Operating Room Managerial Decision-Making on the Day of Surgery With and Without Computer Recommendations and Status Displays

Franklin Dexter, MD, PhD*†

Ann Willemsen-Dunlap, CRNA, PhD*

John D. Lee, PhD‡

BACKGROUND: There are three basic types of decision aids to facilitate operating room (OR) management decision-making on the day of surgery. Decision makers can rely on passive status displays (e.g., big screens or whiteboards), active status displays (e.g., text pager notification), and/or command displays (e.g., text recommendations about what to do).

METHODS: Anesthesiologists, OR nurses, and housekeepers were given nine simulated scenarios (vignettes) involving multiple ORs to study their decision-making. Participants were randomized to one of four groups, all with an updated paper OR schedule: with/without command display and with/without passive status display.

RESULTS: Participants making decisions without command displays performed no better than random chance in terms of increasing the predictability of work hours, reducing over-utilized OR time, and increasing OR efficiency. Status displays had no effect on these end-points, whereas command displays improved the quality of decisions. In the scenarios for which the command displays provided recommendations that adversely affected safety, participants appropriately ignored advice.

CONCLUSIONS: Anesthesia providers and nursing staff made decisions that increased clinical work per unit time in each OR, even when doing so resulted in an increase in over-utilized OR time, higher staffing costs, unpredictable work hours, and/or mandatory overtime. Organizational culture and socialization during clinical training may be a cause. Command displays showed promise in mitigating this tendency. Additional investigations are in our companion paper.

(Anesth Analg 2007;105:419–29)

Anesthesia providers and operating room (OR) nurses make managerial decisions throughout each day of surgery. For example, they decide when to call for the next patient coming to their OR. They decide whether to start a case in one OR or another.

The statistical basis for operational decision-making on the day of surgery is understood (1,2). Economically rational decisions arise from the use of the following ordered priorities (1): i) performing all scheduled cases unless there is a patient safety concern, ii) reducing over-utilized OR time, iii) reducing patient and surgeon waiting times, and iv) satisfying personal priorities, etc. The decision-making is summarized in Table 1, the Methods, and two review articles (1,2). Profitable facilities that are focused on satisfaction of surgeons can have sufficient excess capacity to prevent almost all over-utilized OR time (3). Decisions on the day of surgery are then based on reducing surgeon waiting.

Commercial products to communicate OR management information on the day of surgery include both passive status displays (e.g., big screens or whiteboards) and active status displays (e.g., text pagers). These displays show data about where the patient is (their status) as they advance through the perioperative period. It is unknown how best to use these passive and active displays to present clinicians with information derived from statistical methods (1,2) to achieve optimal decision-making. Previous papers that have reported the use of status displays in ORs for managerial decision-making have been limited to decisions involving individual ORs (Table 2).

OR staff likely use displays to increase the speed of work in each OR (i.e., increase the work per unit time in

From the Departments of *Anesthesia, †Health Management and Policy, and ‡Mechanical and Industrial Engineering, University of Iowa, Iowa City, Iowa.

Accepted for publication April 16, 2007.

Dr. Dexter, Section Editor for Economics, Education, and Policy, was recused from all editorial decisions related to this manuscript.

This research was performed in part during a consultation for Stottler Henke Associates, Inc., by the Division of Management Consulting of the Department of Anesthesia of the University of Iowa. FD is the Director of that Division. He receives no funds personally other than his salary from the State of Iowa, including no travel expenses or honoraria, and has tenure with no incentive program.

An abstract describing this work was presented at the American Society of Anesthesiologists' Annual meeting on October 17, 2006, in Chicago, Illinois.

Address correspondence and reprint requests to Franklin Dexter, PhD, Division of Management Consulting, Department of Anesthesia, University of Iowa, Iowa City, IA 52242. Address e-mail to Franklin-Dexter@UIowa.edu.

DOI: 10.1213/01.ane.0000268539.85847.c9

Table 1. Ordered Priorities for Operational Decision-Making on the Day of Surgery (1)

1. Patient *safety* and quality of care are preeminent
   - Every clinical deadline should be satisfied, regardless of the resulting workload
   - For urgent cases, deadlines refer to times by which each case must start to prevent patient harm
   - Cases are only moved/added if able to ensure clinical safety and quality
2. *Access*—Every surgeon has open access for cases on any future workday, provided the cases can be done safely
   - This principle promotes flexibility and growth of the surgeons' practices
   - The limit to case booking is safety, not an arbitrary time of day (e.g., 3:30 PM, 3:45 PM, or 6:00 PM)
3. Operating room *efficiency*
   - If staffing is planned from 8:00 AM to 3:45 PM, and the last case of the day in the OR ends at 2:15 PM, there are 1.5 h of under-utilized OR time
   - If staffing is planned from 8:00 AM to 6:00 PM, and the last case of the day in the OR ends at 7:00 PM, there is 1 h of over-utilized OR time
   - On the day of surgery, OR efficiency is maximized by minimizing the hours of over-utilized OR time
4. Patient service—Reducing *patient waiting time* on the day of surgery
   - If case is scheduled to enter its OR at 10:00 AM, and instead enters at 09:40 AM, there is 20 min of earliness
   - If case is scheduled to enter OR at 2:00 PM, and instead enters at 2:40 PM, there is 40 min of tardiness
   - For elective cases, reducing waiting times means reducing total tardiness of all patients scheduled in the OR on a given day
   - For urgent cases, reducing patient waiting times means reducing waiting in hours from when the patient and surgeon are available
5. Other priorities such as personal priorities, professional satisfaction, and education

When Table 1 was posted at the study hospital's OR control desk as used in Study #2 of our companion paper, there were two other lines below priority 3, operating room efficiency. One line was that staffing is planned for three (3) ORs from 6:00 PM to 7:00 AM all days of the week. The other line was that staffing is planned for three (3) ORs from 7:00 AM to 7:00 AM on Saturdays and Sundays.

each OR). Surgeons can perceive that the most important attributes of a high-performing anesthesia group include those related to working quickly in individual ORs: timely starts and brief times for patient awakening, turnover, and patient entrance to positioning (11). Time is a consistent catalyst for tension and interpersonal conflict among physicians and OR nurses (12,13). When time is the subject of communication during cases, 46% of interactions can involve high tension and blame (14). OR staff can perceive that the desire to perform as many surgical cases in a day as possible is so pervasive as to be the most dominant organizational factor underlying unsafe practices at work (15). Staff working in outpatient surgery centers can perceive time pressures to be so strong as to commonly affect quality of care and patient safety (16). Some anesthesia departments publish each anesthesiologist's time of entrance of the first patient of the day into each OR, anesthesia induction time, wakeup time, and turnover time (17). Financial incentives have been provided for anesthesiologists meeting measures of "productivity . . . on the day of surgery," defined as timely entrance of the first patient of the day into his or her OR, brief time to positioning, and lack of waiting for the anesthesiologist (18).

