Surgical Intelligence Can Lead to Higher Adoption of Best Practices in Minimally Invasive Surgery

Gerald M. Fried, MD,✉ Monica Ortenzi, MD,†‡ Danit Dayan, MD,§ Eran Nizri, MD,§ Yuval Mirkin, MD,† Sari Maril, PhD,† Dotan Asselmann, MSc,† and Tamir Wolf, MD, PhD †*

Objective: To examine the use of surgical intelligence for automatically monitoring critical view of safety (CVS) in laparoscopic cholecystectomy (LC) in a real-world quality initiative.

Background: Surgical intelligence encompasses routine, artificial intelligence-based capture and analysis of surgical video, and connection of derived data with patient and outcomes data. These capabilities are applied to continuously assess and improve surgical quality and efficiency in real-world settings.

Methods: Laparoscopic cholecystectomies conducted at 2 general surgery departments between December 2022 and August 2023 were routinely captured by a surgical intelligence platform, which identified and continuously presented CVS adoption, surgery duration, complexity, and negative events. In March 2023, the departments launched a quality initiative aiming for 75% CVS adoption.

Results: Two hundred seventy-nine procedures were performed during the study. Adoption increased from 39.2% in the 3 preintervention months to 69.2% in the final 3 months ($P < 0.001$). Monthly adoption rose from 33.3% to 75.7%. Visualization of the cystic duct and artery accounted for most of the improvement; the other 2 components had high adoption throughout. Procedures with full CVS were shorter ($P = 0.007$) and had fewer events ($P = 0.011$) than those without. OR time decreased following intervention ($P = 0.033$).

Conclusions: Surgical intelligence facilitated a steady increase in CVS adoption, reaching the goal within 6 months. Low initial adoption stemmed from a single CVS component, and increased adoption was associated with improved OR efficiency. Real-world use of surgical intelligence can uncover new insights, modify surgeon behavior, and support best practices to improve surgical quality and efficiency.

Keywords: artificial intelligence, bile duct injury, computer vision, critical view of safety, laparoscopic cholecystectomy, surgical intelligence

(*Ann Surg* 2024;280:525–534)

Though standardization in surgery is starting to take root, the implementation of widely accepted evidence-based practices has been a largely insurmountable obstacle.[1] Practical and logistical considerations have hindered outcomes research in surgery, making it difficult to identify best surgical practices and relate them to measures of quality and efficiency. When we do manage to identify such practices, it is extremely challenging to determine whether surgeons accurately and consistently adopt them, outside of small clinical trials.

Perhaps the most familiar illustration of this phenomenon is the critical view of safety (CVS) in laparoscopic cholecystectomy (LC), a rare example of a broadly accepted evidence-based surgical best practice. Since its introduction,[2] CVS has been solidified as an effective measure for reducing the risk of bile duct injury (BDI)[3,4] and is commonly included by surgical societies in guidelines for safe LC.[5–7]

Despite this consensus, BDI in LC occurs more frequently than the 0.1% to 0.2% rate reported in the open cholecystectomy era, with 2300 to 3000 cases reported annually in the United States alone.[3,8] Research prompted by these numbers reveals that CVS is by no means a universally adopted standard.[9] Moreover, studies show dramatic incongruencies between the adoption rates reported by surgeons and those objectively documented in intraoperative videos,[10] with the latter as low as 10.8%.[11] Though several factors likely contribute to low CVS rates, it does appear that education and training can increase the adoption of the technique.[12]

These findings indicate that widespread implementation of CVS specifically and of validated surgical best practices, in general, require accurate, continuous documentation, and analysis of adoption trends over time. Today, developments in artificial intelligence–augmented video analysis render this task feasible. Computer vision algorithms have been used successfully to identify CVS in existing video data sets.[13,14] They have yet to be applied, however, in real-world clinical settings.

By integrating such algorithms with advanced electronic data collection, surgical intelligence platforms enable routine, automated capture, storage, and analysis of surgical video and can connect resulting data with patient factors and postoperative outcomes.[15] We report on the first documented use of surgical intelligence to assess CVS adoption routinely and automatically over time in a real-world surgical department. Our primary aim was a quality initiative to increase practice adoption among surgeons and shed light on the relationship between adoption and operational efficiency over time.

METHODS

Study Design and Quality Initiative

The study was conducted at 2 general surgery departments of a medical center where an installed surgical

From the *Department of Surgery, McGill University, Montreal, QC, Canada; †Department of Research and Development, Theator Inc., Palo Alto, CA; ‡Department of General and Emergency Surgery, Polytechnic University of Marche, Ancona, Italy; and Department of General Surgery, Tel Aviv Sourasky Medical Center, Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.
✉gerald.fried@mcgill.ca.
T.W. and D.A. hold leadership roles, Y.M. and S.M. are employees, and M.O. and G.M.F. are advisors at Theator Inc. The remaining authors report no conflicts of interest.

DOI: 10.1097/SLA.0000000000006377

intelligence platform (Theator Inc.) is used routinely to capture, analyze, and securely store procedure videos. The study included videos of all LCs conducted between December 2022 and August 2023. It was approved by the medical center's Institutional Review Board.

The employed surgical intelligence platform incorporates validated algorithms that accurately break down surgical videos into procedure-specific steps and identify intraoperative actions and events.[16,17] The platform presents captured videos and collected intraoperative data through a dedicated user interface, accessible to all department surgeons and clinical leadership on their personal computers and cellular phones (Fig. 1). An integrated algorithm that blurs extracavitary frames[18] maintains patient and surgeon confidentiality, in line with recently published standards for surgical video deidentification.[19]

The study time frame was divided into three 3-month periods: (1) preintervention to surface the rate of adoption between December 2022 and February 2023, (2) a quality initiative intervention between March and May 2023, and (3) continuous assessment to evaluate postintervention change between June and August 2023. The initiative intervention was planned and executed by the department heads, who set a cross-departmental goal of 75% CVS adoption. They communicated this goal at 2 departmental meetings, at which all surgeons and trainees engaged in a tutorial on platform capabilities, including how to view their procedures, receive and provide feedback, and review their own and aggregated metrics.

Algorithmically Derived Measures

CVS Adoption

The computer vision algorithms integrated into the employed surgical intelligence platform identify CVS based on 3 criteria, each associated with 1 of the 3 primary components of the method: (1) the hepatocystic triangle is cleared of fat and fibrous tissue, (2) the lower one-third of the gallbladder is separated from the liver to expose the cystic plate, and (3) the cystic duct and the cystic artery are the only 2 structures visible entering the gallbladder (Fig. 2). The platform provides a binary designation of "adopted" or "not adopted" for each component and full CVS if a procedure includes all 3. In studies detailing their development several years ago,[13,14] CVS adoption algorithms reached over 80% accuracy. The accuracy of the platform has since increased. Regarding the data set used in the current study, accuracy in reference to the annotations of 2 experienced general surgeons was over 95%.

