This paper describes the communication analyses performed from both live and post-hoc review of 48 surgery cases at an academic medical center. Team performance in the operating room is relatively understudied and non-standardized compared to other high-risk industries. Our research group has worked to evaluate interventions for improving team performance. In this study, we compared the use of a checklist by the senior surgeon to current practices (where no checklist is used), and evaluated teams based on technical errors and communication skills. Communication was measured quantitatively by dividing the procedure into specific events and scoring individual team members using a grading scale from 1 to 5. Using this measure, the intervention group had significantly higher scores in several communication categories. Next, a subset of team member quotes were qualitatively analyzed and divided into four phases: setting the course, gathering data, exchanging information, and tribulations. This qualitative data was used to modify our analysis tools for later studies to take advantage of the patterns we were seeing.

**INTRODUCTION**

The operating room is a high-risk environment where the life of a patient is in the hands of the surgical team. However, performance in the operating room (OR) is relatively non-standardized compared to other high-risk industries. This paper describes our approach to analyzing communication patterns of operating room personnel, which can be challenging to analyze the communications and collaborations of up to 8 people at a time – see Table 1. Due to the challenges of collecting data in the OR, teamwork has primarily been studied in smaller, controlled, simulated environments with anesthesiologists and nurses using an anesthesia simulator, with the role of the surgeon played by whatever surrogate individual is available regardless of training or specialty (Gaba, 1989). This work has yielded important information about the use of crew resource management techniques in a simulated environment. However, currently there are no viable surgery simulators that would allow for realistic whole-team simulation of operative procedures. Thus, our focus is on studying surgery team performance “live” in the operating room.

**TEAM EVALUATION METHODS**

Much of the current team performance literature discusses what makes an effective team – see (McNeese, Salas, and Endsley, 2001) and (Cannon-Bowers and Salas, 1998) for a thorough overview of these issues. Relatively few articles, however, discuss how team measures are obtained.
Another methodology involves having one trained observer watch a whole team, and give overall scores after procedure segments, as performed during aviation crew resource management ratings given to pilot teams (Helmreich, Butler, Taggart, and Wilhelm, 1995). This can be difficult, however, with larger teams. Furthermore, with both of these techniques, team performance is often measured using some form of a note-taking methodology. With this method, there are risks of missing certain aspects of events when taking notes, as well as variations in different observers, without the ability to go back and reanalyze the events. With no “integrated trace” of actual activities, scoring can not be re-evaluated nor can events be easily recalled for debriefing purposes.

We thus developed a hardware/software system entitled Remote Analysis of Team Environments (RATE) to automate as much as possible the ability to digitally record, score, annotate, and analyze team performance for a team of up to 8 people.

Four computers are used to record four separate videos and 8 separate audio tracks (one video feed and two audio feeds are encoded into each computer). In our operating room research, the four videos are comprised of a room camera (with the circulating nurse’s voice), an overhead camera (with the scrub technician’s and camera operator’s voices), a view of the physiological data (with the anesthesiology attending’s and resident’s voices), and the intracorporeal laparoscopic image (with the surgery attending’s and resident’s voices). The purpose of recording each individual’s voice in separate audio track is to make transcription/listening easier (through the use of headsets while watching the team by selectively listening to certain voices during replay).

We have used the tool as a data collection and analysis system for a randomized, controlled trial of the use of a checklist on the improved technical and communication performance of surgery teams performing laparoscopic cholecystectomy (gall bladder removal). Initial results of that project have been reported elsewhere (Calland, 2001). The focus of this paper is on the communication analysis conducted during that study.

METHODS

Subjects

Ten general surgery attendings participated in the study, and were randomly assigned to a Control Group or an Intervention Group. An unavoidable confound was that team members working with an attending vary from case to case, so Control cases were those in which the surgery attending was assigned to the Control Group and Intervention Cases were those in which the surgery attending was assigned to the Intervention Group. A case had a minimum of 6 participants (anesthesia attending and resident; surgery attending and resident; circulating nurse and scrub nurse), but often included a medical student as the laparoscopic camera operator and sometimes other team members in training. Forty eight cases (25 intervention and 23 control) were studied.