Nevertheless, increasing each clinician's clinical work per unit time is *not* the same as following the ordered priorities, described above, when decisions involve more than one OR. As summarized below, decisions that increase clinical work per unit time can result in decisions that are suboptimal economically for the surgical suite as a whole (1,2). We used an experimental simulation study to assess the degree to which passive status displays (i.e., information upon request) or active command displays (i.e., recommendations) can result in decision-making that matches the ordered-priorities of Table 1.

BACKGROUND

This section is a summary of definitions and results needed to develop and interpret our study. Terms defined are shown in italic. References, details, justifications, explanations of economic rationality, and dozens of examples are in two recent review articles of OR management operational decision-making based on OR efficiency (1,2). A summary is in Table 1.

Allocated OR time is an interval of OR time, with a specified start and end time on a specified day of the week, assigned by the facility to a service for scheduling cases (e.g., otolaryngology in OR #23 from 7:15 AM to 3:30 PM).

A few months before the day of surgery, based on the OR allocations, the process of *staff scheduling* determines the individuals who will work each shift on each day. The allocated OR time affects not just appropriate staff scheduling in anesthesia departments, but also in OR nursing, postanesthesia care unit, surgical pathology, surgical wards, and so forth.

If the allocated time for an OR was planned from 7:15 AM to 3:30 PM, and the last case of the day in the OR exited the OR at 1:30 PM, there were 2 h of *under-utilized OR time*. If the last case of the day exited at 4:30 PM, there was 1 h of *over-utilized OR time*. Depending on staff scheduling and compensation, each hour of over-utilized OR time may be an hour of overtime.

On the day of surgery, the cost of an hour of under-utilized OR time is negligible relative to the cost of an hour of over-utilized OR time (2,19). Therefore, decisions can be made to complete the existing cases with as little over-utilized OR time as possible, and without regard for the under-utilized OR time (1).

In the ordered priorities above (Table 1), increasing OR efficiency is a lower priority than completing the cases each day. The reason is that, overall, virtually every surgeon's cases contribute positively to hospital and professional margin when there is prospective payment (e.g., diagnosis related groups) (3).

Table 2. Previous Studies of Use of Status and Command Displays in Operating Rooms (ORs) for Managerial Decision-Making

Passive status displays displaying OR occupancy (4)
Passive status displays provided knowledge of OR occupancy in real-time from vital sign data. Qualitative end-points included a manager reporting that he used the patient-in times & patient-out times to prioritize the tasks of getting the next cases ready. Another OR manager reported that she used the patient-in times & patient out-times to follow up on the rooms that were unoccupied for more than half an hour
Passive status displays of OR activity located throughout surgical suite and wards (5)
Passive airport style status displays were implemented throughout a hospital's ORs and wards. Qualitative assessments were that displays gave nurses on the surgical wards access to knowledge of ongoing OR activity so that they could anticipate when to prepare their patients. There were 7 min mean reductions in times for patients to arrive from hospital wards and 4.5 min mean reductions in turnover time
Passive status displays on computer screens showed OR case status (6)
Passive status displays on computer screens were used to show OR case status. Qualitative assessments were that post-anesthesia care unit (PACU) assignment decisions were improved
Active status displays (text pagers) assigned to surgical wards (7)
Active status displays, specifically text pagers, were assigned one to each surgical ward to be carried by the charge nurse during his or her shift. Each patient's post-anesthesia care unit nurse would send a message to the pager approximately 20 min before providing a phone report for patient transfer. Qualitative assessments were that staff assignment decisions and staff communication were improved
Active status display (text pagers) assigned to surgical ward charge nurses (8)
Patient admission into the PACU was monitored automatically. An automated paging system alerts clinical unit managers to "pull" their patients from the PACU after a set recovery period. The mean PACU length of stay decreased in the orthopedic intervention group by 26 min, while the mean length of stay increased nonsignificantly in the general surgery control group
Active status displays in hallways and ORs alerting need for assistance in ORs (9)
Keypads in ORs displayed auditory and visual alerts throughout relative portions of surgical suites (e.g., for medically directing anesthesiologist to come to an OR). Mean response time of anesthesiologists to ORs after being paged was reduced by 1 min versus when digital pagers were used. Qualitative assessment was that digital pagers were inferior communication tools, because only one person was paged instead of an individual and available colleagues, incorrect beeper numbers were sometimes reported to personnel, multiple pages were sometimes made because the caller was not certain that the correct number had been dialed, missed or ignored pages required repeat paging, and there was lack of public accountability for responsiveness
Active status displays (text pagers) provided to patients' friends and family (10)
Providing friends or family members with a pager so that they could leave the surgical waiting room if they wanted to do so, and yet be notified when the patient was finished with surgery and the surgeon was available to speak with them, did not reduce their underlying state anxiety (confidence interval for mean reduction in anxiety −16.5 to 1.5; control group anxiety 38.6 ± 12.0, $N = 28$; pager group anxiety 46.1 ± 13.3, $N = 29$)

Even without any incremental revenue, canceling cases on the day of surgery increases costs, whether analyzed from the perspective of the hospital, patient, physicians, or society (20). Therefore, completing all scheduled cases each day is economically rational. The consequence is that minimizing over-utilized OR time serves to maximize the ratio of the output (clinical care) to input (staffing costs), which is efficiency. Thus, on the day of surgery, minimizing over-utilized OR time maximizes OR efficiency (1,19,21).

When decision-making involves multiple ORs, decisions to reduce over-utilized OR time can conflict with decisions to increase clinical work per unit time in each OR.

