A PROSPECTIVE ANALYSIS OF SURGERY USING THE OPERATING ROOM BLACK BOX: A PILOT STUDY

James J. Jung, MD (SSTP)[1]; Teodor Grantcharov, MD, PhD[1]
[1]Department of Surgery, University of Toronto and Institute of Health Policy, Management, and Evaluation

Introduction

Errors resulting in adverse events are among the leading causes of death in hospitalized patients[1]. About 10% of hospitalized patients had at least one adverse event[2-9]. Among the adverse events, 60% occurred in surgery and about half of them were preventable[8-12]. Further, 75% of surgical adverse events took place in the Operating Room (OR)[11, 13]. Thus, investigating adverse events in the OR and implementing interventions to mitigate their potential harmful effects are important steps to achieve improvement in patient safety.

In a highly complex system like the OR, small variations in individual and team performances and environmental and organizational factors can accumulate to result in patient harm[14, 15]. Currently, the analysis of intra-operative adverse events (iAE) often relies on post-event reconstruction of self-reported data from Morbidity and Mortality rounds, incident reports, and retrospective chart review. This kind of analysis has serious limitations including reporter recall bias, low compliance rate, and a lack of details[16]. A prospective observation is a more sensitive method to detect small variations in the OR. However, prospective observations required human analysts in the OR, rendering this process laborious, costly, and potentially intrusive. Due to its limited sustainability and spread potential, this method has only been used in small series.

Our group developed a synchronized multiport recording system called the OR Black Box (ORBB). This technology continuously captures and synchronizes data feeds from the OR. Research analysts review the recordings and collect relevant data outside of the OR. Thus, the ORBB allows us to overcome the limitations posed by traditional methods of surgical safety research using self-reported data and direct human observers. The purpose of this paper is to describe surgical team performance, and environmental and organizational factors in the OR for a one-year period from May 2015 to April 2016.

Methods

Overview

We conducted a prospective study of consecutive operations performed by a single surgeon in St. Michael's Hospital (Toronto, ON), a tertiary academic center in urban setting. The study was approved by the research ethics board of St. Michael's Hospital.

The Operating Room Black Box (ORBB)

The ORBB used audio and video feeds during the study period. It captured video feeds from a laparoscopic camera and two small wall-mounted room cameras. It also received audio feeds from several microphones installed behind laparoscope monitors. The cameras and microphones were strategically placed in the least conspicuous locations. The ORBB feeds were then synchronized, encrypted, and stored in a secure server. This technology allowed research analysts to perform a "fly on the wall" assessment of surgery without the intrusiveness and laboriousness of being in the OR. Further, the analysts could pause, zoom in, and replay the ORBB records to carry out accurate and detailed assessment. Two analysts are board certified surgeons and received at least three months of training related to using the ORBB data. Both have more than two years of experience in analyzing the ORBB records as of April 2016. A separate research analyst collected patient-specific characteristics and perioperative variables. They did not have clinical responsibilities during the study period.

The ORBB recording begins when the patient is fully draped and ends right before the drapes come off. The completed ORBB recording is assigned a random five-digit case number

linked to the patient log, which is kept in a secure server for 30 days. Patients and members of the OR team have the right to withdraw their consents within 48 hours, at which point the recording is deleted permanently. All recordings are automatically deleted after 30 days, in compliance with the research ethics protocol.

Study Subjects

All elective laparoscopic general surgery procedures performed on adult patients (≥18 years old) were considered for inclusion. Cases with invalid or missing informed consent from any member of the OR team or patients were excluded. The OR team often comprised of nurses, anesthetists, and an attending surgeon and their respective trainees (clinical fellows, residents, and students). The attending surgeon has fellowship certificate in minimally invasive foregut surgery and nine years of practice. During the study period, performance of anesthetists and their trainees were not measured.

Data Collection

OR Black Box Recordings

The ORBB analysts reviewed the recordings and annotated relevant data points along the timeline. The data points were categorized into procedure steps (Table 1), environmental factors and organizational factors (Table 2), and global assessment of surgeons' technical skills (Table 3).