Procedure

Surgery attendings in the Intervention Group were briefly trained on the use of a standardized intraoperative paper checklist prior to the study period. The checklist required review of critical steps during laparoscopic cholecystectomy prior to and throughout the procedure. All laparoscopic cholecystectomies were scored using the scoring instruments described below. During the study, surgery attendings in the Intervention Group were instructed to use the checklist and it was taped up in sight of the surgery attending. Surgery attendings in the Control Group performed standard laparoscopic cholecystectomy without any formalized checklist or pre-training. All cases were recorded as described above and scored intraoperatively by two of the investigators on: 1) technical proficiency, using a standardized laparoscopic cholecystectomy scoring instrument (Eubanks et al 1999) and a team communication score system based loosely on Line Operations Safety Audits (LOSA) (Helmreich et al, 1995).

Our initial team communication scale was found deficient once the study was underway (it only rated the whole team on a scale from 1 to 5) so we used the experience to develop a new rating scale post-hoc that rated each team member on a 5-point scale, for each segment of the surgery. The first author re-scored all the cases using the audio-video recordings from the cases.

In addition to these global communications scores, the investigator noted and transcribed “interesting events” (arguments, positive comments or anything perceived as noteworthy) and then analyzed this qualitative data for patterns. A modification of Steeves (1992) outline of Kockelmans’ (1975) “hermeneutic circle” was used to analyze the qualitative data. The interesting event quotes were transcribed and reduced to the smallest amount of data. The units of data with similar meaning were placed into categories. 313 statements from participants (attending and resident surgeons, attending and resident anesthesiologist, medical students, scrub and circulating nurses) and 62 field note entries were obtained and analyzed.
RESULTS

Use of the checklist resulted in significantly higher team communication scores using the post-hoc rating scale, particularly in the preparation phase of the case (see Table 3). No differences in technical scores were noted between the two groups.

Table 2. Communication Results

<table>
<thead>
<tr>
<th></th>
<th>Control (n=23)*</th>
<th>Intervention (n=24)*</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Operational</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introductions</td>
<td>1.11</td>
<td>2.77</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Patient presentation</td>
<td>1.39</td>
<td>2.92</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Role and responsibilities</td>
<td>1.07</td>
<td>2.49</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Contingency planning</td>
<td>1.42</td>
<td>2.82</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Equipment and personnel check</td>
<td>2.25</td>
<td>2.78</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>Intra-Operational</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Team planning</td>
<td>3.03</td>
<td>3.09</td>
<td>0.34</td>
</tr>
<tr>
<td>Communication</td>
<td>3.27</td>
<td>3.54</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>Peripheral environment</td>
<td>3.14</td>
<td>3.14</td>
<td>0.79</td>
</tr>
<tr>
<td>Coordination of flow</td>
<td>3.01</td>
<td>3.28</td>
<td>0.12</td>
</tr>
<tr>
<td>Team expression of concerns</td>
<td>3.04</td>
<td>3.10</td>
<td>0.45</td>
</tr>
<tr>
<td>Situation awareness</td>
<td>3.35</td>
<td>3.36</td>
<td>0.63</td>
</tr>
<tr>
<td><strong>Post-Operational</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-case communication</td>
<td>3.08</td>
<td>3.32</td>
<td>0.09</td>
</tr>
<tr>
<td>Performance review</td>
<td>1.61</td>
<td>2.10</td>
<td>&lt; 0.05</td>
</tr>
</tbody>
</table>

From the qualitative analysis, twenty four categories and nine themes emerged from the data. These were grouped into four phases: 1) Setting the course, 2) Gathering data 3) Exchanging Information, and 4) Tribulations.

Setting The Course

The first phase, setting the course, includes the events prior to the initial skin incision such as establishing rapport with team members and case preparation.