Example 1: A six OR ambulatory surgery center had five ORs allocated (staffed) for short hours (7:15 AM to 3:30 PM) and one OR allocated for long hours (7:15 AM to 6:00 PM). One day, the last case of the day in the 6:00 PM OR was to be performed by a different surgeon than the preceding cases. Based on progress in that OR, that 1.5 h case was expected to start at 4:00 PM, and thus result in no expected over-utilized OR time. Nonetheless, at 2:30 PM the case was moved to the same anesthesiologist's 3:30 PM OR. Starting the case earlier increased the anesthesiologist's clinical work per unit time. The anesthesiologist may be rewarded intangibly or tangibly for doing more of the work. Nevertheless, the consequence was over-utilized OR time. The OR nurses were scheduled so that there was no staff available to relieve the OR team in the 3:30 PM OR other than the staff in the 6:00 PM OR once they finished their cases. Thus, the over-utilized OR time meant, in practice, not only unpredictable work hours and increased staffing cost, but also mandatory, unexpected, and unnecessary overtime to finish the case, albeit briefly.

When a decision does not affect over-utilized OR time, the next ordered priority is to reduce patient and surgeon waiting time (Table 1) (1). For elective cases, this refers to reducing waiting from the scheduled start time (Table 1) (1) (i.e., tardiness as defined in the following Example).

Example 2: Three cases of the same procedure were performed in an OR. Each case took 2 h. The turnover times were 30 min. Scheduled start times were 8 AM, 10:30 AM, and 1 PM. The first case entered the OR at

Table 3. Scenarios Listed in Sequence Viewed by All Participants

Scenario #1
Basis for the decision was over-utilized OR time
Vignette asked staff to decide whether first to focus their attention on one OR versus another
Anesthesiologist had to decide whether first to assist the CRNA in one OR or the other OR
Scenario #2
Basis for the decision was reduction in expected tardiness from scheduled start times, with no effect on over-utilized OR time
Vignette asked staff to decide whether first to focus their attention on one OR versus another
Housekeeper had just started cleaning one OR when receiving a request to clean a different OR
Scenario #3
Basis for the decision was over-utilized OR time
Vignette asked staff whether to move a case from one OR to another
Decision to move a case from one OR into another differs from other management decisions on the day of surgery in that the decision does not have to be made. There is no medical necessity to move an elective case. This is why there is heterogeneity (22) of views among OR managers in the amount of over-utilized OR time to be saved to move a case. To avoid this dilemma, Scenario #3 addressed the moving of a case from one OR with OR time allocated from 7:15 AM to 6:00 PM to another OR with OR time allocated from 7:15 AM to 3:30 PM, resulting in mandatory overtime (see *Example 1*)
Scenario #4
Basis for the decision was reduction in expected tardiness from scheduled start times, with no effect on over-utilized OR time
Vignette asked staff to decide whether first to focus their attention on one OR versus another
Scenario #5
Basis for the decision was over-utilized OR time
Vignette asked staff to decide whether first to focus their attention on one OR versus another Information provided by the displays was out of date, and should be ignored
Scenario #5 was used to test what happens when there are poor recommendations from displays, but safety is unaffected (see *Displays* section). The information on which displays were based neglected that the attending neurosurgeon and chief resident in an OR had been called unexpectedly to the emergency department to evaluate a patient with head trauma during the OR turnover
Scenario #6
Basis for the decision was over-utilized OR time
Vignette asked staff to assign add-on cases to ORs during the clinical day
Scenario #7
Basis for the decision was over-utilized OR time
Vignette asked staff to decide whether first to focus their attention on one OR versus another
Scenario #8
Basis for the decision was safety
Vignette asked staff to decide whether first to focus their attention on one OR versus another
Information provided by the displays incorrectly suggested that the basis for the decision was tardiness. Information on which displays were based failed to consider that the housekeeper was needed in one OR in which the preceding patient had extensive bleeding instead of another OR that finished earlier with a breast biopsy. Scenarios #8 and #9 were put last so that any resulting reduction in trust (23) in information from displays would not bias responses to other scenarios
Scenario #9
Basis for the decision was safety
Vignette asked staff to decide whether to assign a nurse to one OR or another
Information provided by the displays incorrectly suggested that the basis for the decision was over-utilized OR time. Information on which displays were based failed to consider that an OR nurse lacked the skills appropriate to circulate at night for a jugular foramen tumor resection

8:20 AM. For purposes of a decision made at 9 AM, the *tardiness* of the start of the OR's remaining cases from their scheduled start times could be considered to be 40 min, where 40 min = 20 min late for the second case plus 20 min late for the third case. The minimum value of tardiness was 0 min.

In practice, most decisions should be made based on reducing tardiness (i.e., surgeon and patient waiting). The reason is that when OR allocations are made appropriately or additional capacity is provided to ease case scheduling (2,3), most ORs have no over-utilized OR (1,2). Yet, often an OR without any expected over-utilized OR time has cases with expected tardiness of start (e.g., an OR with four cases totaling 6.0 h starts its first case of the day 15 min late).

Although personal priorities may intuitively seem to often be a higher priority than reducing waiting, they cannot because bizarre behaviors would then be appropriate. For example, consider Example 2 with three cases totaling 6 h. An OR nurse could then reasonably take a 1.5 h lunch break in the middle of this day, as it would not cause over-utilized OR time, just further increase tardiness from scheduled start times.

Table 4. Portion of the Paper Operating Room Schedule Relevant to Simulated Scenario #1

In	Out	Room	Time	Case	HCase	Age	Operation	Surgeon
7:18	8:25	ASC1	7:15	1′15	1′25	35	Hysteroscopy, surgical; with sampling (biopsy) of endometriu	Gynecology 1
Four other cases in ASC1 deleted here for presentation, because Scenario #1 involved only ASC2 and ASC4								
7:32	10:35	ASC2	7:15	3′15	2′57	1	Repair of hypospadias complications (i.e., fistula, stricture, divert	Urology 1
10:43		ASC2	10:45	2′00	1′41	10	Orchiopexy, inguinal approach, with or without hernia repair	Urology 1
		ASC2	13:00	2′30	3′41	0	Pyeloplasty (Foley Y-pyeloplasty), plastic operation on renal pel	Urology 1
Seven rows of cases in ASC3 deleted here for presentation, because Scenario #1 involved only ASC2 and ASC4								
7:15	7:33	ASC4	7:15	0′30	0′32	2	Tympanostomy (requiring insertion of ventilating tube), general	Otolaryngology 3
7:43	8:05	ASC4	8:00	0′30	0′32	3	Tympanostomy (requiring insertion of ventilating tube), general	Otolaryngology 3
8:15	8:38	ASC4	8:45	0′15	0′23	4	Unlisted procedure, palate, uvula	Otolaryngology 3
8:56	10:34	ASC4	9:15	1′00	1′09	20	Tonsillectomy, primary or secondary; age 12 or over	Otolaryngology 3
10:45		ASC4	10:30	1′15	1′19	2	Laryngoscopy direct, with or without tracheoscopy; diagnostic,	Otolaryngology 3
		ASC4	12:00	1′00	1′03	5	Tonsillectomy and adenoidectomy; under age 12	Otolaryngology 3
		ASC4	13:15	1′15	1′03	6	Tonsillectomy and adenoidectomy; under age 12	Otolaryngology 3
Seventy-six rows of cases deleted below here for presentation, because they were for ORs other than ASC2 or ASC4								