Additional Measures

Using previously validated algorithms,[15] the surgical intelligence platform also routinely analyzes and presents the following data, per procedure and aggregated by variables such as specific surgeon, department, and time frame: procedure duration from scope in to scope out; negative intraoperative events, including common hepatic duct injury, cystic duct rupture, cystic duct injury, gallstone spillage, pus spillage, bile leak, bowel injury, liver injury, hydrops fluid leak, notable hemorrhage onset, repair of parenchymal injury, and sludge spillage; and complexity rating based on the Parkland grading scale, a 5-tiered, intraoperative grading system to stratify gallbladder disease severity during LC based on anatomic and inflammatory changes, which is associated with the technical complexity of the procedure.[20,21]

Outcome Measures and Statistical Analysis

To reveal patterns throughout the study and compare the periods before and following the quality initiative intervention, we extracted the following variables from the platform: monthly and 3-month adoption percentages (corresponding to the 3 aforementioned periods) for full CVS and each component; procedure duration in minutes; number of negative events per procedure; and complexity ratings.

We used χ^2 tests to assess the changes in adoption percentages over time for full CVS and each of its 3 components, comparing the 3-month preintervention period to the 3-month postintervention period. To minimize the possibility that results could be explained by associations between CVS adoption and procedure complexity, we used χ^2 tests to assess whether full adoption percentages differed between complexity levels during the first and final study periods and in the entire sample of procedures.

We also used χ^2 tests to compare the percentage of procedures with at least 1 negative event among procedures in which CVS was adopted versus those in which it was not. As Shapiro-Wilk tests indicated that procedure duration was not normally distributed, we used nonparametric analysis of variance (Kruskal-Wallis) to compare the 3 preintervention months with the 3 postintervention months and to compare procedures in which CVS was adopted with those in which it was not.

RESULTS

CVS Adoption

The final data set included 279 LC videos conducted by 46 surgeons, including 102 in the first period, 86 in the second, and 91 in the third. Of the patients recorded, none had postoperative bile duct injuries and 1 required reoperation due to bleeding.

Overall, surgeons adopted full CVS in 154 (55.2%) of the procedures. Figure 3 shows monthly adoption percentages for full CVS and each of the 3 components, indicating a gradual increase in full CVS from 33.3% in December 2022 to 75.7% in August 2023, with the departmental goal of 75% reached within 6 months of starting the quality initiative. Using 3-month adoption percentages to evaluate this change, we found a significant rise from 39.2% in the 3 months before intervention to 69.2% in the 3 months following intervention ($\chi^2 = 17.4$, $P < 0.001$). The "two structures visibility" component similarly showed a significant rise from 42.2% in the 3 months before intervention to 74.7% in the final 3 months ($\chi^2 = 20.9$, $P < 0.001$). However, the 2 additional components did not mirror this pattern, with the "clear hepatocystic triangle" starting and remaining high (88.2%–92.3%; $\chi^2 = 0.899$, ns) and the "gallbladder separation" component going from 83.3% to 87.9% ($\chi^2 = 0.813$, ns).

CVS adoption percentage did not differ based on complexity in the 3 months before intervention ($\chi^2 = 1.47$, ns), in the 3 months following intervention ($\chi^2 = 2.70$, ns), or overall ($\chi^2 = 6.28$, ns).

Efficiency-related Factors: Duration and Negative Events

Of the procedures in which CVS was adopted, there was at least 1 negative intraoperative event in 62%,

Copyright © 2024 Wolters Kluwer Health, Inc. All rights reserved.