The steps of procedures were marked along the ORBB timeline. The steps included abdominal access, exposure, dissection, resection, reconstruction, inspection, and closure. We also identified when the primary surgeon role changed from the attending surgeon to a surgical trainee or vice versa. Environmental and organizational factors deemed disruptive to the flow of operations were annotated. They included occurrences of loud noise, door opening and closing, pager or telephone ringing, and nursing changeovers. Objective Structured Assessment of Technical Skills (OSATS) was used to measure seven aspects of surgeons' technical skills[17]. Each category was rated on a 5-point scale to give a total score out of 35[17]. This instrument demonstrated their validity in laparoscopic surgery and on surgeons and trainees.

Statistical Analysis

We performed descriptive statistics including frequency analysis (percentage) for categorical variables and means (standard deviation (SD)) or medians (interquartile ranges (IQR)) for continuous variables. We tested differences in OSATS score among the attending surgeon, clinical fellows, and residents using the one-way analysis of variance (ANOVA). The 95% confidence intervals (CIs) and P values were reported using a two-tailed α level of 0.05. We used SAS 9.4 (SAS Institute Inc., Cary, NC) to perform statistical analyses.

Results

Characteristics of Study Patients and Procedures (Table 1)

Between May 1, 2015 and April 30, 2016, 168 patients underwent elective surgery at the pilot OR. Eight patients did not consent to our study. Of the remaining 160 cases, seven had withdrawn consents from members of the OR team and 24 contained disrupted recordings. Thus, 129 ORBB recordings were included for analysis. Sixty-three (47%) cases were bariatric procedures, either laparoscopic roux-en-y gastric bypass or sleeve gastrectomy. Oncologic gastrectomy cases had the longest procedure duration with a mean (SD) of 147 minutes (58.4 minutes). Paraesophageal hernia repair (62%) and diagnostic laparoscopy (55%) were the only procedures where the attending surgeon participated as the primary surgeon for the majority of case duration (Table 1).

Disruptive Environmental and Organizational Factors (Table 2)

The OR door opened and closed median 42 times [IQR, 32-54] or once every two minutes per procedure. Loud noise was heard median 18 times [IQR, 6-29]. Pager or telephone

rang median 6 times [IQR, 3-8]. Together, these environmental disruptions occurred 66 times, or once every 76 seconds per case. Nursing changeover occurred median once [IQR, 0-2] per case for a mean (SD) of 2.9 (4.4) minutes. Technology in the OR, such as surgical instruments, patient warmers, and laparoscope consoles were absent or malfunctioning in 40 (23%) cases. Finally, “time pressure”, or an inquiry about the estimated time of case completion was communicated in 14 (11%) procedures.

Global Rating of Surgeons’ Technical Skills (Table 3)

The attending surgeon had the highest mean (SD) OSATS score at 32.0 (1.7), followed by clinical fellows at 29.8 (2.0) and residents at 28.0 (3.8). The difference among the three mean OSATS scores was statistically significant ($p< 0.001$).

Discussion

This study presents a prospective audio-video analysis of consecutive elective laparoscopic procedures in the OR for a one-year period. First, we described duration of various procedural steps that make up an operation. A significant amount of time was spent on dissection, resection, or reconstruction. Then, we demonstrated the proportion of time spent as the primary surgeon by the attending, fellows, and residents. On average, the attending spent more time as an assistant than the primary surgeon. Several studies have focused on the increased operating time due to the presence of surgical trainees in the OR[18, 19]. However, no study demonstrated how much time the trainees spent as the primary surgeon.

We described frequency of disruptive environmental and organizational factors. The OR door opened and closed once every two minutes. Door opening is associated with change in airflow, which is hypothesized to increase the risk of wound infection[20]. Disruptions such as loud noises, phone rings, nursing changeovers, and technological failure are linked with increased chance of surgical errors[21]. Avoidable disruptions represent latent failures in the system that may lead to adverse events.

Our study has several strengths. This is the largest sample of prospectively collected consecutive direct observations of the OR. The data points are identified and annotated by dedicated analysts with extensive training in surgery and our study method. The ORBB technology enables the analysts to pause, rewind, and replay the recordings to provide detailed and accurate descriptions, free of recall bias and low compliance.

Our study has several limitations. All procedures were performed by the same attending surgeon at a single center. Further, all cases were elective laparoscopic general surgery. Thus, our findings may not be generalizable to other institutions. Only two ORBB analysts performed data collection. However, these two analysts demonstrated satisfactory level of inter-rater reliability throughout the study period (Jung et al. work in progress). Every 3 months, both analysts rated 10% of the ORBB records to determine the inter-rater reliability. Lastly, there were a significant number of ORBB cases that were disrupted. This might represent a sampling bias. However, the disrupted records contained similar pattern of procedure types as our final sample. Also, they occurred at random temporal distribution.