Columns from left to right are times patients actually entered (In) operating rooms (ORs), time patients actually exited (Out), case's scheduled OR (Room), case's scheduled time of patient entry into the OR (Time), time that the patient is scheduled to be in the OR plus 15 min for turnover time (Case), historical average time for cases of the same surgeon and scheduled procedure(s) by Current Procedural Terminology code(s) plus 15 min for turnover time (HCase), patient age in years (Age), scheduled procedure(s) printed on one line often with truncation (Operation), and Surgeon. Explanations for differences between the simulated paper schedule and actual schedules are in the *Methods* section.

The displays that we studied not only showed tardiness, but also earliness. Suppose that a 2 h case enters an OR at 11:30 AM. The turnover times are 15 min. The scheduled start time of the next and final case is 2:30 PM. The *earliness* of start of the OR's case is 45 min, where 45 min = 2:30 PM to 1:45 PM. The minimum value of earliness is 0 min.

METHODS

Scenarios

Nine scenarios were created from actual cases (Table 3). The scenarios involved anesthesiologists, OR charge nurses, and/or turnover personnel. Each scenario described a decision that involved two ORs and had a correct answer based on expected over-utilized OR time, under-utilized OR time, earliness, and/or tardiness (Table 1).

The nine scenarios were created so that one updated paper OR schedule for 1 day [i.e., status display with raw data (24)] was sufficient for all scenarios (Table 4).

Room assignments, scheduled start times, scheduled OR times, historical average OR times, ages, times of patient entrance and exit from ORs, and procedures were realistic and internally consistent. The paper OR schedule matched the ones printed at the participants' hospital other than by not listing the patient name or medical record number. In addition, under the column with the surgeon's name, the specialty was listed with sequential numbers for each surgeon (e.g., Gynecology 1 and Gynecology 2), instead of the surgeon's name.

Which ORs were allocated to 3:30 PM or 6:00 PM were evident from the paper schedule based on the time at which the last case of the day was scheduled to end in the OR. For example, Table 4 shows scenario #1 in which both ASC2 and ASC4 were allocated to 3:30 PM.

Displays

Paper-based graphical depictions were created of passive status displays (24) showing hours of over/under-utilized OR time and minutes of tardiness/earliness (Fig. 1). When the last case of the day in an OR was ongoing or had finished, the words "Last case" were displayed instead of the tardiness or earliness. The displays offered data about each of the ORs in a scenario. The displays were depicted by pictures of LED displays positioned over OR doors like those at airline gates (Fig. 1). These passive

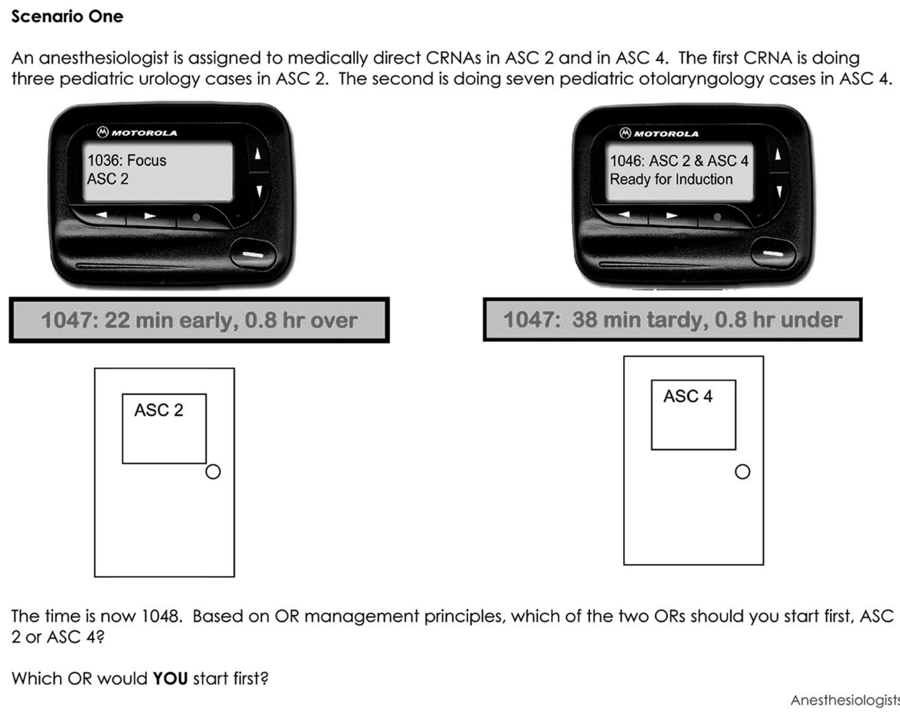

Figure 1. Active status displays and passive status displays. Each scenario was printed in color in landscape view on one piece of paper. The command display and/or status display, when applicable, was printed on that one piece of paper. There were 108 different versions, where 108 = (9 scenarios) × (4 groups) × (3 types of providers). The nine scenarios are given in Table 3. The four groups are given in Table 7: door status display, pager active status display, neither of the two displays, or both displays. The three types of providers were anesthesiologists, operating room (OR) charge nurses, and turnover personnel. The particular version shown is for Scenario #1, "Both" pager and door, and provider of anesthesiologist. The Scenario #1 was chosen, because it matches Table 4. Readers can compare the times shown on the pagers and doors to those of Table 4. In the figure, the abbreviation "CRNA" stands for Certified Registered Nurse Anesthetist. The "ASC" refers to the study hospital's Ambulatory Surgery Center. The correct decision was to focus on ASC2 in order to reduce expected over-utilized OR time.

displays were chosen for the experiment since they provided all data required for the decision, no redundant information (e.g., other ORs), and focused participants on the fact that they were being asked to address the deliberately over-simplified problem of just two ORs. In the "real world," the spatial distance between such passive displays would reduce their effectiveness. Thus, the study was, by design, biased to show a benefit of status displays.

