fortune Journals

Research Article

JOURNAL OF SURGERY AND RESEARCH

ISSN: 2640-1002

Development and Validation of a Novel Intraoperative Scoring Tool Using Video Analytics to Standardise the Complexity of Gallbladder Surgery

H Naseem[1], A Tam[1], H Naseem[2*], M Hutton[1], M Clarke[1], I Finlay[1], J Clark[1]

Abstract

Background: Cholecystitis is a common disease that is often treated with laparoscopic cholecystectomies. Existing grading scales lack dynamism. Introducing a dynamic intraoperative scoring tool would help standardize training, utilisation and scheduling of operating rooms,that later could be incorporated into artificial intelligence.

Aim/Hypothesis: • To develop and validate a dynamic scoring tool that incorporates the procedure's difficulty and severity of cholecystitis in correlation to operation time.

• We hypothesise that a higher score correlates positively with increasing time, which could dynamically enhance operating room efficiency and turnover.

Method: A retrospective review of n = 123 LC videos of anonymised patients of Royal Cornwall Trust Hospital, the UK, between April 2019 - September 2021 was performed. It was to propose a scoring tool that analysed the correlation between cholecystitis severity, the difficulty of the operation and operative time. The system considers gall bladder characteristics, operative steps, anomalies, and additional procedures.

Results: In the study, 123 cholecystectomy procedures were analysed, n = 81(65.85%) met inclusion criteria and n=42 (34.15%) were excluded. The average total score per video was 8.63, and the completion time was around 65 minutes. Higher scores correlated positively with extended operative time (R=0.791). The initial Assessment Score positively correlated with the Initial Assessment Time (R=0.607). Total Score, Completion Time, and Initial Assessment Score had a strong positive correlation (R=0.944).

Conclusion: The proposed grading scale is a practical predictor of the complexity and the duration of the operation. As envisaged it should be helpful in surgical training and enhance theatre efficiency.

Introduction

Cholelithiasis is a common ailment globally, and especially in most developed countries, it predominates in women [1]. However, only 20% of those affected are symptomatically treated surgically with laparoscopic cholecystectomy (LC), which has become a gold standard since 1992 [1-3].

In the UK, LC rates range from 112 to 371 per 100,000 population. Cholelithiasis prevalence peaks in Norway (21.9%) and former East Germany (19.7%), contrasting with Italy's lowest rates (5.9%). Cholecystectomy rates vary globally: Belgium (189/100,000), France (133/100,000), Sweden (121/100,000), and Norway (62/100,000) [4]. In the USA, 10%-20% of adults

Affiliation:

[1]Department of Surgery Royal Cornwall Hospital Treliske, Truro, Cornwall, United Kingdom

[2]Altnagelvin Area Hospital, Londonderry, Northern Ireland, United Kingdom

***Corresponding author:**
Humair Naseem, Department of Surgery, Royal Cornwall Hospital Treliske, Truro, Cornwall, United Kingdom.

Citation: H Naseem, A Tam, H Naseem, M Hutton, M Clarke, I Finlay, J Clark. Development and Validation of a Novel Intraoperative Scoring Tool Using Video Analytics to Standardise the Complexity of Gallbladder Surgery. Journal of Surgery and Research. 8 (2025): 423-436.

Received: August 15, 2024
Accepted: August 23, 2024
Published: August 22, 2025

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

(approximately 20 million) have gallstone disease, and only 1-3% become symptomatic with associated complications, costing an estimated $5 billion [1]. Annually, the USA performs 300,000 to 600,000 cholecystectomies, while the UK conducts 70,000 [5,6]. These statistics underscore cholelithiasis's global impact and emphasize laparoscopic cholecystectomy as the gold standard, highlighting it is imperative in healthcare systems.

