surgery. Surg Innov 2005;12:80.
50. Cook RI, Woods DD. Adapting to new technology in the operating room. Hum Fact 1996;38:593.
51. Xiao Y, Hunter WA, Mackenzie CF, et al. Task complexity in emergency medical care and its implications for team coordination. LOTAS Group. Level one trauma anesthesia simulation. Hum Fact 1996;38:636.
52. Pierluissi E, Fischer MA, Campbell AR, et al. Discussion of medical errors in morbidity and mortality conferences. JAMA 2003;290:2838.
53. Hutter MM, Rowell KS, Devaney LA, et al. Identification of surgical complications and deaths: An assessment of the traditional surgical morbidity and mortality conference compared with the American College of Surgeons–National Surgical Quality Improvement Program. J Am Coll Surg 2006;203:618.
54. Thomas E. Aviation safety methods: Quickly adopted but questions remain. AHRQ Web M and M 2006.
55. Thomas E, Helmreich R. Will airline safety models work in medicine? In: What do we know? What do we do? Rosenthal M, Sutcliffe K, Eds., San Francisco, CA: Jossey-Bass, 2002;217.
56. McGreevy JM. The aviation paradigm and surgical education. J Am Coll Surg 2005;201:110.
57. Muller MH. How does aviation find the ideal pilot? Suitability testing: Applicability to surgery? Methods for determining basic occupational suitability. Zentralbl Chir 1999;124:889.

58. Cuschieri A, Francis N, Crosby J, et al. What do master surgeons think of surgical competence and revalidation? Am J Surg 2001;182:110.

59. Bann S, Darzi A. Selection of individuals for training in surgery. Am J Surg 2005;190:98.

60. Wanzel KR, Hamstra SJ, Caminiti MF, et al. Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. Surgery 2003;134:750.

61. Deary IJ, Graham KS, Maran AG. Relationships between surgical ability ratings and spatial abilities and personality. J R Coll Surg Edinb 1992;37:74.

62. Francis NK, Hanna GB, Cresswell AB, et al. The performance of master surgeons on standard aptitude testing. Am J Surg 2001;182:30.

63. Trunkey DD, Botney R. Assessing competency: A tale of two professions. J Am Coll Surg 2001;192:385.

64. Feldman LS, Hagarty SE, Ghitulescu G, et al. Relationship between objective assessment of technical skills and subjective in-training evaluations in surgical residents. J Am Coll Surg 2004;198:105.

65. Vassiliou MC, Ghitulescu GA, Feldman LS, et al. The MISTELS program to measure technical skill in laparoscopic surgery: Evidence for reliability. Surg Endosc 2006;20:744.

66. Grantcharov TP, Kristiansen VB, Bendix J, et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 2004;91:146.

67. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. Ann Surg 2002;236:458; Discussion 463.

68. Feldman LS, Sherman V, Fried GM. Using simulators to assess laparoscopic competence: Ready for widespread use? Surgery 2004;135:28.

69. Greene AK, Zurakowski D, Puder M, et al. Determining the need for simulated training of invasive procedures. Adv Health Sci Educ Theory Pract 2006;11:41.

70. Moorthy K, Munz Y, Forrest D, et al. Surgical crisis management skills training and assessment: A simulation[corrected]-based approach to enhancing operating room performance. Ann Surg 2006;244:139.

71. Khuri SF. The NSQIP: A new frontier in surgery. Surgery 2005;138:837.

72. Rissmiller R. Patients are not airplanes and doctors are not pilots. Crit Care Med 2006;34:2869; Author Reply 2869.

73. Cook T, Campbell D. Quasi-experimentation: Design and analysis issues for field settings. Boston: Houghton Mifflin Co. 1979: 207–214.