Participants could try to use the data from the scenarios and the status displays to estimate uncertainties in OR times (25) and incorporate the uncertainty in their decision-making. Therefore, we designed the status displays to provide not just the expected over-utilized OR time and tardiness, but also the expected under-utilized OR time and earliness (1).

Graphical depictions of active command displays were pictures of the hospital's alphanumeric text pagers with recommendations (Fig. 1). The text recommendations pertained either to which OR to focus on (e.g., start first or clean first), to which OR to assign an add-on case, to which OR to assign staff, or whether or not to move a case to a different OR.

Although the use of command displays may serve to mitigate the influence of the heuristic of working fast on decision-making, the displays may also inappropriately lull participants into relying on recommendations in lieu of clinical judgment. Previous studies in other subject areas found participants complied with poor recommendations from command displays (Table 5). In three of the scenarios (#5, #8, and #9), incorrect recommendations were made by the active command display and would be suggested by the passive status display or use of the paper OR schedule.

Participants

Following institutional review board approval, eight certified registered nurse anesthetists (CRNAs) from the facility where the research was completed piloted the study in November 2005. Minor modifications were made to scenario presentation, depictions of displays, and to presentation of the paper OR schedule. Starting December 2005, enrollment began. Enrollment continued through the end of January.

The subject pool for the study included anesthesiologists, OR charge nurses, and turnover personnel whose responsibilities included cleaning and assisting in the preparation of ORs and equipment for each case. Examples of decisions by nurses and housekeepers are in Table 3. Participants included 12 anesthesiologists, 15 registered nurses, and 18 housekeepers and anesthesia assistants. Employment at the hospital had been 5 yr or less for 44% and greater than 10 yr for 40%.

Table 5. Decision-Making with Unreliable Status and Command Displays from Other Subject Areas (Domains)

Unreliable command and status displays (26)
Pilots flew simulated approaches involving icing encounters. The baseline group received raw data, the status display group received information about the icing situation but had to determine which course of action to take, and the command display group received a computer generated recommendation. When the status and command displays were accurate, decision-making was better than baseline. When the displays were inaccurate, performance was worse, particularly for the group with a command display
Unreliable command displays (27)
Air traffic controllers managed aircraft in a free flight environment. Each simulation contained events of aircraft flying too close to one another. The group with a perfectly reliable command display performed significantly better than the group with manual operation based on data from a passive status display. The participants with command displays that missed some events had significantly worse performance than participants in the group operating manually
Unreliable command display (28)
Participants performed repeated trials of a simulated task that required them to diagnose the validity of pump failures within a plant based on a status display showing raw data. A command display specified whether the system had truly failed. Participants' did not adjust their decision-making appropriately based on the accuracy of the command display. Agreement between participants' decision and the display averaged 70, 78, and 83% when the display's accuracy was 60, 80, or 100%, respectively
Unreliable command display (29)
Participants viewed slides of forest terrain, each presented for about 0.75 s. A command display then recommended whether or not a camouflaged soldier had been within the picture. The participants did not adjust decisions based on accuracy of the display. When the display was correct, participants' error rates differed by less than 1% regardless of whether the aid's accuracy was set at 60, 75, or 90%. When the display was incorrect, each increase in the overall accuracy of the display was associated with an increase in the participants' error rates
Unreliable command display from computer or human (30)
Participants viewed slides of forest terrain, half containing a soldier. The accuracy of the command display was adjusted to make half as many errors as the participant. Half of the participants were told the aid was a computer program, and the other half that it was the prior participant. Before the trial, participants' expectations were that the computer would outperform a person. Even knowing that the computer performed better than they did, significantly more chose to ignore the aid than did participants receiving human aid. More participants working with automated aids than with human aids justified self-reliance by detailing a mistake their aid had made (i.e., users expected automated systems to be near perfect, but not humans)

Previous studies selected met three criteria. First, participants compared their decisions to those from displays. Second, participants could not evaluate what were correct responses. The qualities of OR management decisions depend on decisions made by other clinical services (e.g., both internists and OR managers affect emergency department bed availability). Even if data were available in real-time and people could back propagate decision-trees, there would be substantial uncertainty in the quality of any one decision. Third, displays did not critique participants' decisions. For critiquing to occur in ORs, likely clinicians would need to be tracked by indoor positioning device and have their communications monitored.

Procedure

The nine scenarios were presented to each of four groups, all with an updated paper OR schedule: with/without command display and with/without passive status display. After each participant agreed to participate in the study, he or she was randomized to one of the four groups based on a card in an opaque envelope. No limit was placed on the time that each participant chose to take to complete the study. All participants responded to every scenario.

Preceding the presentation of the scenarios, each participant was shown a 10-min slide presentation (www.FranklinDexter.net/education.htm, first half of the training materials, accessed Sept. 15, 2006). He or she was also given a one page summary of principles as posted at the OR control desk (Table 1).

Each scenario was printed in color in landscape view on one piece of paper. The command display or status display, when applicable, was printed on that one piece of paper (Fig. 1).

Primary Analyses

To evaluate whether displays would mitigate what we expected to be clinicians' tendency to make decisions that increased the clinical work per unit time in each OR, we compared the accuracy of responses to the four scenarios involving reducing expected hours of over-utilized OR time (#1, #3, #6, and #7) to random chance by the binomial test. Correct answers could be determined from the paper OR schedule (Table 4) using either the listed mean historical average OR times (1) or the scheduled OR time. Statistical significance of worse than random chance for the status displays would be particularly convincing, because the study was, by design, biased to show benefit to the status displays (see above *Displays* section).