Theatre rooms are significant for the hospital finance systems, constituting approximately one-third of healthcare expenditure due to surgical care. However, inefficiencies in theatre utilisation contribute significantly to avoidable healthcare costs. According to a study by Daniel J Lee et al. [6], the average cost of operating a theatre room is reported to be around $30-$100 per minute. This underscores the critical need to maximize theatre utilization for enhancing surgical care. Furthermore, in 2007, unexpected surgical cancellations cost UK hospitals almost £88 million, rising to £400 million annually by 2018 [6,7]. Prolonged waiting times, often resulting from the aforementioned factors, have led to reported morbidity rates of 14% and mortality rates of 2% in patients waiting an average of 238 days [8]. These statistics highlight the urgency of addressing factors contributing to delayed surgical procedures, not only from a financial perspective but also for the well-being and outcomes of patients.

The heterogeneity within gallstone disease's spectrum and its resulting complications often lead to procedural complexities and challenging dissections. To standardize its management, numerous attempts have been made to classify this condition based on clinical, operative, pathological, and imaging criteria. However, existing classifications, such as Tokyo Guidelines 2018 [9], Nassar's intraoperative scale [10], Parkland grading [11], and AAST EGS grading scale [12], though comprehensive, lack dynamic intra-operative applicability and practical surgical or organizational utility. These classifications incorporate various factors but fail to adapt in real-time during surgery. This has led to a proliferation of classifications and guidelines, highlighting the need for more refined tools to manage acute cholecystitis effectively. Recent research has aimed at modifying existing grading tools like Tokyo Guidelines 2018, acknowledging their limitations and suggesting improvements for flawless management of acute cholecystitis [13].

Among these efforts, studies like Chole S [14] have incorporated operative time as a significant factor, while CAAD [15] has attempted to establish criteria for day-case scheduling. However, it's worth noting that the primary data collected for these studies had different objectives, indicating a gap in directly addressing the needs of these secondary studies. Therefore, there is a pressing need to refine and innovate upon existing classifications to create more effective tools.

The recent advancements in data capture and artificial intelligence (AI) practices have ushered in a new era of real-time surgical decision-making, providing invaluable support to clinical teams in identifying and neutralizing adverse intraoperative events. Leonardo et al. [16] showed in one of their studies that AI is revolutionising surgical practices, particularly in theatre management. Therefore, a significant improvement in the surgical workflow can be anticipated with the progression of AI and other technologies.

Medical training has undergone significant advancements and is currently undergoing fundamental changes worldwide. A notable example is the laparoscopic cholecystectomy, which serves as an index procedure for trainee surgeons with varying durations, ranging from 20 minutes to 3 hours [17]. The duration is contingent upon the complexity of gallstone disease and the surgeons' proficiency levels. Moreover, medical institutions are remodelling education and evaluation methodologies to address various impediments and challenges arising from patients, societal expectations, physicians, and students. Hashimoto et al. [18] propose the integration of AI in machine-based learning to classify, summarize, and segment recorded surgical videos into distinct procedural steps. Their findings reveal a remarkable 92% accuracy in comparisons between AI-generated segmentations and classifications and those performed by surgeons. This application of AI holds promise in reducing human error and enhancing efficiency, particularly in the validation of the algorithm for clinical use during intraoperative situations, where adverse incidents are a prevalent cause of surgical morbidity [19].

We are cognizant of the facts cited above, such as improved patient care and waiting-list time, surgical training, and theatre utilisation and efficiency. We aim to develop a tool that would encompass challenges addressed individually by previous classifications and a wholesome scoring system that not only serves an academic purpose but is also practical and dynamic. It would standardise the complexities of surgical procedures and disease severity by using video-based analytics, where the score is reflected in real time. Alongside the potential benefits of this in providing enhanced scheduling opportunities for the hospital. We envisage that in the future, it will be incorporated with or without some modifications in laparoscopic cameras for automation with the help of artificial intelligence, enhancing surgical efficiency and standardization.