In conclusion, our study describes procedural steps, environmental and organizational factors, and surgeons’ technical skills at the level of detail that was not feasible without the ORBB technology.

Table 1. Procedural steps and proportion of case as primary surgeon by training level

	Roux-en-Y Bypass	Sleeve Gastrectomy	Oncologic gastrectomy	Paraesophageal Hernia	Internal Hernia	Antrireflux	Cholecystectomy	Diagnostic laparoscopy	Other
Case number, frec	47	16	5	6	4	7	26	11	7
Duration in minutes, mean (SD)									
Total	102 (17)	84.0 (22.6)	146.9 (58.4)	89.5 (37.2)	60.0 (9.1)	70.8 (4.3)	59.8 (21.0)	55.2 (31.4)	72.9 (26.3)
Access	6.5 (1.8)	6.4 (2.1)	7.1 (2.4)	4.6 (1.6)	5.8 (2.0)	5.4 (1.7)	6.3 (2.3)	5.8 (2.2)	5.5 (1.3)
Exposure	3.4 (1.4)	3.9 (1.9)	2.1 (2.0)	3.6 (1.0)	0.6 (1.3)	2.9 (0.9)	1.6 (1.8)	0.8 (1.3)	1.1 (1.4)
Dissection	10.9 (3.8)	28.5 (7.1)	47.7 (28.0)	47.0 (24.0)	0.1 (0.3)	29.8 (13.1)	27.8 (12.8)	16.8 (15.6)	34.9 (23.5)
Resection	0.1 (0.5)	15.1 (4.8)	13.6 (7.0)	0	0	0	7.6 (6.1)	0	1.1 (2.4)
Reconstruction	51.0 (12.4)	0	38.6 (20.6)	15.8 (7.0)	0	21.5 (8.7)	0	0	15.6 (15.3)
Removal of specie	0.7 (3.3)	5.5 [2.2]	6.4 (4.4)	0	0	0	3.8 (3.1)	0.8 (1.8)	1.0 (1.3)
Inspection	3.5 (1.8)	3.5 (3.3)	10.4 (4.6)	4.9 (3.4)	41.3 (7.3)	0.9 (0.8)	3.1 (2.5)	10.9 (4.6)	2.3 (1.6)
Delay	1.5 (1.5)	2.7 (1.6)	5.0 (5.6)	2.1 (2.4)	4.6 (3.1)	0.8 (0.9)	1.4 (1.6)	10.4 (7.6)	1.7 (2.4)
Other	0.3 (1.1)	2.7 (7.0)	2.5 (2.7)	2.8 [4.4]	0	0	0.2 (0.9)	7.4 (17.0)	1.6 (3.0)
Secondary Procec	0.2 [1.2]	2.5 (5.9)	0	2.2 [5.4]	0	0	0	0	0
Closure	2.7 (1.7)	7.1 (3.7)	3.3 (3.9)	1.9 (1.2)	2.3 (2.1)	2.9 (1.8)	2.6 (2.6)	2.6 (2.4)	0.8 (0.9)
Proportion of case as primary surgeon, %									
Attending	35%	25%	44%	62%	40%	30%	5%	55%	42%
Fellow	38%	43%	34%	32%	45%	56%	45%	29%	39%
Resident	27%	32%	22%	7%	16%	15%	50%	16%	19%