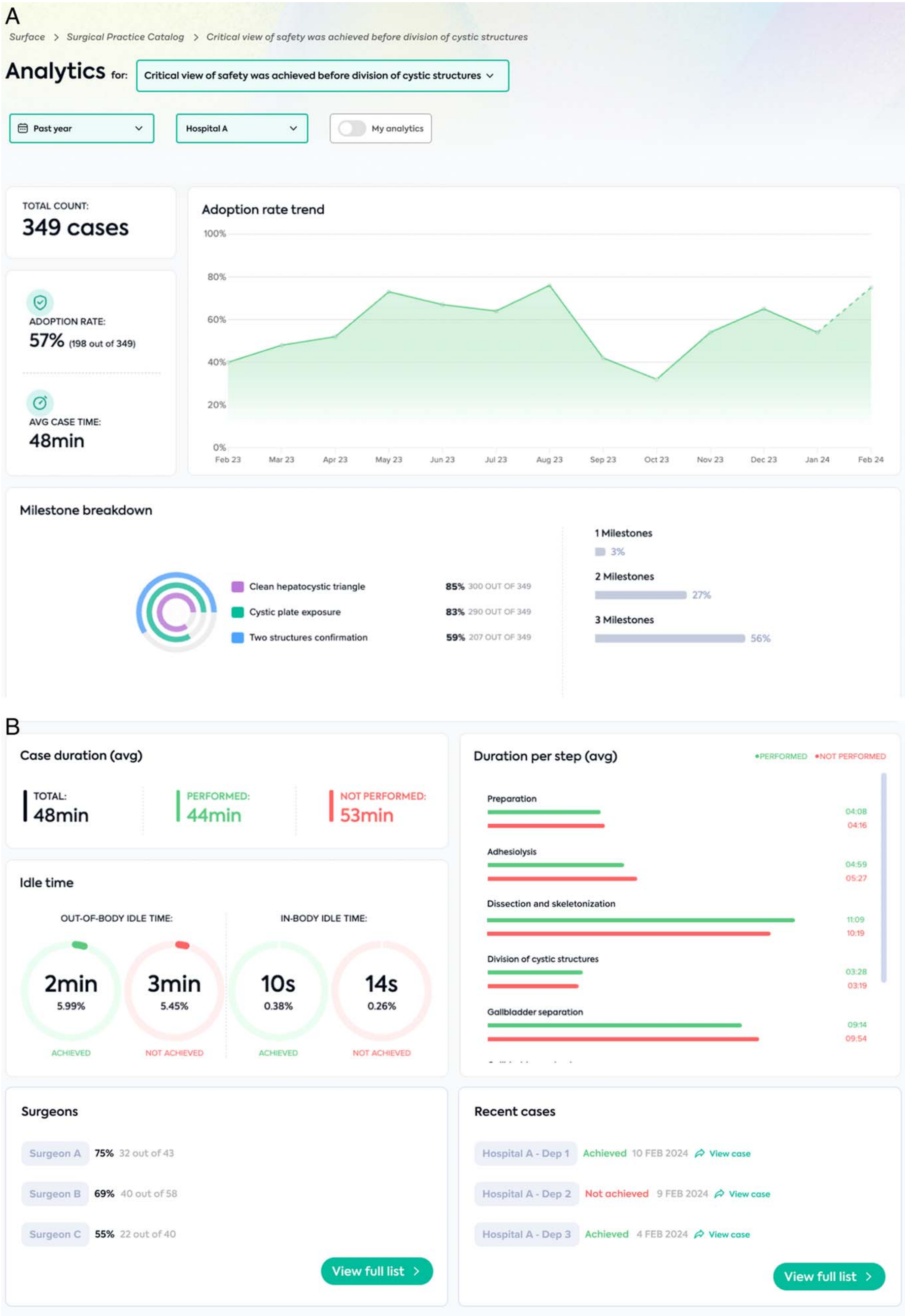

FIGURE 1. Screenshots (A, B) from the surfacing interface of the surgical intelligence platform employed in the current study, showing examples of the type of aggregated intraoperative data that were available to surgeons and clinical leadership participating in the quality initiative.

Copyright © 2024 Wolters Kluwer Health, Inc. All rights reserved.

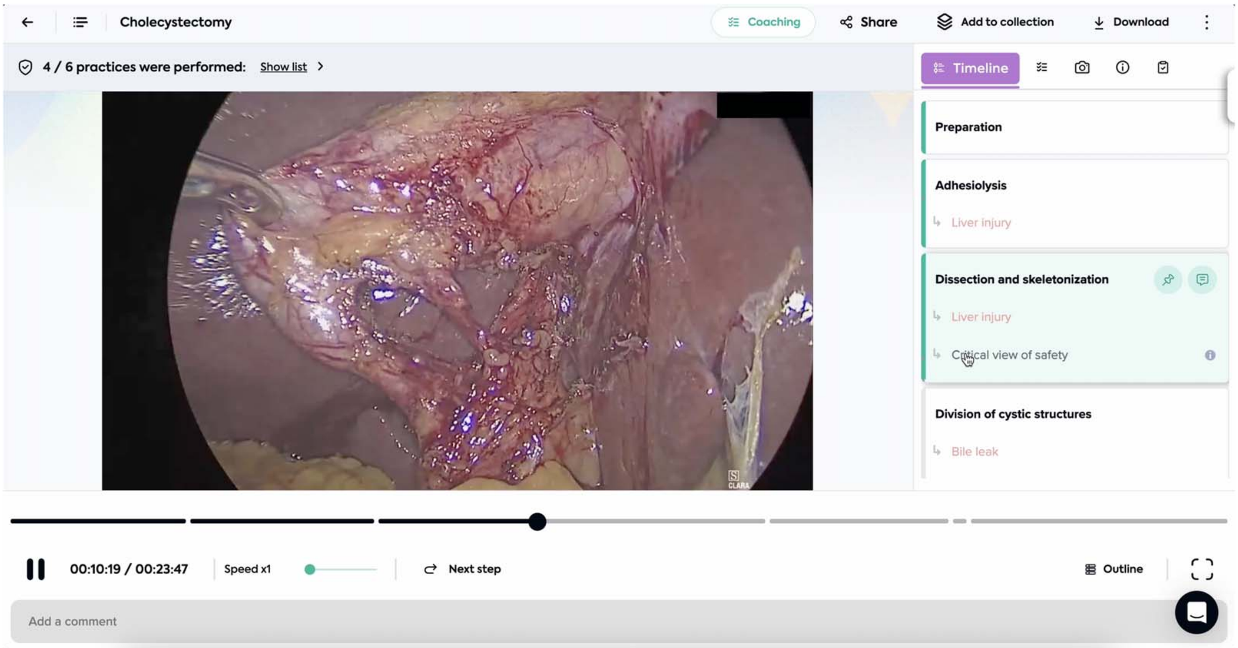

FIGURE 2. A screenshot from the video review interface of the surgical intelligence platform employed in the current study, showing the frame in which a critical view of safety was achieved.

compared with 76% in procedures without CVS adoption. This difference was significant ($\chi^2 = 6.51$, $P = 0.011$). The documented intraoperative events included no hepatic duct injuries and 27 hemorrhage onsets deemed notable by the surgical intelligence platform. Other intraoperative events predominantly included gallstone spillage, bile and fluid leaks, and sludge spillage. Table 1 shows the distribution of algorithmically identified intraoperative events in procedures with, compared to without, achievement of all 3 elements of the CVS.

Procedure duration was shorter when CVS was adopted (mean = 44 minutes, SD = 24.5 minutes) compared with when it was not adopted (mean = 57 minutes, SD = 41.1 minutes; $P = 0.007$). Procedure duration decreased significantly following intervention, from 54 minutes (SD = 38 minutes) in the 3 months before intervention to 44 minutes (SD = 26 minutes) in the final 3 months ($P = 0.033$). In Fig. 4, monthly CVS adoption percentages are plotted alongside average monthly procedure durations for the entire study period, to illustrate the relationship between these measures.

DISCUSSION

Through the specific case of CVS in LC, we examined the use of surgical intelligence to automatically assess surgical practice adoption and enable a real-world quality initiative. Two general surgery departments successfully employed a surgical intelligence platform to assess CVS adoption routinely, objectively, and accurately over time. Coupled with a focused quality intervention involving goal-setting, training, and routine assessment, this easily scaled process led to a stable increase that took the departments from just over 30% adoption to their goal of 75% within half a year. Procedure durations were lower following the intervention, and CVS adoption was associated with fewer negative events, suggesting that the implementation of validated surgical practices could be inherently linked to surgical efficiency.

In the context of CVS, these findings support a potential solution to a problem that has challenged the general surgery community in the 35 years since LC was introduced.[3] LC has notable advantages compared with open cholecystectomy but is also associated with higher rates of BDI.[22] Since LC is among the most commonly performed procedures worldwide,[23] even a low percentage of associated BDIs is a significant concern, because of its association with increased morbidity, prolonged hospitalization, high costs, and frequent litigation.[9]

CVS is associated with decreased BDI risk but video-based studies show that it is not universally adopted by surgeons.[10,11,14,24] Various reasons have been proposed for low CVS adoption, from lack of knowledge to personal preference for alternative methods.[25] While education and training might increase CVS rates,[10,12] this is unlikely to happen at the scale required to affect quality unless we can easily and accurately assess adoption over time in real-world clinical settings. The current work demonstrates that general surgery departments with access to routine video capture and artificial intelligence-based CVS identification can monitor, increase, and maintain adoption rates. As these technologies become more common, we can expect to see increased CVS adoption, which is subsequently reflected in lower BDI rates.