Methodology

In this study, we conducted a comprehensive review of one hundred and twenty-three shared recorded videos of laparoscopic cholecystectomy operations performed on adult patients of Royal Cornwall Hospital NHS Trust, Treliske, Cornwall, UK, between April 2019 - September 2021. Of these, eighty-one laparoscopic cholecystectomy videos were deemed suitable for inclusion in our study. These videos were

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

anonymised and retrospectively analysed over eight months. A scoring tool was developed initially using steps and challenges of laparoscopic cholecystectomy as a procedure is performed and progresses. This tool evolved and reached its final iteration by the study's conclusion. Data was collected and used to analyse the correlation of procedure scoring and operative time considering gall bladder characteristics, Calots Triangle anatomy and dissection, essential operative steps, and additional procedures such as operative cholangiography (OTC) and intraoperative ultrasonography (IOUS) (Figure 1). Consent was sought from seven upper gastrointestinal consultant surgeons via email to access their recorded videos, with five consenting to share their recordings for the study.

Inclusion criteria for video selection were stringent, requiring fully recorded videos enabling accurate scoring tool calculations. Exclusion criteria encompassed incomplete videos or those lacking determinable gallbladder characteristics, ensuring the integrity and reliability of the scoring process. Consequently, forty-two videos were excluded from the study, resulting in a sample size of eighty-one.

Our calculated sample size, set at forty with a power of 0.903 using a large effect in the equation, was doubled to eighty-one to enhance the robustness of the data. This decision was made to elevate the study's reliability and statistical power.

The intraoperative predictive scoring model was validated by showing the same five LC videos, to three experienced upper GI / General surgeons. In addition, the surgeons were briefed on the study, provided with intraoperative reference images to guide them, and requested to apply the suggested scoring tool for this study. The validators were blinded to the surgeons and did not participate in the procedures in the reviewed videos.

Complexity assessment and score assignments

The proposed scoring tool aims to assess the complexity of Laparoscopic Cholecystectomy (LC) procedures by assigning scores based on operative, clinical, pathological, and intraoperative findings (Figure 1). The LC procedure is systematically divided into twenty-six distinct steps, and tasks, assembled into seven categories, with corresponding scores reflecting the increasing difficulty and severity of the disease. This tool employs a composite of five grades, ranging from 1 to 5, to indicate potential degrees of disease severity.

Categories and Scoring Guidelines:

The scoring tool encompasses seven distinct categories incorporating various aspects, each contributing to the overall assessment:

- Categories A & B: Focus on gallbladder-related factors, such as appearance, size variations (e.g., shrunken, fibrosed, distended), wall thickness, and varying degrees of adhesions (Picture 1 & 2). Scores in these categories reflect disease severity and potential procedural complexity, allowing only one variable selection per category.

Humair-Clark LC Scoring System

Select One from categories A to E

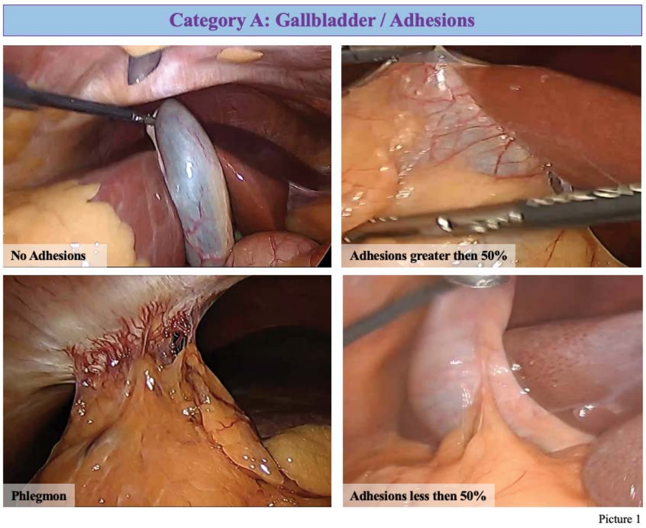
		Scoring
Category A: GB Adhesions		
	No Adhesions	1
	Adhesions <50%	2
	Adhesions >50%	3
	Phlegmon with / without bowel / hepatic adhesions	4
Category B: GB Size / Wall Thickness		
	Normal with thin walls	1
	Distended can be grasped	2
	Distended unable to grasp / Need decompression	3
	Shrunken / fibrosed / thick walled	4
	Necrotic / Gangrenous / Perforated	5
Category C: Calot's Triangle		
	Normal / Non-Fatty	1
	Fatty	2
	Fibrosed	3
Category D: Difficult Access		
	Pervious Surgeries / Hostile Abdomen	1
Category E: Calot's Triangle Dissection		
	Easy + CVS	1
	Moderately difficult dissection + CVS	2
	Very difficult +- with CVS / without CVS	3
	Impossible / Subtotal cholecystectomy / No CVS	4