Establishing Rapport. Establishing rapport included the categories of greeting, socializing, and joking. Team rapport was established in a variety of ways by physicians. Some chose a more formal greeting, “Bob, I have not met you,” others, a more lighthearted approach such as, “Everyone knows everyone, hello, hello.” Socializing was used to create a light atmosphere and humor was used to diffuse tedious or tense situations. One surgeon quipped, “coffee break?” to a resident who had just broken a needle while sutureing.

Case Preparation. Preparing for a case included the categories of preplanning and lack of planning. Preplanning occurred in a minority of cases and was best illustrated by the clear statement from the attending surgeon, “I’m going to need additional fluids for this case.” Additional issues that were discussed included: informing staff about a special intraoperative study, requests for particular instruments, and specifying changes in the skin prep solution for the patient’s surgery. Lack of planning resulted in delays for two patients. In one case, staff had not been informed that an intraoperative study was going to be performed and exclaimed, “What cholangiogram now...?” In a second case, lead shields were not available as required to perform an intraoperative diagnostic test.

Gathering Data

Data gathering, the second phase, reflects attempts by team members to increase their understanding of the immediate situation. Team members gathered data through inquiry and illumination.

Inquiry. Inquiries were categorized as questions, seeking advice, requests, and call outs. Questions were asked by team members to fill a knowledge gap, or seek additional information, for example, “Do you know how to work the suction irrigator?” In a few instances, advice was sought from another physician, “Call the GI fellow to see this liver” or from team members, “Does anyone have any ideas?” Requests were seldom directed at individuals but were more likely to be directed to the team in general, “Can we have someone turn on the insufflator for us?” Few callouts (an unsolicited announcement of a completed task) occurred and those that did were related to appliances, e.g., “Foley’s out.”

Illumination. Illumination included the categories of checking, situation awareness, and clarifying. Checking focused on the performance of tasks, “Cut the rope,...really?” while questions regarding the patient’s condition, “How’s he doing? Stable?”, or the status of
the team, "You guys OK?", were recorded in the category of situation awareness.

The majority of clarifying statements related to questions about patients’ appliances or surgical equipment. Questions such as “Foley out?”, “OGT for this one?”, “Seat belt on, right?”, “[cautery] hook plugged in?”, “[will we use a] closure device?”. Clarification of medications, “Antibiotics?”, was a frequent issue as were clarifications about the surgical approach “transverse or supraumbilical [incision]?”, and care management “[wound dressing] wet or dry?”.

Exchanging Information

Exchanging information, the third phase, reflects the variety of ways that the teams conversed and included the themes of discourse, coaching, and closure.

Discourse. Discourse included the categories of commentary and discussion. Commentary that related to the case, but did not directly contribute to the patient’s care was seen in several instances, and ranged from, “These are the gallbladders I live for...thin, easy,”, to the more colorful, “This guys a cirrhotic...@#*$&#, I don’t want to be [aggravated].” Alternatively, commentary reflecting on the patient’s care was noted in a smaller number of cases, “Maybe we should have done a cholangiogram. An ERCP pre-op, should have been done.” Several physicians invited discussion during the case with topics from antibiotic prophylaxis to the type of drain to be used.

Coaching. Coaching included the categories of instruction and performance review. Certain instances of instruction were more effective, “Watch your horizon, don’t move around” than others, “Don’t get in the gall bladder; you have done that on two cases with me.”

Individual performance review was uncommon, occurring in only four cases, “You seem to struggle”... “Are you with me? I am talking to you.” On the other hand, individual praise, “Hang in there buddy, you are doing a good job” and team acknowledgment, “All right troops.” was common.

Closure. Closure included the categories of post-case “thanks” and apology. Thanks were directed at both individuals and the team, “Thank you all, appreciate it”. In one case the surgeon offered an apology regarding their behavior, “I feel bad I yelled at you, I was frustrated.”

Tribulations

Tribulations, the fourth and final phase, included events that were real or perceived obstacles to satisfactory completion of the surgical procedure. Two primary themes, inconveniences and problems, were apparent for this phase.