Data extracted from the surgical intelligence platform revealed that in the current sample, variance in CVS adoption was due predominantly to 1 of the 3 criteria comprising the method. While surgeons in the department cleared the hepatocystic triangle and separated the gallbladder from the liver in a high percentage of cases throughout the 3 study periods, the performance of the "two structures visibility" component started low. It rose gradually, essentially driving the improved rate observed for full CVS adoption. This level of detail is a key advantage of

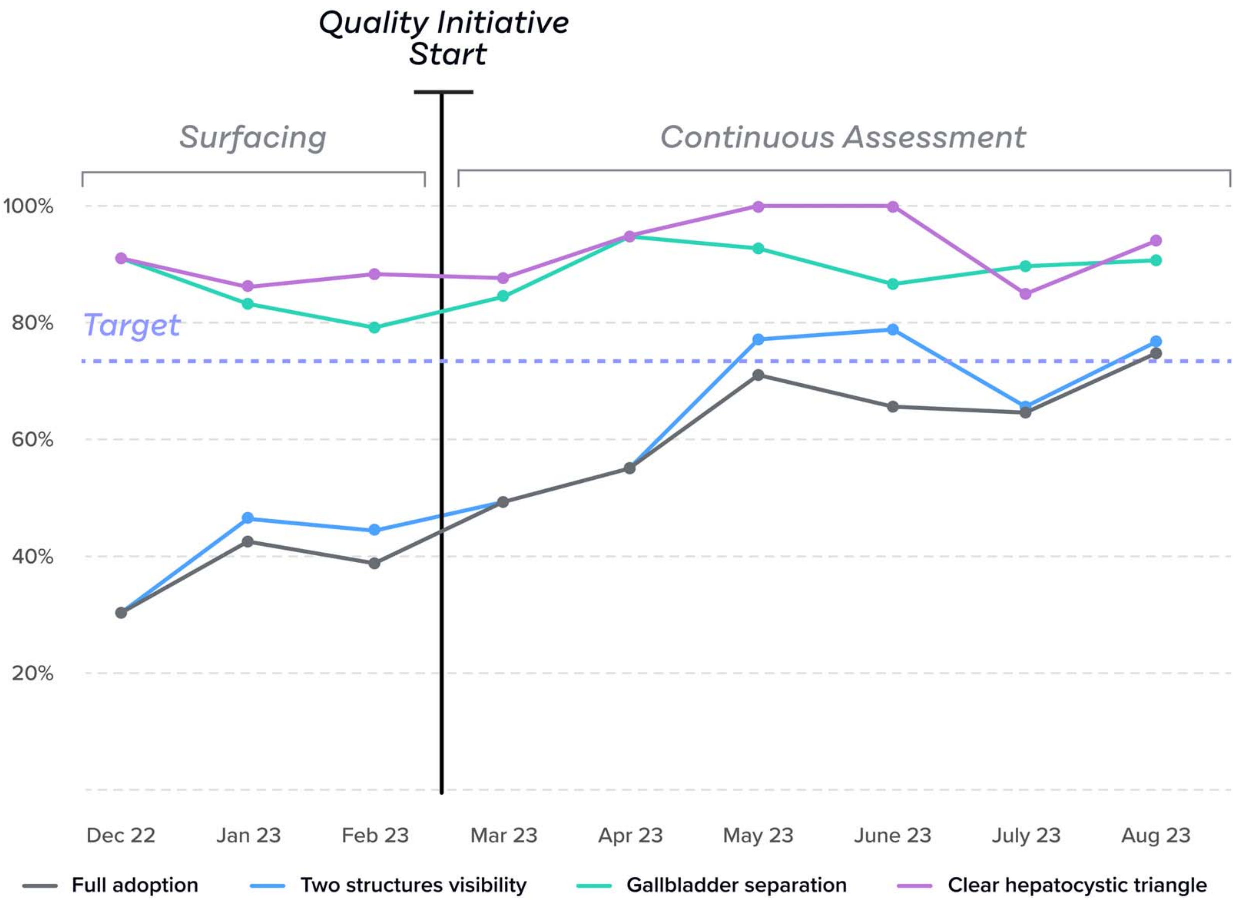

FIGURE 3. Monthly adoption percentages for a full critical view of safety (CVS) and each of the 3 CVS components for the entire study period.

surgical intelligence, noteworthy in that it would never be revealed without routine video capture and analysis since surgeons rarely reach such high resolution in their narrative operative reports. Details like this are critical if artificial intelligence is to provide actionable insights into surgery and tailor quality initiatives to address the true sources of surgical variability.

Additional insights stem from the relationships revealed between CVS adoption and other algorithmically derived intraoperative measures, another core capability of surgical intelligence platforms. Both procedure duration and negative events are tied to surgical efficiency through the costs of operating room time and operative complications, respectively. The idea that standardization leads to better efficiency, for example, by improving postoperative outcomes and decreasing surgical training times, is not new.[1] In the current study, lower procedure durations following the quality initiative supported the hypothesis that adopting safety views and other validated best practices directly facilitates the surgeon's ability to work efficiently.

Given the complexities of surgical workflows and the wide range of clinical expertise found among different surgeons in a department, we also considered the possibility that other variables could explain the relationships between the rate of CVS adoption and the duration and event measures. We did not find evidence of influence by the complexity measure and, as elaborated in the limitations paragraph below, were unable to reliably examine the effect of individual surgeon expertise. Further exploration through studies controlling these and possibly additional variables would be required to fully understand the determinants and effects of CVS adoption both within the context of surgical efficiency and outside it. In the meantime, the current findings indicate that beyond enhancing safety, CVS adoption and procedure standardization do not come at the expense of efficiency.

Beyond the aims specifically related to CVS, we sought to demonstrate the role of surgical intelligence in surgical quality improvement and to highlight the capabilities made possible by this emerging technology. The understanding that software-based automation is required to unlock the full potential of video analysis for surgical quality improvement has resulted in an outpouring of research on computer vision algorithms that accurately recognize intraoperative actions and events in a range of procedure types.[26] To impact real-world surgical quality and efficiency, however, algorithms must be packaged into a usable platform and integrated into daily surgeon workflows. The novelty and strength of surgical intelligence lies in incorporating computer vision capabilities into the same platforms responsible for routine video capture and storage so that all recorded procedures are automatically deidentified, structured, and analyzed.