Table 2. Disruptive environmental and organizational factors in the OR

	Overall	Roux-en-Y bypass	Sleeve gastrectomy	Oncologic gastrectomy	Paraesophageal Hernia	Internal Hernia	Antireflux	Cholecystectomy	Diagnostic Laparoscopy	Other
Case number, count	129	47	16	5	6	4	7	26	11	7
Case duration in minutes, mean (SD)	83.8 (32.0)	102 (17.0)	84.0 (22.6)	146.9 (58.4)	89.5 (37.2)	60.0 (9.1)	70.8 (4.3)	59.8 (21.0)	55.2 (31.4)	72.9 (26.3)
Door count, median [IQR]	42 [32-54]	45 [35-54]	54 [34-56]	105 [85-105]	40 [35-53]	40 [32-50]	37 [30-46]	34 [26-47]	39 [31-46]	38 [31-66]
Bleep, median [IQR]	69 [43-103]	87 [61-126]	65 [51-84]	95 [89-109]	64 [42-114]	57 [44-82]	51 [38-75]	44 [33-66]	69 [30-116]	73 [47-153]
Externa loud noise, median [IQR]	18 [6-29]	22 [9-32]	10 [6-32]	37 [29-39]	14 [4-25]	14 [5-24]	18 [6-33]	16 [3-25]	14 [4-32]	20 [10-32]
External communication, median [IQR]	6 [3, 8]	6 [4, 9]	6 [3, 8]	9 [6, 9]	5 [5, 9]	2 [1, 6]	8 [4, 12]	4 [2, 7]	7 [5, 9]	4 [2, 7]
Number of nursing changeover per case, median [IQR]	1 [0, 2]	2 [1,3]	1 [0,2]	3 [3, 4]	2 [0,2]	1 [0, 1]	1 [0, 3]	1 [0, 2]	1 [0, 2]	1 [1, 1]
Number of cases with technology malfunction (%)	40 (31.0)	17 (36.2)	6 (37.5)	1 (20.0)	0	0	1 (14.3)	7 (26.9)	5 (45.4)	2 (28.6)
Number of cases with time pressure (%)	15 (11.6)	10 (21.3)	1 (6.2)	1 (20.0)	0	0	0	3 (11.5)	0	0

Table 3. OSATS scores of attending, fellow, and residents

	Attending	Fellow	Resident	p-value
OSATS, mean (SD)	32.0 (1.7)	29.8 (2.0)	28.0 (3.8)	<0.001

References

1. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. *BMJ* 2016; 353:i2139.
2. Leape LL, Brennan TA, Laird N, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med* 1991; 324(6):377-84.
3. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 1991; 324(6):370-6.
4. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000; 38(3):261-71.
5. Wilson RM, Runciman WB, Gibberd RW, et al. The Quality in Australian Health Care Study. *Med J Aust* 1995; 163(9):458-71.
6. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001; 322(7285):517-9.
7. Davis P, Lay-Yee R, Briant R, et al. Adverse events in New Zealand public hospitals I: occurrence and impact. *N Z Med J* 2002; 115(1167):U271.
8. Davis P, Lay-Yee R, Briant R, et al. Adverse events in New Zealand public hospitals II: preventability and clinical context. *N Z Med J* 2003; 116(1183):U624.
9. Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 2004; 170(11):1678-86.
10. Gawande AA, Thomas EJ, Zinner MJ, et al. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery* 1999; 126(1):66-75.
11. Kable AK, Gibberd RW, Spigelman AD. Adverse events in surgical patients in Australia. *Int J Qual Health Care* 2002; 14(4):269-76.
12. Zegers M, de Bruijne MC, Wagner C, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. *Qual Saf Health Care* 2009; 18(4):297-302.
13. Zegers M, de Bruijne MC, de Keizer B, et al. The incidence, root-causes, and outcomes of adverse events in surgical units: implication for potential prevention strategies. *Patient Saf Surg* 2011; 5:13.
14. Barach P, Johnson JK, Ahmad A, et al. A prospective observational study of human factors, adverse events, and patient outcomes in surgery for pediatric cardiac disease. *J Thorac Cardiovasc Surg* 2008; 136(6):1422-8.
15. Catchpole KR, Giddings AE, Wilkinson M, et al. Improving patient safety by identifying latent failures in successful operations. *Surgery* 2007; 142(1):102-10.
16. 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(5):618-24.
17. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997; 84(2):273-8.
18. Babineau TJ, Becker J, Gibbons G, et al. The "cost" of operative training for surgical residents. *Arch Surg* 2004; 139(4):366-9; discussion 369-70.
19. Allen RW, Pruitt M, Taaffe KM. Effect of Resident Involvement on Operative Time and Operating Room Staffing Costs. *J Surg Educ* 2016; 73(6):979-985.
20. Pokrywka M, Byers K. Traffic in the operating room: a review of factors influencing air flow and surgical wound contamination. *Infect Disord Drug Targets* 2013; 13(3):156-61.
21. Wiegmann DA, ElBardissi AW, Dearani JA, et al. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. *Surgery* 2007; 142(5):658-65.