When anesthesia providers were provided prompts in real-time for quality assurance documentation, they complied with the recommendations, even though doing so reduced their clinical work per unit time (31). Therefore, we expected that command displays would increase the percentage of managerial decisions that match the ordered priorities of Table 1, particularly when increasing clinical work per unit time in each OR results in additional over-utilized OR time. Analysis of variance with interaction was used to study the number of scenarios answered correctly by each participant. This was a 2×2 design with one factor being command display Yes/No and the other factor being status display with processed data Yes/No. All four groups received the status display with updated raw data (i.e., the OR

Table 6. Responses to Each Scenario

	% Participants with correct response						
Scenario sequence	Both	Pager	Door	Paper	Both or pager	Door or paper	Basis for decision
1	50	80	36	43	63	38	Over-utilized OR time
3	43	70	50	14	54	38	Over-utilized OR time
6	86	70	43	71	79	52	Over-utilized OR time
7	57	50	21	14	54	19	Over-utilized OR time
2	93	60	57	71	79	62	Tardiness
4	86	100	93	100	92	95	Tardiness
							Wrong recommendations
8	50	40	50	43	46	48	Safety
9	79	80	86	71	79	81	Safety
5	71	40	86	86	58	86	Surgeon absent
Participants	*14*	*10*	*14*	*7*	*24*	*21*	

The "Both or Pager" group contains participants from two groups: Both or Pager. The "Door or Paper" group contains the other two groups of participants: Door or Paper. "OR" is the abbreviation for "operating room." Description of the groups is given in the Methods and in Table 7. Among the four scenarios involving over-utilized OR time, there are 84 responses in the groups without command displays, where $84 = 4 \times 21$. Supporting question validity, among participants answering correctly and asked what they would do, 96% of correct answers were associated with matched actions (26 of 27 responses). Incorrect answers were associated with fewer ($P = 0.011$) matched actions (67%, 18 of 27 responses). There was no effect of job category on the impact of command displays on the correct response rate ($P = 0.77$). There were two scenarios (#1 and #3) involving over-utilized OR time before the poor recommendation of Scenario #5 and two scenarios (#6 and #7) afterwards. There was no before/after effect of the incorrect recommendation on the percentage increase in the correct response rate from the command display ($P = 0.54$). Statistical analysis with exact P values was performed using StatXact-7 (Cytel Software Corporation, Cambridge, MA).

schedule) (Table 4). Significant differences were checked by the nonparametric Mann–Whitney test. The χ^2 test was used to assess the impact of command displays on the accuracy of responses to the four scenarios involving reducing expected hours of over-utilized OR time.

Although in other subject areas poor recommendations from command displays can result in unsafe decisions (Table 5), we expected such situations to be recognized by participants. We assessed the accuracy of responses to the three scenarios with incomplete information (#5, #8, and #9). In addition, relevant unsolicited verbal comments were transcribed and analyzed qualitatively.

Secondary Analyses

Additional tests evaluated validity of our experimental design. As all tests supported validity, these findings are reported in the legend of Table 6.

Although five scenarios involved reducing expected hours of over-utilized OR time (#1, #3, #5, #6, and #7), one of the scenarios (#5) focused on incomplete information. That scenario (#5) was put in between the first pair of scenarios (#1 and #3) and the second pair of scenarios (#6 and #7) so that the accuracy of responses to the first pair of scenarios (#1 and #3) could be compared with that of the second pair of scenarios (#6 and #7). The impact of the scenario with incomplete information (#5) on the accuracy of responses to the other four scenarios was tested by the test for the homogeneity of the odds that the command displays increased the accuracy of responses.

All scenarios crossed job categories. Four scenarios had a decision made by anesthesiologists, three had decisions made by OR charge nurses, and two had decisions made by turnover personnel. Every participant responded to every scenario by specifying how the decision should be made. In addition, when the job category of the scenario matched that of the participant, the participant was also asked what he or she would do. For example, the text of scenario #1 (Tables 3 and 4) seen by anesthesiologists was as in Figure 1, with the last two sentences: ". . . which of the two ORs should you start first, ASC2 or ASC4? Which one would you start?" The OR nurses and turnover technicians did not have the last sentence. Concordance between participants' responses and reports of actions was tested by Fisher's exact test.

An additional *post hoc* test examined whether there was an effect of job category on the impact of command displays on decisions involving over-utilized OR time. Recommendation from a command display may have a larger incremental benefit on getting an answer correct for scenarios that consider decisions for which the job category of the scenario does not match the job category of the participant. The opposite could also be argued. Statistical analysis was performed by testing for the homogeneity of the odds ratio.

RESULTS

Participants without command displays answered the scenarios involving over-utilized OR time less accurately than random chance ($P = 0.011$, 31 of 84 responses) (Table 6). This result was consistent with the staff using the status displays to increase the clinical work per unit time in each OR (i.e., keeping the ORs busy). Command displays significantly increased the correct response rate ($P < 0.001$).

Among all scenarios, simulated status displays with processed information (e.g., expected over-utilized OR time) had no effect on the accuracy of decision-making versus displays with raw data (i.e.,

Table 7. Decision-Making Involving More Than One OR

		Correct responses out of 9		
Group name	Brief description	Mean ± standard error	Median	*Participants*
Paper	Status display with raw data (i.e., updated paper OR schedule)	5.1 ± 0.1	5.0	7
Door	Status display with processed data (i.e., text above OR door) plus paper OR schedule	5.1 ± 0.3	5.0	14
Pager	Command display providing recommendation (i.e., text pager) plus paper OR schedule	5.9 ± 0.4	6.0	10
Both	All three aids	6.4 ± 0.3	6.5	14

Table 8. Verbal (Unsolicited) Comments by Participants in Groups Receiving Command Displays Regarding the Scenarios with Wrong Recommendations

From an anesthesiologist: "The priority is always supposed to be safety, but the pagers don't always seem consistent with that. What are you supposed to do?"
From an anesthesia technician: "Some of the pager messages didn't make sense. Why would you do what the system tells you to when you know it's not right?"
From a nursing assistant: "The pager said to do something in a couple of the scenarios that seemed wrong. It would make me think twice about doing what it said."

updated paper schedule) ($P = 0.40$) (Table 7). Command displays increased the accuracy of decision-making ($P = 0.010$; Mann–Whitney $P < 0.001$). The combination did not increase the accuracy further ($P = 0.40$).

Previous studies in other subject areas found participants comply with poor recommendations from command displays (Table 5). In our scenario for which the command displays provided a poor recommendation and safety was unaffected, participants complied with the incorrect recommendation more often than did the other participants (Table 6). In the two scenarios for which the command displays provided recommendations that adversely affected safety, participants appropriately ignored advice (Table 6). Unsolicited verbal comments showed a resulting lack of trust (23) in the command displays (Table 8), matching results of studies in other subject areas (23).

DISCUSSION

When decision-making on the day of surgery involves multiple ORs, decisions to increase clinical work per unit time in each OR often conflict with decisions to reduce over-utilized OR time. Command displays (recommendations) increased consideration of over-utilized OR time, but status displays did not.

Reducing over-utilized OR time has advantages for OR nurses, housekeepers, and anesthesia providers: consistent decision-making on the day of surgery, more predictable work hours, fewer hand-offs during cases, and reduced staffing costs (1,2). Depending on staff scheduling, reducing over-utilized OR time can also reduce scheduled overtime (e.g., late call list) and unscheduled (mandatory) overtime. The premise of making decisions to reduce the hours that ORs finish late is so simple that in our experience people often find the concept to be so obvious that they are annoyed by being taught it. Even though we removed organizational or time pressure by using simulation, the participants made decisions that increased the clinical work per unit time in each OR.

Nurses and physicians receive clinical training while completing lists of cases in single ORs, not while working at OR control desks or while medically directing multiple ORs. Thus, socialization of OR culture occurs in situations for which making decisions to increase clinical work per unit time is advantageous. Our results suggest that the tendency is then applied inappropriately to managerial decisions involving multiple ORs. This socialization likely is reinforced by the presence of intangible and tangible rewards for working fast (see Introduction). This tendency is explored further in our companion paper (32).

Although many communication episodes at the OR control desk involve decisions that affect multiple ORs (1,33,34), previous studies of status displays have not considered such decisions (Table 2). Status displays (Tables 6 and 7) were insufficient to change decision-making, even though the situations studied were deliberately simple. In contrast, command displays were efficacious (Table 7), particularly for the economically most important decisions (1) (i.e., those to reduce over-utilized OR time) (Table 6). Thus, we recommend that research in computer–human interaction for decision-making on the day of surgery focuses on decision aids that make recommendations. Such displays are *not* Gantt charts displaying information about multiple ORs, whether as large displays in hallways or on web pages, and whether the information provided is raw [i.e., like an updated OR schedule (24)] or processed [e.g., expected over-utilized OR times (24)].

Command displays have disadvantages when there are errors in recommendations due to the decision aids having incomplete knowledge (Table 5). Such situations should be expected when coordinating ORs, because much of the information is not obtained from information systems, but from direct observation (e.g., a patient being wheeled down the hall) and social networks (e.g., asking a surgeon in passing) (35). One disadvantage is that decision-makers often comply with the erroneous recommendations (Tables 5 and 6). Preliminary findings are that when safety is adversely affected, erroneous recommendations are recognized (Table 8) and ignored (Table 6), unlike in other subject areas (domains). The resulting disadvantage is that there is a reduction in trust in all of the command display's recommendations, good or bad (23). Findings from other domains are that, when some recommendations are erroneous, appropriate recommendations are followed more often when from a human than from a computer (Table 5) (30). Therefore, we expect continuation of the current managerial model whereby most decisions involving multiple ORs are made at an OR control desk and then communicated verbally to other clinicians (33). We also recommend strongly against using a reduction in communication episodes (e.g., phone calls) at OR control desks (34) as an end-point of success of information systems, in lieu of end-points of decisions made. Our results show that the successful distribution of information on the day of surgery *cannot* be taken as sole evidence of value of communication tools (e.g., displays).

Our finding of a benefit for accurate real-time OR information system data is novel, even though it may seem intuitive. The economic importance of decision-making on the day of surgery is very small (1) compared to having the distribution of services among ORs (blocks) planned (3) right several months before the day of surgery, OR allocations (staffing) planned (1,21,36) right a few months before the day of surgery, and scheduling (1,19) the cases right a few weeks to days before the day of surgery. In contrast to the findings in the current paper, decisions for allocating blocks, planning staffing, and scheduling cases are robust to even large errors in OR data (1,3,19,25,36).

Usefulness Varies Among Facilities

Because OR allocations are driven by the relative cost of an hour of over-utilized OR time to an hour of under-utilized OR time, if a facility sets this ratio to a high value, there can be sufficient excess capacity (i.e., under-utilized OR time) that rarely would one surgeon need to wait for another (2,3,21). On the day of surgery, decision-making by the ordered priorities would then usually simplify to reducing patient and surgeon waiting (Table 1). For example, if at a reader's facility there are <8 h of cases in most ORs every workday, then the results of our study are unlikely to be useful.

The studied recommendations on the day of surgery are likely to be useful only for facilities that also follow the ordered priorities (Table 1) when making OR allocation (2,21,36) and case scheduling (1,2,19) decisions before the day of surgery, as is economically rational (1–3).

Limitations

Reducing over-utilized OR time can provide direct benefits to OR staff, but less likely to customers of the surgical suite, the surgeons. The argument could be made that anesthesiologists acted to reduce surgeon waiting in lieu of over-utilized OR time to reduce surgeon and patient waiting (i.e., to increase surgeon satisfaction). Alternatively, they could have made decisions to finish the available clinical work as early in the day as possible. Additional observational studies in our companion paper show that neither hypothesis can explain our results (32). Rather, the behavior was consistent with the heuristic of keeping each clinician busy when present.

Our study has the weakness that it was a simulation (i.e., not naturalistic decision-making). This weakness is mitigated in our companion paper (32) in which we show that our results (Tables 6 and 7) are matched by two observational studies.

The studies were performed at one academic hospital. We suspect that confounding effects of organizational culture had a small influence on results based on the results explaining previous findings from other facilities [Table 2, Ref. (31)] and matching findings from other subject areas (Table 5). Nevertheless, we studied decision-making at just one hospital, and that one hospital was an academic hospital.

Finally, we did not detect heterogeneity of decision-making among clinicians with different backgrounds and jobs (Table 6). However, our statistical power was likely weak to detect such differences if present.

CONCLUSIONS

Clinicians at a hospital made managerial decisions that were consistent with a tendency to increase the clinical work per unit time in each OR. This work ethic likely is supported by intangible and tangible rewards, and seems to be a reasonable basis for decisions involving individual cases. However, it can be disadvantageous for decisions involving multiple ORs. Command displays with recommendations may be more effective at changing decisions than education and distributed status displays. Future research can focus on how to create better recommendations and how to increase the accuracy of the information used to make the recommendations to prevent a reduction in trust.

REFERENCES

1. Dexter F, Epstein RH, Traub RD, Xiao Y. Making management decisions on the day of surgery based on operating room efficiency and patient waiting times. Anesthesiology 2004;101:1444–53
2. McIntosh C, Dexter F, Epstein RH. Impact of service-specific staffing, case scheduling, turnovers, and first-case starts on anesthesia group and operating room productivity: tutorial using data from an Australian hospital. Anesth Analg 2006;103:1499–516

3. Dexter F, Ledolter J, Wachtel RE. Tactical decision making for selective expansion of operating room resources incorporating financial criteria and uncertainty in sub-specialties' future workloads. Anesth Analg 2005;100:1425–32
4. Xiao Y, Hu P, Hao H, Ho D, Dexter F, Mackenzie CF, Seagull FJ, Dutton R. Algorithm for processing vital sign monitoring data to remotely identify operating room occupancy in real-time. Anesth Analg 2005;101:823–9
5. De Deyne C, Heylen R. Introduction of an operating room information management system improved overall operating room efficiency. Stud Health Technol Inform 2004;110:61–7
6. Sullivan EE. How to work smarter, not harder: a PACU tracking system. J Perianesth Nurs 2003;18:123–5
7. Sullivan EE. The patient flow beeper: a system to decrease PACU transfer delays. J Perianesth Nurs 2005;20:438–40
8. Meyer MA, Sokal SM, Sandberg W, Chang Y, Daily B, Berger DL. INCOMING!–A web tracking application for PACU and post-surgical patients. J Surg Res 2006;132:153–8
9. Abenstein JP, Allan JA, Ferguson JA, Deick SD, Rose SH, Narr BJ. Computer-based anesthesiology paging system. Anesth Analg 2003;97:196–204
10. Topp R, Walsh E, Sanford C. Can providing paging devices relieve waiting room anxiety? AORN J 1998;67:852–61
11. Vitez TS, Macario A. Setting performance standards for an anesthesia department. J Clin Anesth 1998;10:166–75
12. Espin SL, Lingard LA. Time as a catalyst for tension in nurse-surgeon communication. AORN J 2001;74:672–82
13. Riley R, Manias E. Governing time in operating rooms. J Clin Nurs 2006;15:546–53
14. Lingard L, Garwood S, Poenaru D. Tensions influencing operating room team function: does institutional context make a difference? Med Educ 2004;37:691–9
15. Espin S, Lingard L, Baker GR, Regehr G. Persistence of unsafe practice in everyday work: an exploration of organizational and psychological factors constraining safety in the operating room. Qual Saf Health Care 2006;15:165–70
16. Carayon P, Hundt AS, Alvarado CJ, Springman SR, Ayoub P. Patient safety in outpatient surgery: the viewpoint of the healthcare providers. Ergonomics 2006;49:470–85
17. Overdyk FJ, Harvey SC, Fishman RL, Shippey F. Successful strategies for improving operating room efficiency at academic institutions. Anesth Analg 1998;86:896–906
18. St. Jacques PJ, Patel N, Higgins MS. Improving anesthesiologist performance through profiling and incentives. J Clin Anesth 2004;16:523–8
19. Dexter F, Traub RD. How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. Anesth Analg 2002;94:933–42
20. Tessler MJ, Kleiman SJ, Huberman MM. A "zero tolerance for overtime" increases surgical per case costs. Can J Anaesth 1997; 44:1036–41
21. Strum DP, Vargas LG, May JH. Surgical subspecialty block utilization and capacity planning. A minimal cost analysis model. Anesthesiology 1999;90:1176–85
22. Dexter F, Smith TC, Tatman DJ, Macario A. Physicians' perceptions of minimum time that should be saved to move a surgical case from one operating room to another: internet-based survey of the Association of Anesthesia Clinical Directors' (AACD) members. J Clin Anesth 2003;15:206–10
23. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors 2004;46:50–80
24. Spangler WE, Strum DP, Vargas LG, May JH. Computer-based operational and decision support for surgical services. Int J Healthc Technol Manag 2000;2:15–36
25. Dexter F, Ledolter J. Bayesian prediction bounds and comparisons of operating room times even for procedures with few or no historical data. Anesthesiology 2005;103:1259–67
26. Sarter NB, Schroeder B. Supporting decision making and action selection under time pressure and uncertainty: the case of in-flight icing. Hum Factors 2001;43:573–83
27. Metzger U, Parasuraman R. Automation in future air traffic management: effects of decision aid reliability on controller performance and mental workload. Hum Factors 2005;47:35–49
28. Wiegmann DA, Rich A, Zhang H. Automated diagnostic aids: the effects of aid reliability on users' trust and reliance. Theor Issues Ergon Sci 2001;2:352–67
29. Dzindolet MT, Pierce LG, Beck HP, Dawe LA, Anderson BW. Predicting misuse and disuse of combat identification systems. Mil Psychol 2001;13:147–64
30. Dzindolet MT, Pierce LG, Beck HP, Dawe LA. The perceived utility of human and automated aids in a visual detection task. Hum Factors 2002;44:79–94
31. Vigoda MM, Gencorelli F, Lubarsky DA. Changing medical group behaviors: increasing the rate of documentation of quality assurance events using an anesthesia information system. Anesth Analg 2006;103:390–5
32. Dexter F, Dow AJ, Lee JD, Lubarsky DA. A psychological basis for clinicians' operating room managerial decision-making on the day of surgery. Anesth Analg 2007;105:430–4
33. Moss J, Xiao Y, Zubaidah S. The operating room charge nurse: coordinator and communicator. J Am Med Inform Assoc 2002;9:S70–4
34. Moss J, Xiao Y. Improving operating room coordination: communication pattern assessment. J Nurs Adm 2004;34:93–100
35. Seagull FJ, Xiao Y, Plasters C. Information accuracy and sampling effort: a field study of surgical scheduling coordination. IEEE Trans Syst Man Cybern A 2004;34:764–71
36. Epstein RH, Dexter F. Uncertainty in knowing the operating rooms in which cases were performed has little effect on operating room allocations or efficiency. Anesth Analg 2002;95: 1726–30