The benefits of this technology in identifying evidence-based best practices and ensuring their implementation in surgical departments were well-illustrated in the current

Copyright © 2024 Wolters Kluwer Health, Inc. All rights reserved.

TABLE 1. Distribution of intraoperative events in procedures with versus without critical view of safety

	Critical view of safety	
Intraoperative event	**Adopted**	**Not adopted**
Common hepatic duct injury	0	0
Esophagus injury	0	0
Stomach injury	0	0
Cystic duct rupture	0	2
Cystic duct injury	0	1
Gallstone spillage	21	16
Pus spillage	1	2
Chyle leak	0	0
Bile leak	72	66
Stomach content spillage	0	0
Ureter injury	0	0
Spleen injury	0	0
Bowel injury	0	0
Pancreas injury	0	0
Liver injury	5	8
Hydrops fluid leak	5	15
Notable hemorrhage onset	13	14
Bowel content spillage	0	0
Bladder injury	0	0
Parenchymal injury repair	5	8
Sludge spillage	35	41
Total	157	173

Table entries represent the number of procedures in which the noted event was identified by the surgical intelligence platform at least once.

study. Automatic, routine video capture and analysis require minimal resources, making the process highly scalable and feasible. Importantly, routine capture prevents the biases often created by patient selection in surgical outcomes research,[27] while the use of algorithms with objective criteria for determining practice adoption further minimizes human bias and error.[26] Another key advantage over traditional research methodologies is continuity: while looking at practice adoption over a limited time frame can provide important insights, it cannot capture the dynamic nature of intraoperative activity within a functioning surgical department. When a surgical intelligence platform is installed, it acts as a sort of ongoing quality intervention by continually showing surgeons and clinical leaders where they stand and reminding them where they should be. Finally, as noted above, surgical intelligence enables the examination of intraoperative data at a greater level of detail than previously possible, revealing the actions most likely to improve practice adoption and subsequent quality and efficiency.

The findings of this study should be considered in the context of its limitations. First, due to the nature of data collection, factors with potential effects on the study variables were not controlled. In some cases, this limited our ability to analyze their impact. For example, while the platform documented the name of the attending surgeon for every procedure, we did not analyze these data due to the unequal distribution of cases among the surgeons, and particularly the fact that most of them performed a limited number of surgeries. While there did not appear to be an

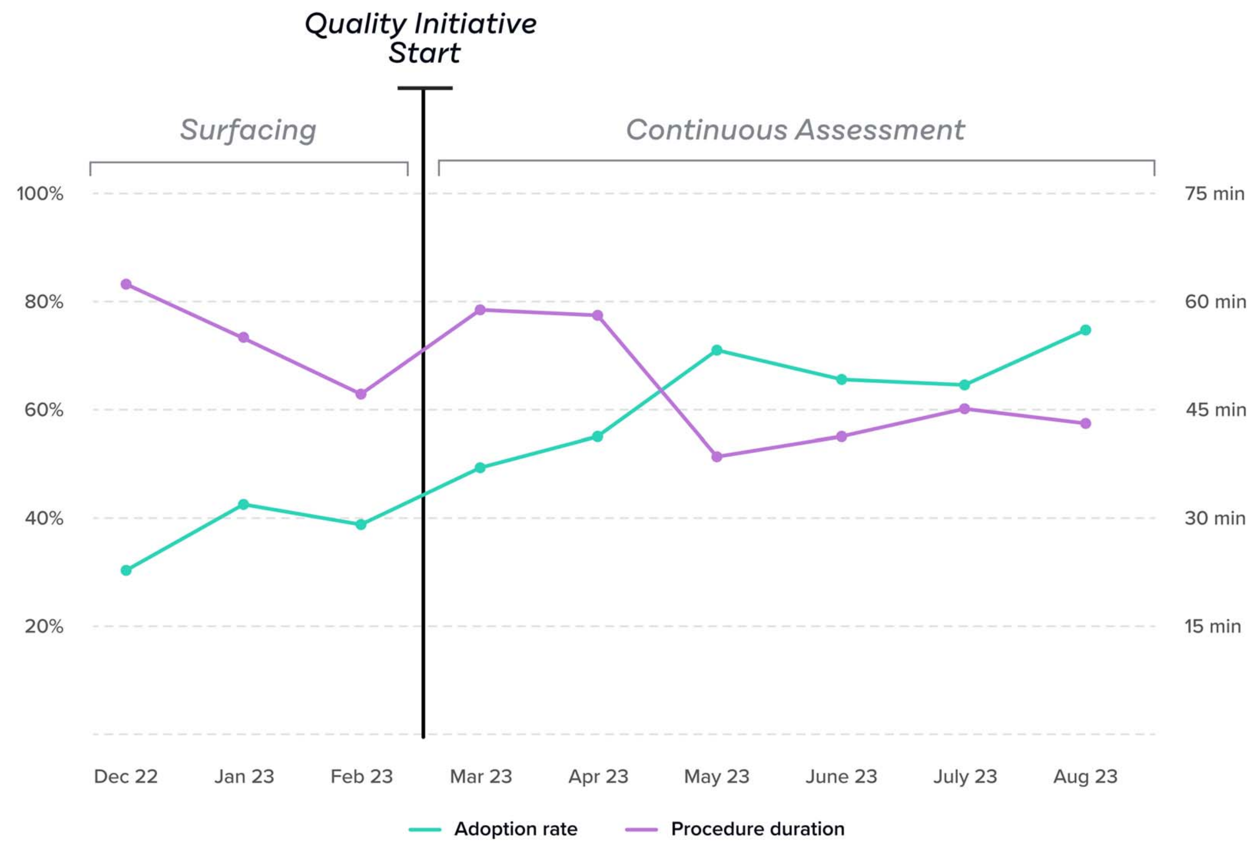

FIGURE 4. Monthly critical view of safety adoption percentages plotted alongside average monthly procedure durations for the entire study period.