Inconveniences. Inconveniences included the categories of distractions, repetition, and other cases. The majority of inconveniences was due to environmental distracters, mainly telephone calls and peripheral conversations in the room. A number of instances of repetition were recorded, reflecting uncertainty on a team member’s part as to whether the information had been adequately relayed to other team members. The surgeon lamented, “Is anybody listening? This won’t be effective if you guys don’t pay attention”. Other cases were also discussed during the surgery and took the surgeon’s focus off the task at hand.

Problems. Problems included the categories of arguments, errors, and adverse events. Arguments occurred in five instances, best illustrated by the heated statement, “it’s not my job to do irrigation.” Additional errors (n=77) were identified that were not captured completely by the original technical or communication scoring tools. Device errors occurred in 33 cases, from, “The needle broke on closure, I need pick ups” to “suction not working”. Errors in technique were also present, including small areas of visceral injuries from an errant cautery, to a surgical clip being applied when the back end of the clip could not be fully seen. Lack of knowledge errors were noted in seven cases. These included being unable to operate a camera and take a picture during an ultrasound procedure, to the inability to manipulate the operating room table.

Additional errors included exceedingly long delays while waiting for equipment and in one case after the long wait, the wrong tool was brought. In two instances, surgical field contaminations occurred and the case proceeded without corrective action. In one case, a basket for stone extraction was contaminated on the mask of the attending surgeon (prior to insertion). In another, a contaminated camera bag was moved into the sterile field. During the administration of the post-case questionnaire, five individuals were caught cheating (looking up clinical information they did not know), once by the attending physician, indicating they were not fully aware of relevant patient history information.

Several actions (n=15) resulted in adverse events. One example, “If you push it...oops,” resulting in an instrument puncturing the liver capsule. Surgeons sustained two needle stick injuries, “You stuck my hand” and 1 near needle stick injury.

LIMITATIONS

While this study provided invaluable information that has directed our safety research, it was limited by the lack of transcripts. Transcripts would potentially allow
for more accurate analysis of the data as time constraints of scoring during a procedure would not exist. Transcripts would also provide the opportunity to test inter-rater reliability, ensuring the same information exchange was being scored. Double blinding the scorers would also strengthen the study design by reducing the risk of bias emanating from the scorers’ knowledge of control vs intervention groups.

CONCLUSION

Supplementation of audio-video data and the other objective measures with qualitative data was done to capture infrequent or unique events missed by other measures. This data provided additional insights into issues not anticipated at the study’s inception. In addition, this data analysis has led to further refinement of our observational study methods and tools. We identified a number of important contributors to team interactions in the operating room environment that had not been captured by our other data collection instruments. This additional qualitative data generated the framework for modifying our RATE event tracking software for our current study, which focuses on measuring the effectiveness of team communication training through structured debriefs after cases. We now track conversation “types”, such as asking, questioning, confirming, introducing, so that we can track such communication events on the fly and refer back to these events during debriefs. We also look for repetitions of information as an indication that information was not sufficiently communicated the first time. We also track distractions, device failures, personnel injuries, and delays in procedure, either due to poor planning and equipment failure, or personnel not present. These communication types and failures are presumably relevant to other domains, and thus we hope that the tracking software will easily port to other domains of analysis.

The themes from the qualitative analysis, particularly that clarifying statements related to patients’ appliances or surgical equipment and medications and approach may indicate a lack of planning in some cases, a lack of training in other cases, or just a communication event that could perhaps have been explained more clearly or loudly. This type of analysis of existing whole-team behaviors in a corpus of cases lends insight to the high-priority, low-hanging fruit for further research. These include ideas for making information more visible to the whole team, training of communication skills such as crew resource management training, and ensuring that team members are trained to the extent possible prior to participation in live surgeries. Furthermore, the qualitative analysis lends insight into the cues that team members are using to make decisions during laparoscopic surgery, and thus with further analysis could lead to valuable off-line training program development.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 0092985, the National Patient Safety Foundation, and Karl Storz Endoscopy of America. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES