Remote video auditing with real-time feedback in an academic surgical suite improves safety and efficiency metrics: a cluster randomised study

Frank J Overdyk,1 Oonagh Dowling,2 Sheldon Newman,1 David Glatt,1 Michelle Chester,3 Donna Armellino,4 Brandon Cole,3 Gregg S Landis,5 David Schoenfeld,6 John F DiCapua1

ABSTRACT
Importance Compliance with the surgical safety checklist during operative procedures has been shown to reduce inhospital mortality and complications but proper execution by the surgical team remains elusive.

Objective We evaluated the impact of remote video auditing with real-time provider feedback on checklist compliance during sign-in, time-out and sign-out and case turnover times.

Design, setting Prospective, cluster randomised study in a 23-operating room (OR) suite.

Participants Surgeons, anaesthesia providers, nurses and support staff.

Exposure ORs were randomised to receive, or not receive, real-time feedback on safety checklist compliance and efficiency metrics via display boards and text messages, followed by a period during which all ORs received feedback.

Main outcome(s) and measure(s) Checklist compliance (Pass/Fail) during sign-in, time-out and sign-out and case turnover times.

Results Sign-in, time-out and sign-out PASS rates increased from 25%, 16% and 32% during baseline phase (n=1886) to 64%, 84% and 68% for feedback ORs versus 40%, 77% and 51% for no-feedback ORs (p<0.004) during the intervention phase (n=2693). Pass rates were 91%, 95% and 84% during the all-feedback phase (n=2001). For scheduled cases (n=1406, 71%), feedback reduced mean turnover times by 14% (41.4 min vs 48.1 min, p<0.004), and the improvement was sustained during the all-feedback period. Feedback had no effect on turnover time for unscheduled cases (n=587, 29%).

Conclusions and relevance Our data indicate that remote video auditing with feedback improves surgical safety checklist compliance for all cases, and turnover time for scheduled cases, but not for unscheduled cases.

INTRODUCTION
The operating room (OR) is a complex system that requires a carefully coordinated workflow to ensure both safe and efficient operation. Video monitoring of ORs has been described on a small scale by Hu et al1 who recorded 10 operations and used the video offline to identify and analyse safety and efficiency deficits. Makary discussed the potential of video recording in driving surgical quality improvements and eliminating surgical ‘never events’ as well as the importance of effective teamwork in driving patient safety.2 3 The use of a surgical safety checklist has been shown to reduce mortality and complication rates in a large multicentre trial and improve outcomes during simulated events.4 5 We hypothesised that remote video auditing (RVA) with real-time feedback would improve compliance with safety processes and OR efficiency by improving accountability and communication, among the OR team. We engaged a commercial company, Arrowsight, which provides RVA services to promote safety and efficiency in fast food, meatpacking and other industries, to audit our large OR suite in real time with two objectives. The first was to study the impact of RVA with real-time feedback (FB) on compliance by OR staff with the...
sign-in, time-out and sign-out elements of the WHO’s surgical safety checklist. The second was to quantify the impact of RVA with feedback on driving OR efficiency, by measuring the effect of feedback on turnover time, defined as ‘patient out of OR’ to next ‘patient in OR’ for consecutive cases.

METHODS
The North Shore-LIJ Health System Institutional Review Board approved this study (Ref: 13-208B) with a waiver of informed consent and deemed study subjects all OR staff, including OR nurses, and physicians (not patients) employed at the Long Island Jewish Medical Center’s main OR suite. A video stream from wide-angle cameras in each of the 23 ORs included was transmitted through an encrypted virtual private network to offsite auditors (Arrowsight) trained in recognising and timing milestones typical of an OR case. The auditors had no knowledge of, or involvement in, the study protocol or details such as OR allocation to treatment groups of the OR cases they reviewed. Cameras were not placed to monitor the operative field and the audio component of the video stream was not included for auditors to review. The video stream was blurred to minimise the identification of individuals, and deleted upon audit completion (figure 1). The safety and efficiency metrics collected were chosen to demonstrate OR team performance only and did not identify individual team members.

Arrowsight used a 10-person audit team to identify and time OR milestones from the video stream to the nearest 20 s. Video was audited every 2 min and ‘real-time’ feedback metrics were posted to the display boards or sent as email or text alerts to the OR team within 3 min of each audit start. For the OR team to receive a ‘PASS’ for each of the three elements of the checklist (sign-in, time-out, sign-out), auditors looked for the surgeon, circulating nurse and anaesthesia provider to (1) read the script from yellow backed checklist, clearly visible on camera (The Power of Now—Eckhart Tolle, see online supplementary figure S1), (2) be engaged and participate without distraction, (3) take the minimum amount of time required to complete the checklist and (4) complete each element in the proper order during the course of the case. For instance, a time-out had to be performed prior to the last suture being placed during closure of the incision. From 1 August through 30 September 2013, baseline compliance with the elements of the checklist and case turnover times were measured daily until 19:00 h for both scheduled and unscheduled (‘add-on’) cases. During October 2013, staff and physicians were re-educated on the appropriate protocol for performing the elements of the surgical safety checklist and the objectives and content of the feedback provided by RVA. During the intervention phase from 18 November 2013 until 14 February 2014, ORs were stratified by mean non-operating time and randomised to receive RVA with or without feedback. Feedback consisted of OR status and OR team efficiency metrics conveyed by 127 cm light crystal display boards installed throughout the OR hallways, and nursing, surgical and anaesthesia lounges (figures 2 and 3). To improve communication and efficiency of the OR, staff and surgeons received text alerts on their smart phone and by email on OR status changes and safety alerts, such as ‘next patient in OR’, ‘drape down’, ‘time out failed’ or ‘team distracted’. These notifications allowed surgeons, anaesthesia providers and nurse managers to intervene to minimise case delays or remedy checklist non-compliance. OR Efficiency boards displayed safety and efficiency metrics with OR-specific targets, established by OR leadership, and colour coded to indicate how the metrics compared with benchmark values derived from historical data specific for each OR and day of the week (figure 3). Aggregate safety and efficiency data for the OR suite was displayed at the top of the board, alongside a projected ‘minutes gained’ per OR for that day (assuming three cases per day, per OR). Patient safety metrics were labelled ‘PS’ so as not to alarm patients who may view the boards during transport to the OR.

A team comprising an OR manager, an anaesthesiologist and an anaesthesia technician monitored the status and performance boards, and provided logistical assistance and reinforcement for OR teams whose metrics deviated significantly from the targets. This team immediately notified ORs that received a ‘FAIL’ for any component of the surgical safety checklist during a case so that corrective action could be taken and the sign-in, time-out or sign-out could be repeated in a compliant manner.

During the final phase of the study, from 17 February through 18 April 2014, all ORs received RVA with feedback.

Statistical methods
Statistical analyses were performed with the use of the Statistical Analysis System statistical software package, V9.1 (SAS Institute, Cary, North Carolina, USA). Outcomes were the impact of RVA with feedback on (a) adherence to sign-in, sign-out and time-out components of the checklist and (b) case turnover time.

ORs were selected as units for stratified randomisation to feedback or no feedback during the intervention period. To ensure balanced randomisation of rooms with surgical services that perform predominantly complex and lengthy operations (longer turnover times, fewer cases per day) and those with shorter operations (shorter turnover times, more cases per day) to feedback and no-feedback treatment groups, ORs were classified as ‘slow’ or ‘fast’ based on a preliminary analysis of their mean daily non-

operative time during the baseline phase (see online supplementary figures S2 and S3). Rooms with mean non-operative time of less than 300 min and average of 3.8 cases per day were classified as ‘fast’ rooms while rooms with mean non-operative time of greater than 300 min and average of 2.1 cases per day were classified as ‘slow’ rooms (see online supplementary figure S2). Subsequently, the ORs were stratified into fast and slow groups and randomised to feedback and no-feedback treatment groups.

The analysis was conducted using the generalised estimating equation approach for analysing cluster-randomised clinical trials. The independent variables were the mean baseline for the variable in the OR, the cluster (fast or slow) and the treatment (feedback or no feedback). Safety was assessed as an audit ‘PASS’ to the four requirements during sign-in, time-out and sign-out. The frequency of adherence was analysed using a logistic model with the independent variables listed above. Case segment durations were log-transformed and turnover time durations were analysed using a linear model. Raw rates, geometric means and SDs of each of the variables are presented as descriptive statistics.

Figure 1  Wide-angle camera displaying blurred image of operating room and use of yellow surgical safety checklist.

Figure 2  OR ‘Status’ display board during all-feedback phase of study. OR, operating room.
Inter-rater reliability for RVA with a sample of four auditors was estimated by examining the surgical safety checklist compliance and time segment metrics for 30 randomly selected cases. Intraclass correlation coefficients and Fleiss’ kappas were determined for segment durations and checklist compliance metrics.

RESULTS
From 18 November 2013 until 14 February 2014, 23 ORs (n=2693) were randomised to receive RVA with or without feedback. Online supplementary figure S3 shows the Consort flow diagram for the intervention phase of the study. Cases for days with OR disruptions such as major holidays (four) and weather emergencies (four) were excluded.

Surgical safety checklist compliance
The overall rate of sign-in compliance for ORs increased from 25.4% during the baseline phase to 64.1% in feedback rooms and 39.7% in no-feedback rooms during the intervention phase. Feedback conferred a 2.75-fold (p<0.0001) significantly increased odds of compliant sign-in execution versus no feedback (table 1).

Time-out compliance increased from 16.3% during baseline to 84.4% in feedback rooms and 62.7% in no-feedback rooms during the intervention. Feedback conferred a 3.37-fold (p=0.0004) increased odds of time-out compliance versus no feedback.

The overall baseline sign-out compliance rate improved from 31.7% to 65.8% in feedback rooms during the intervention with feedback resulting in a 2.40-fold (p<0.0001) increased odds of sign-out compliance versus no feedback.

Overall compliance rates during the final, all-feedback phase improved to 90.8%, 95.3% and 84.3% for sign-in, time-out and sign-out, respectively. There were no wrong-patient, wrong-site/side or retained-foreign body adverse events during the study.

Efficiency
Baseline raw mean turnover times for all cases, scheduled cases and unscheduled cases were 52.9, 44.5 and 80.3 min, respectively. During the intervention phase, the mean turnover time for all ORs receiving feedback (n=1130) was 48.4 min, not statistically lower than the 54.7 min turnover time for ORs not receiving feedback (n=863, table 2). When the analysis was limited to scheduled cases (n=1406, 71% of cases), mean turnover times in ORs with feedback (41.4 min, n=795) were 14% shorter than in ORs with no feedback (48.1 min, n=611, p=0.0041). Within this

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Regression results, checklist% compliance (n) during the intervention phase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw rate no feedback</td>
</tr>
<tr>
<td>Sign-in</td>
<td>39.7% (724/1824)</td>
</tr>
<tr>
<td>Time-out</td>
<td>62.7% (1281/2043)</td>
</tr>
<tr>
<td>Sign-out</td>
<td>40.9% (739/1809)</td>
</tr>
</tbody>
</table>

*Surgical safety checklist. †General Estimating Equations Model.
group, mean turnover time in ‘fast’ ORs was 20% less in feedback (n=580) versus no-feedback (n=505) ORs (37.2 min vs 46.8 min, p=0.0004). Feedback had no significant effect on mean turnover time for unscheduled cases (n=587) or slow OR scheduled cases (n=321). Turnover time improvements for scheduled cases were maintained during the all-feedback phase (‘slow’ ORs: 42.5 min; ‘fast’ ORs: 39.8 min).

**DISCUSSION**

Our Health System adopted the WHO surgical safety checklist in 2007 and despite a near-perfect record of compliance with the checklist as measured by random ‘in OR’ audits, ‘never events’ continued to occur. In the spring of 2013, Health System leadership designed a revised surgical safety checklist process requiring physical use of a scripted checklist in a ‘call and response’ format, adequate engagement by the OR team, and proper timing and duration for all elements of the checklist. This checklist was intended to simplify the process and reduce the addressed behaviour deficits observed during OR audits and previously described in the literature.

For example, Borchard et al.\(^6\) found that although checklists conferred a reduced relative risk for mortality and complications of 0.57 and 0.63, respectively, compliance was highly variable. The most frequent reasons providers admitted to not being compliant with the surgical safety checklist is that they forgot the checklist, forgot to address elements of the checklist or did not ‘have time’. Vats et al. noted that practitioners were frequently hurried, missed elements of the checklists, were dismissive of the process and did not pay attention. The ‘sign-out’, a critical process intended to prevent the most common ‘never event’, a retained foreign object during surgery, was often skipped altogether.\(^7\) Thus, in spite of global adoption of the WHO surgical safety checklist in the USA, ‘never events’ continue to occur at a rate of over 4000 per year.\(^8\)

We recognised that closing critical gaps and improving participation and team communication with checklists, as described by Johnston et al.,\(^9\) would likely lower the odds of a complication resulting in morbidity or mortality.\(^10\) We used our Center for Learning and Innovation (Simulation Center) to educate and engage surgical, anaesthesia and nursing leadership on the benefits of the revised surgical safety checklist process and RVA. Conley et al.\(^11\) have shown that improving outcomes with the checklist depends on effective communication and engagement of the stakeholders. Surgeons, nurses, anaesthesiologists and support staff leaders agreed that the time-out, sign-in and sign-out should be reasonably accomplished in 60, 30 and 30 s, respectively, after simulations at Center for Learning and Innovation (CLI); thus, auditors used these as the required minimum time for these tasks.

Our results confirm that immediate feedback of staff compliance with the audited elements of the checklist is effective in sustainably improving compliance and that improvement in PASS rates continued during the all-feedback period.

**Efficiency**

The vast literature on improving OR efficiency lacks a consistent methodology on collecting objective and reproducible OR milestones. Paper-based or computer-based entry of procedural times by OR staff is often inaccurate because of multitasking, unsynchronised clocks, the Hawthorne effect or bias from stakeholders entering the data.\(^12\) RVA provided us ‘hands free’, objective, accurate, reproducible and unbiased surgical milestones in real time, which has never been described.

Previously, OR managers have used a variety of process improvement techniques to improve OR efficiency and turnover times, including industrial best practices such as Lean Six Sigma, parallel processing and financial incentive programmes with variable results.\(^13–18\) For all scheduled and unscheduled cases, our improvement in efficiency was 10% and not statistically significant (p=0.11). However, our subgroup analysis was compelling and statistically significant with 14% improvement in scheduled cases, and a 20% improvement in ‘fast’ rooms with greater numbers of shorter scheduled cases and less non-operative time (‘fast’ ORs). This last subgroup

---

Table 2  Regression results, operating room turnover times in minutes ±SD (n) during intervention

<table>
<thead>
<tr>
<th>Type of room/procedure</th>
<th>Raw mean no feedback</th>
<th>Raw mean feedback</th>
<th>Difference in log mean</th>
<th>p Value</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All rooms</td>
<td>54.7±30.1 (863)</td>
<td>48.4±27.9 (1130)</td>
<td>−0.104±0.065</td>
<td>0.1090</td>
<td>−0.231</td>
<td>0.023</td>
</tr>
<tr>
<td>All rooms scheduled</td>
<td>48.1±24.3 (611)</td>
<td>41.4±20.8 (795)</td>
<td>−0.151±0.056</td>
<td>0.0041</td>
<td>−0.261</td>
<td>−0.041</td>
</tr>
<tr>
<td>All rooms Unscheduled</td>
<td>74.5±43 (252)</td>
<td>70.2±40.1 (335)</td>
<td>−0.068±0.063</td>
<td>0.2818</td>
<td>−0.191</td>
<td>0.055</td>
</tr>
<tr>
<td>Fast rooms scheduled</td>
<td>46.8±22.5 (505)</td>
<td>37.2±16.2 (580)</td>
<td>−0.206±0.058</td>
<td>0.0004</td>
<td>−0.320</td>
<td>−0.092</td>
</tr>
<tr>
<td>Slow rooms scheduled</td>
<td>69.3±37.2 (106)</td>
<td>65.8±38.0 (215)</td>
<td>0.017±0.118</td>
<td>0.887</td>
<td>−0.215</td>
<td>0.249</td>
</tr>
</tbody>
</table>

Inter-rater agreement: four auditors scored all checklist elements, (1) use of checklist, (2) team attentiveness, (3) required duration and (4) proper sequence with Fleiss’ kappas of 1, 95% CI (1.0 to 1.0) for 30 randomly selected cases. Intraclass correlation coefficients (ICCs) for sign-in, time-out and sign-out ‘minimum time’ durations were 0.976, 95% CI (0.957 to 0.988), 0.953, 95% CI (0.917 to 0.976) and 0.999, 95% CI (0.998 to 0.999), respectively. ICCs for timed segment durations were 1.0, 95% CI (1.0 to 1.0) for all segments.
represents roughly half of the operative room cases and 39% of the non-operative time spent. The most common causes for delays between OR cases are waiting for the surgeon, the patient, cleaning staff or surgical instruments to arrive and our hypothesis that providing OR teams case and patient status information through text messages, emails and display boards would significantly reduce case turnover times was verified for scheduled cases. Non-emergent add-on cases may have long gaps between cases as patients or surgeons may be unavailable, and just a few long gaps can distort data. Unlike previous studies that arbitrarily removed turnover times greater than 90 min to eliminate the effect of non-consecutive cases and add-on cases, we did not use a turnover time cut-off to eliminate the effect of non-consecutive cases and add-on cases. While reducing OR down time for non-consecutive cases may likely be done through scheduling strategies as suggested by Dexter et al., we intend to study additional logistics support embedded in RVA with feedback that may remedy this shortcoming.

This study presented unique technical, logistical, legal and regulatory challenges. The two most frequently cited concerns with video auditing during medical procedures are the risk of legal discovery and patient privacy. To address these concerns, we worked to ensure that (1) the video/auditing process eliminated the likelihood of identifying patients and OR staff, (2) audio not be transmitted to the auditors and (3) individual performance not be targeted. Adherence to these principles facilitated adoption by workers, both union and non-union. The video stream was intentionally blurred so as to make staff indistinguishable beyond size and gender. Offsite audio audits were considered too intrusive and, therefore, were not be audited by Arrowsight. We designed the study to measure team performance and not individual performance, which was a critical step in recruiting the support of all OR stakeholders.

A major limitation of this study was not being able to use the audio channel of the video stream. We recognise that the four visual elements of the checklist audited do not guarantee that each element of the checklist has actually been addressed. However, we found that the flaws in checklist compliance were so common and fundamental, such as distraction and lack of participation, that video alone with feedback dramatically improved safety behaviour. This is a good first step. However, to further improve the effectiveness of the surgical safety checklist, we agree that future studies should incorporate audio monitoring. Another limitation to the study design during the feedback–no feedback phase was that staff could not be limited to work in a feedback or no-feedback OR; thus, a ‘contamination’ effect may have occurred, where providers working in a ‘no-feedback’ OR could use practices acquired in a ‘feedback’ OR. This would have a tendency to reduce the effect of the intervention, but only for the checklist since the efficiency gains are dependent on the case status updates by text and displays. However, given the large sample size, differences in checklist compliance were still statistically significant between groups. Although the goal of our surgical safety checklist protocol is to eliminate ‘never events’, they were not an outcome in this study, as their incidence is too low to detect.

The total cost of implementing RVA had three cost components: one-time video equipment cost of approximately US $4000 per camera; one-time RVA set-up and onsite consulting training fee of US $7500 per OR, and an RVA service charge of US $40/day per OR. While a return on investment analysis is beyond the scope of this manuscript, Long Island Jewish Medical Center leadership deemed the tangible efficiency gains and anticipated improvements in patient safety and surgeon satisfaction sufficient to commit to RVA beyond the study period.

In summary, we find that improved efficiency and compliance to safety protocols in a complex environment such as the OR can result from direct observation, measurement and immediate feedback to the OR team. We also conclude that RVA of safety and efficiency processes in the OR is tolerated and supported by physicians and OR staff. In addition to its role in providing effective logistical support and promoting sustainable performance improvements, we believe that acceptance among staff was facilitated by our emphasis on individual privacy, teamwork and non-punitive reinforcement when metrics diverged from targets. We propose to continue studying the long-term efficacy and acceptability of large-scale RVA of surgical procedures, the impact on patient adverse events as well as improving its technical capabilities.

Author affiliations
1Department of Anesthesiology, Hofstra North Shore-LIJ School of Medicine, Long Island Jewish Medical Center, New Hyde Park, New York, USA
2Department of Medicine, Hofstra North Shore-LIJ School of Medicine, Hempstead, New York, USA
3Department of Anesthesiology, Long Island Jewish Medical Center, New Hyde Park, New York, USA
4Department of Anesthesiology, North Shore-LIJ Health System, Lake Success, New York, USA
5Department of Surgery, Hofstra North Shore-LIJ School of Medicine, Hempstead, New York, USA
6Department of Biostatistics, Massachusetts General Hospital Biostatistics Center, Harvard Medical School, Boston, Massachusetts, USA

Acknowledgements The authors are grateful to the operating room staff and administrators of Long Island Jewish Medical Center for their support and dedication to this research initiative.

Contributors FJO, OD and JFD contributed substantially to the conception and design of the work, the acquisition, analysis and interpretation of data, and the drafting of the manuscript. SNB, DG, MC, DA, BC and GSL contributed substantially to the acquisition, analysis and interpretation of data. DS contributed substantially to the design of the work, the analysis and interpretation of data, and the drafting of the manuscript.
REFERENCES