association between adoption percentage and average procedure duration among surgeons who conducted at least 10 LCs during the study period, suggesting that this was an unlikely alternative explanation for decreased durations in procedures where CVS was adopted, we cannot rule out a mediating role for this factor. Also notable in the context of the surgeon variable is that residents participated in many of the procedures, and we did not determine which parts of the surgery they performed. While attending surgeons were likely to guide decisions about practice adoption, efficiency and events would likely be associated with the level of experience, which could have affected our results. The distribution of procedures between the 5 complexity ratings was also unequal, with the largest number of procedures receiving the middle rating of 3. In this case, we ran analyses to rule out the possibility that complex procedures were less likely to include CVS adoption and more likely to take longer. However, a more uniform distribution would lend greater reliability to this finding. Another limitation involves the accuracy of AI systems. While computer vision algorithms are sufficiently accurate to drive powerful feedback and tracking abilities in surgery and are expected to improve as routine video capture expands data sets, they are not 100%. Surgical intelligence platforms mitigate this limitation by connecting algorithm results to easily accessible raw video footage, enabling manual validation per video as well as broader quality assurance processes. Still, algorithm accuracy should be considered in drawing conclusions based on AI. Finally, we note that we were unable to assess the relationship between CVS adoption and BDI, as our relatively small sample did not include this complication. While prior work allows us to assume that fewer BDIs would occur when CVS was adopted,[28,29] the full capacity of surgical intelligence in quality improvement requires the illustration of ties between intraoperative practices and postoperative outcomes, a key component in identifying new practices and evaluating the effectiveness of quality interventions.

CONCLUSIONS

Surgical intelligence facilitated a quality initiative aimed at increasing CVS adoption in LCs performed at 2 surgical departments of a medical center. The initiative led to a steady increase in full CVS adoption, with the goal reaching 6 months after the initiative was launched. Routine, ongoing assessment also revealed that low initial adoption was related to a single CVS component and that increased adoption was associated with improved OR efficiency. These results are the first to demonstrate how real-world use of surgical intelligence to routinely and automatically assess the adoption of an intraoperative safety measure can uncover new insights, modify surgeon behavior, and enable focused initiatives to implement best practices for improving the quality and uniformity of surgical care.

REFERENCES

1. Hsueh WD, Gibber MJ. Closer to the mean: an argument for decreased variance in surgery. *Otolaryngol Head Neck Surg*. 2015;153:903–904.
2. Strasberg SM, Hertl M, Soper NJ. An analysis of the problem of biliary injury during laparoscopic cholecystectomy. *J Am Coll Surg*. 1995;180:101–125.
3. Brunt LM, Deziel DJ, Telem DA, et al. Safe cholecystectomy multi-society practice guideline and state of the art consensus conference on prevention of bile duct injury during cholecystectomy. *Ann Surg*. 2020;272:3–23.
4. Strasberg SM, Brunt LM. Rationale and use of the critical view of safety in laparoscopic cholecystectomy. *J Am Coll Surg*. 2010;211:132–138.
5. de'Angelis N, Catena F, Memeo R, et al. 2020 WSES guidelines for the detection and management of bile duct injury during cholecystectomy. *World J Emerg Surg*. 2021;16:30.
6. Pucher PH, Brunt LM, Fanelli RD, et al. SAGES expert Delphi consensus: critical factors for safe surgical practice in laparoscopic cholecystectomy. *Surg Endosc*. 2015;29:3074–3085.
7. Wakabayashi G, Iwashita Y, Hibi T, et al. Tokyo guidelines 2018: surgical management of acute cholecystitis: safe steps in laparoscopic cholecystectomy for acute cholecystitis (with videos). *J Hepatobiliary Pancreat Sci*. 2018; 25:73–86.
8. Pucher PH, Brunt LM, Davies N, et al. Outcome trends and safety measures after 30 years of laparoscopic cholecystectomy: a systematic review and pooled data analysis. *Surg Endosc*. 2018;32:2175–2183.
9. Manatakis DK, Antonopoulou MI, Tasis N, et al. Critical view of safety in laparoscopic cholecystectomy: a systematic review of current evidence and future perspectives. *World J Surg*. 2023; 47:640–648.
10. Stefanidis D, Chintalapudi N, Anderson-Montoya B, et al. How often do surgeons obtain the critical view of safety during laparoscopic cholecystectomy? *Surg Endosc*. 2017;31:142–146.
11. Nijssen MA, Schreinemakers JM, Meyer Z, et al. Complications after laparoscopic cholecystectomy: a video evaluation study of whether the critical view of safety was reached. *World J Surg*. 2015;39:1798–1803.
12. Weis J, Brunt LM, Madani A, et al. SAGES safe cholecystectomy modules improve practicing surgeons' judgment: results of a randomized, controlled trial. *Surg Endosc*. 2023;37: 862–870.
13. Colbeci Y, Zohar M, Neimark D, et al. A multi instance learning approach for critical view of safety detection in laparoscopic cholecystectomy. *Proc 7th Mach Learn Healthcare Conf, PMLR*. 2022;182:409–424.
14. Korndorffer JR Jr, Hawn MT, Spain DA, et al. Situating artificial intelligence in surgery: a focus on disease severity. *Ann Surg*. 2020;272:523–528.
15. Bar O, Neimark D, Zohar M, et al. Impact of data on generalization of AI for surgical intelligence applications. *Sci Rep*. 2020;10:22208.
16. Khanna A, Antolin A, Bar O, et al. Automated identification of key steps in robotic-assisted radical prostatectomy using artificial intelligence. *J Urol*. 2024;211:575–584.
17. Ortenzi M, Rapoport Ferman J, Antolin A, et al. A novel high accuracy model for automatic surgical workflow recognition using artificial intelligence in laparoscopic totally extraperitoneal inguinal hernia repair (TEP). *Surg Endosc*. 2023;37: 8818–8828.
18. Zohar M, Bar O, Neimark D, et al. Accurate detection of out of body segments in surgical video using semi-supervised learning. *Proc Third Conf Med Imaging Deep Learn, PMLR*. 2020;121: 923–936.
19. Tollefson MK, Ross CJ. Defining the standard for surgical video deidentification. *JAMA Surg*. 2024;159:104–105.
20. Liu YQ, Wang C, Cai X, et al. Can the Parkland Grading Scale predict the difficulty of laparoscopic cholecystectomy? A new approach to validation. *BMC Surg*. 2023;23:142.
21. Madni TD, Nakonezny PA, Barrios E, et al. Prospective validation of the Parkland Grading Scale for cholecystitis. *Am J Surg*. 2019;217:90–97.
22. The SAGES Safe Cholecystectomy Program. Strategies for minimizing bile duct injuries: adopting a universal culture of safety in cholecystectomy. Accessed March 2024. https://www.sages.org/safe-cholecystectomy-program/
23. Ríos MS, Molina-Rodriguez MA, Londoño D, et al. Cholec80-CVS: an open dataset with an evaluation of Strasberg's critical view of safety for AI. *Sci Data*. 2023;10:194.

24. Daly SC, Deziel DJ, Li X, et al. Current practices in biliary surgery: do we practice what we teach? *Surg Endosc*. 2016;30: 3345–3350.
25. van De Graaf FW, van den Bos J, Stassen LPS, et al. Lacunar implementation of the critical view of safety technique for laparoscopic cholecystectomy: results of a nationwide survey. *Surgery*. 2018;164:31–39.
26. Mascagni P, Alapatt D, Sestini L, et al. Computer vision in surgery: from potential to clinical value. *NPJ Digit Med*. 2022; 5:163.
27. Paradis C. Bias in surgical research. *Ann Surg*. 2008;248: 180–188.
28. Avgerinos C, Kelgiorgi D, Touloumis Z, et al. One thousand laparoscopic cholecystectomies in a single surgical unit using the "critical view of safety" technique. *J Gastrointest Surg*. 2009;13:498–503.
29. McKinley SK, Brunt LM, Schwaitzberg SD. Prevention of bile duct injury: the case for incorporating educational theories of expertise. *Surg Endosc*. 2014;28:3385–3391.

DISCUSSANT

Dr Michael J. Zinner (Miami, FL)

The discussion of this paper will be opened by Dr Carla Pugh.

Dr Carla Pugh (Stanford, CA)

Thank you so much, Dr Fried, for your excellent presentation. Adoption of clinical practice guidelines remains difficult, and we are all fully aware of this persistent problem. For surgery, one of the most significant barriers that plague the adoption of surgical technology is the lack of a plan for thoughtful integration into an already complex surgical workflow. From a process perspective, the value proposition of the work you have presented was facilitated by the department chairs. They set a goal of 75% compliance. This goal was measurable; there was access to training via a highly accessible AI tool, and there was routine assessment, including department-wide adoption and review of progress toward the goal. The department chairs, faculty, residents, and other participants are to be commended for taking a chance on a new implementation that could have been viewed as high risk due to time constraints and could have been derailed for fear of low return on investment. My first question relates to your main conclusions regarding efficiency. In your paper, you state your results suggest that implementing a surgical practice guideline could be inherently linked to efficiency, and I have to say I think that this conclusion might miss the mark. It is pertinent to explore the possibility that what you actually achieved was a department-wide consensus on what the critical view of safety actually is. Until now, our education regarding the critical view of safety and how to achieve it has been based on textbook drawings, images, and a few videos, and it is rare to have access to a large number of highly focused video highlight reels displaying the variety of ways in which surgeons achieve the critical view of safety. In essence, you achieved a department-wide focused learning opportunity and facilitated a team consensus on the operative goals. Please share with us your thoughts on this, is it really inherent efficiency alone or could something else be at play?

My second question is that in your paper you noted that those who adopted the critical view of safety had lower negative events compared to those who did not. However, your list of negative events spans the gamut of inconsequential bile duct leaks, which are extremely common in inflamed gallbladders, to bowel injuries which is an unexpected, egregious error for a lap chole and is highly unlikely to be related to achieving the critical view of safety. Can you comment on the distribution of highly consequential events across the 2 groups and the relationship to achieving the critical view of safety? Lastly, there is some discrepancy in the literature regarding the actual definition of the Parkland scale. Some groups include inflammation as well as anatomical complexity, and others only use inflammation. How was your AI model trained, and what was the focus? Thank you so much for this exciting contribution showing what can be accomplished with department-level consensus, goal-setting and regular, focused review of clinical practice guidelines using AI-assisted learning. Thank you.

Response from Gerald M. Fried

Thank you, Dr Pugh, for your insightful comments. I believe that successful implementation in this particular study was a result of leadership that was strong and highly respected and a safety initiative that the whole department could get behind. They believed not only in the value of the critical view of safety but also trusted the data. They also appreciated the fact that this did not add additional burden on them. So, all these elements came into play in the adoption of this best practice. You asked next about the role of computer vision technology in education. I think this is a powerful opportunity and certainly, after hearing the discussion at the symposium before lunch, I see a lot of opportunities for it to be able to be used for video-based assessment, feedback, and coaching. This particular department had this technology in place for several years before the current study and it was already quite well integrated into their educational programs. So, perhaps it was used in a little bit more directed way with respect to the critical view of the safety of this case, but otherwise, I don't think the additional value of the technology explained the changes that we saw in efficiency or other types of outcomes. We evaluated the efficiency and rate of adverse events to ensure that adding the mandate of full dissection of the critical view of safety prior to the division of the cystic duct and artery did not lead to a decrease in efficiency or technical errors. We hypothesized that the added dissection might also cause a higher risk of gallbladder perforation or cystic duct injury. Adopting the critical view of safety made the operation more efficient, possibly by standardizing the process so that everybody in the team became familiar with a routine way of doing things, or possibly because there was a lower rate of "events" that took time to deal with. This would require more analysis to truly understand the relationship between the critical view of safety and efficiency. With respect to the events that occurred, you are right that they ran a gamut from more minor ones to more major ones, and certainly, no one is equating making a hole and the gallbladder with making a hole in the bowel, but the nice thing is that the technology will track all of these and it will give us data on this. I don't think that making a hole in the gallbladder and spilling bile, stones, or sludge is an inconsequential event. We certainly know of cases where this has led to very severe complications even years down the road, so automatically capturing these events is helpful. Your final question was about case complexity. The software evaluated case complexity based on the visual evaluation of the gallbladder at the time of initial placement of the scope. It assessed

Copyright © 2024 Wolters Kluwer Health, Inc. All rights reserved.

adhesions and inflammatory changes as cited in the article in 2018 referenced in the paper. So, thank you.

Dr Abe Fingerhut (Poissy, France)

Abe Fingerhut from France. Thank you, Gerry, that was an excellent, excellent talk and I thank you also for sharing the manuscript with me. I have 2 questions. You alluded to the fact that there are other applications to this technology; I would be interested to know which application you are thinking about. My second question is, what is the proportion of the actual learning process due to this technology and not a simple Hawthorne effect, as the participants were not blinded?

Response from Gerald M. Fried

Let me try to address your second point first. I certainly believe that the Hawthorne effect is a very powerful one, but usually, it applies to a period of study. When you are doing a research study, it is over a finite period, and people are conscious that they are being measured. The nice thing about using this type of technology is that it is relatively permanent, and the data collection is ongoing. In this case, even though the study I reported was over 9 months, this technology remains in place in the department. Surgeons will get monthly reports on their adoption of the critical view of safety, and if the rate of adoptions starts to fall off, they will be alerted, and it gives them a prompt to improve their performance. Some of the other exciting applications of this technology relate to surgical education. It is a great way to use video technology for coaching, you could add elements of assessment and feedback into the video directly. This allows you to send a link that will take you directly to that part of the video where a faculty member can assess performance or embed feedback directly into the video file. So, this is really powerful. We need to look at other best practices, or putative best practices, and look at opportunities to relate short and long-term outcomes to these best practices, or to surface best practices we have never really thought about.

Dr Michael Brunt (St. Louis, MO)

Michael Brunt, St. Louis. Dr Fried and colleagues, congratulations. This is a really important work, and I think it is just the beginning of seeing more artificial surgical intelligence technologies coming into play in the operative environment. If there is one strategy you could adopt to enhance safety around cholecystectomy and reduced bile duct injuries, it is understanding and applying the critical view of safety in every case; so we have been talking about this for years and years, and yet as your data show, only 33% of surgeons were actually using it when you started this study. So, one of my questions is, did you actually look at the operative notes, and are you communicating to surgeons about how they are documenting in the operating notes since that traditionally has not been a very good way of documenting because surgeons say they are achieving the critical view if you look at the objective data oftentimes that is not the case. I presume you provided feedback continuously over the course of the study, but do you see a role for this being utilized real-time going forward potentially in the operating room environment, and something comes up and says, yes, you have the critical view or not and might lead to an opportunity for a real-time impact. My final question is, with 276 gallbladders, one would predict that there are a few that are going to be really difficult to get to the CVS, and were there any cases in which surgeons did not get to it for that reason or had to alter the approach for example to a subtotal cholecystectomy. Again, congratulations on the study.

Response from Gerald M. Fried

Thank you, Dr. Brunt, and I would like to acknowledge you and your colleagues at Washington University, St. Louis, for your long-standing work in reducing the risk of bile duct injuries and for contributing this best practice to the literature. In response to your questions, there were no partial cholecystectomies in the 297 cases in the study. We plan to review the operative reports in a future study to assess whether the accuracy of documentation in the OR record evolves when surgeons know that the procedure has been documented on video and analyzed by computer vision. I did miss one of your questions

Dr Michael Brunt (St. Louis, MO)

Real-time use in the OR.

Response from Gerald M. Fried

The data are acquired in real time. The annotations are not generally displayed on the monitor that the surgeon is working on, because they may be distracting. However, they can be so when an element of the critical view of safety has been achieved you will see it.

Dr Michael J. Zinner (Miami, FL)

Dr Lillemoe.

Dr Keith D. Lillemoe (Boston, MA)

Congratulations, Jerry and your group. My questions follow up with Mike's as in 2018, when the American College of Surgeons was in Boston, Mike ran a great multidisciplinary consensus conference on the prevention of bile duct injury. At the end of this 7-hour process, I stood up and introduced a young surgical faculty member and a resident at the MGH, Oz Meireles and Dan Hashimoto, who were working on an AI method of intraoperative prevention of bile duct injuries or other technical errors. They then had a paper at the ASA a few years ago looking at technical errors in sleeve gastrectomy. The goal is to have the equivalent of a dog collar "shock" when you are about ready to put a clip on the common bile duct that will stop you from making that error. Dan is now at Penn, and Oz is down at Duke. So, I am going to take the prerogative of the chairman of the Foundation to proudly share that we awarded Dan Hashimoto one of our next year's ASA Foundation Fellowships to continue this great work. I do think the real-time application will be the step that will really make the difference in preventing laparoscopic or robotic surgical errors in the future.

Response from Gerald M. Fried

Well, thank you, Dr Lillemoe. I would say there is a leap of faith that we need to take, and it would be wonderful if this organization took a leadership role in that regard. First and foremost, the surgical community needs to be willing to routinely capture videos of our procedures. All of us know that what we do in the operating room really matters, but we do not capture it. There is this opaque world there, and all that we have to go on is the operative report, which we all know is biased and contains errors of omission and commission. So, the first step is to routinely record. This provides the data that will allow these algorithms to improve their accuracy and their power. We as surgeons need to work with industry to help direct the products so that they provide us the information that we need to improve the quality of our work, and for that reason, I truly appreciate having this on the podium.

Dr William Richards (Mobile, AL)

Bill Richards from Mobile, Alabama. Great study, Dr Fried and colleagues, because this creates really substantial improvement in our ability to increase efficiency, safety, efficacy, and reduce time in OR. So, I have 2 quick questions—number one, you basically got to 75% adoption of the critical view of safety, but it plateaued there, so why did you not get to 100%? Was it because 25% of the surgeons were stuck at 30% or was it that all the surgeons were stuck at 75%. The second question is, will Big Brother use this for credentialing surgeons? Specifically, will Big Brother say that since you have a 25% critical view of safety, Dr Richards, you can no longer do lap chole?

Response from Gerald M. Fried

Dr. Richards, I am sure your rate of critical view of safety is greater than 25%, but if not, I am happy to work with you, and I am sure we can get it better. I don't think we will ever get to a 100% rate of critical view of safety. There are some cases where, in the surgeon's judgment, that is not the safest way to proceed. Dr Brunt may argue with me, but I don't think we will ever get to 100%. There are 46 surgeons involved in the study, so there are some surgeons who do a lot of lap choles and others who do these procedures relatively infrequently. It is a mixed bag. The most important thing is that it becomes part of the departmental culture. If there are individual outliers, there is an opportunity for the chief to look at their practices and ask if they should be doing lap choles, and if they do lap choles, they should be held to a high-quality standard. I think this is a work in progress. The data set is still relatively young, but the essence of what I would like to say is computer vision technology is here, and that we as surgeons should work to both generate the data that is going to be powerful by videoing our cases and helping to identify best practices that we want to be implemented and ensure that they actually are.

Dr Michael J. ZInner (Miami, FL)

Dr Hunt, final question.

Dr Kelly Hunt (Houston, TX)

Yes, I just had a question about the documentation of the operative records. So, are you planning to do that as synoptic operative reporting because we know that when it is a narrative report that frequently people leave out elements that are critical.

Response from Gerald M. Fried

Right, I think that is a very good point. The technology has the capability to automatically generate a synoptic operative report based on what you can see in the operating room and illustrate each step with an image. We are interested in the report that the surgeon creates as part of the EMR. This department does not use synoptic operative reporting, so we will have to manually go through the 300 cases to evaluate accuracy. We plan to then periodically review the data set over time to see if surgeons' documentation improves because they know their cases are being captured on video, allowing the operative report to be compared with video as the ground truth.

Dr Michael J. Zinner (Miami, FL)

Thank you, Dr Fried. Thank you for this presentation.

Copyright © 2024 Wolters Kluwer Health, Inc. All rights reserved.