Select what applies from categories F to G

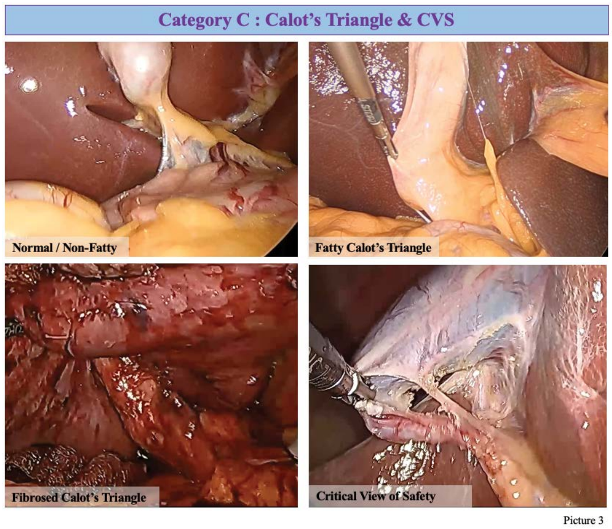
		Scoring
Category F: Anomaly		
	Accessary ducts	1
	RHD in Calot's triangle	2
Category G: Additional Procedures		
	Drain / Haemostatic Control	1
	OTC	1
	IOUS	1
	Extensive Adhesiolysis	2
	Lap Choledochoscopy +/- Stone Extraction	2
	Laparoscopic CBD Exploration	3
	Conversion to open	4
	Total Score	34

Figure 1: Humair-Clark LC Scoring System.

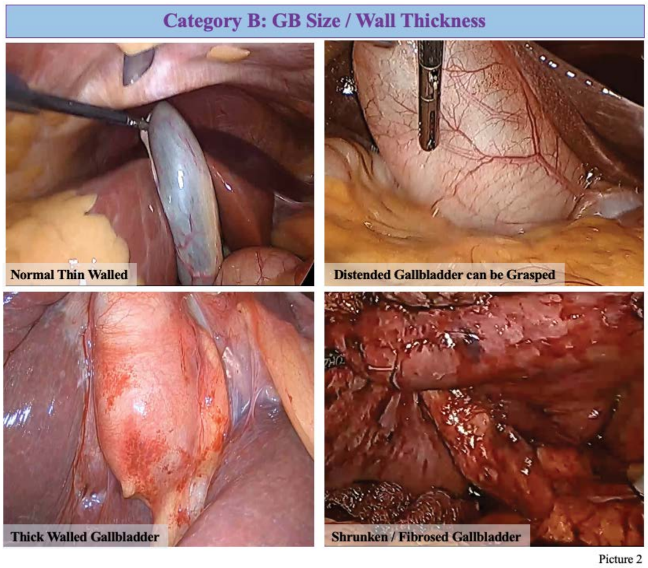
H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468



Picture 1:



Picture 3:



Picture 2:

- Categories C & E: Focus on Calot's triangle appearances and dissection difficulties. (Picture 3). Only one variable can be selected from each category.
- Category D: Assesses and addresses the accessibility challenges in patients with prior surgeries and hostile abdomens. This category highlights the potential challenges and adds complexities to the procedure during and after access. The variable is either selected or left alone.
- Category F & G: Account for anomalies and challenges encountered during the procedure, including identification of accessory and right hepatic ducts, utilisation of additional techniques like intraoperative ultrasound (IOUS), on-table cholangiogram (OTC), drainage procedures, and diverse haemostatic controls to manage excessive bleeding (e.g., haemostatic agents, cauterisation, suturing). These categories permit the selection of multiple variables in any combination.

Scoring parameters

- Scoring Range: A maximum score of seventeen is achievable in Categories A to E, while Categories F & G also allow a maximum score of seventeen, culminating in a total achievable score of 34.
- Difficulty Grading: Scores of 1 are considered easy, 2 to 3 moderate, and 4 to 5 difficult, reflecting increasing disease severity and procedural complexity.

Significance of scores:

- Categories A to E signify acute and chronic disease severity, along with the associated surgical challenges.
- Categories F & G capture anatomical variations and additional procedures, showcasing the evolution of the procedure as it progresses.

Clinical relevance

The scores derived from this tool provide a comprehensive assessment of acute and chronic disease severity, along with associated surgical challenges. Simultaneously, they reflect anatomical variations and additional procedures, offering valuable insights into the evolving nature of the LC procedure.

Data analysis

The total score grades and time taken for each recorded video were calculated using the scoring tool. By compiling and analysing these data points, a time value is assigned to the numerical number one, providing a quantitative measure of procedural efficiency.

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

Statistical analysis

We employed Pearson Correlation Coefficient analysis to assess the relationship between derived grades and procedure finishing times. Linear regression models were created for each hypothesis where Pearson's R-value of 1 indicates a strong positive relationship, -1 indicates a strong negative relationship, and an R-value of 0 shows no relationship between the two variables. Descriptive data were presented as mean, standard deviation and range.

To identify significant associations between variables and outcomes, Fisher's exact test was applied, with a p-value < 0.05 considered indicative of statistical significance. Additionally, Cook's Distance test was utilised to calculate the relative influence of each case in a sample of data on the results of a regression analysis. It is used to identify the influential outliers with such a large effect that they might distort the overall results of the regression model. In all models, both time and grades were continuous dependent variables. The interclass correlation coefficient (ICC) was used for inter-rate reliability. Statistical analysis was performed with SPSS v.25 statistical software, and graphs were refined using MS Excel v.16.50.

Results

During the study period from April 2019 to September 2021, a retrospective analysis was conducted on one hundred twenty-three (n=123) shared and pre-recorded laparoscopic cholecystectomy videos. Out of these, eighty-one videos (n = 81, 65.85%) met the predetermined inclusion criteria while forty-two (n = 42, 14.56%) were excluded from the study.

Total Score (TS) and Completion Time (CT)

The Total Score (TS) serves as the conclusive measure for each Laparoscopic Cholecystectomy (LC) in our proposed scoring tool. Completion Time (CT) is recorded as the duration from the camera's insertion point to the closure of the abdominal sheath for each video. To determine the time per score, we computed the ratio of the average completion time to the average total score, resulting in a value of 7:34 minutes for each score of 1, termed "time per score." The formula is expressed as follows:

"Time per score = Average CT / Average TS"

The mean Total Score was 8.630 (range 4 - 24), accompanied by an average Completion Time of 01:05:20, equivalent to approximately sixty-five minutes (range 00:08:29 - 03:40:45 minutes) as illustrated in figure 2. Notably, robust linear regression analysis between Total Score and the Completion Time, as depicted in graph-1, revealed a significant correlation (R= 0.791) and a regression R squared value (R^2) of 0.626. The remarkably low residual SS value of 539.25 further attests to the goodness-of-fit of our regression analysis, as shown in figure 2c. These findings highlight the association between Total Score and Completion Time in our study, emphasizing their correlation and impact on procedural outcomes in laparoscopic cholecystectomy.

2.a Descriptive Statistics

	TS	CT
Range	20	03:32:16
Minimum	4	00:08:29
Maximum	24	03:40:45
Mean	8.630	01:05:20
Standard Deviation	4.247	00:43:48

2.b Regression Statistics[a]

R[b]	R Square	Adjusted R Square	Std. Error of the Estimate	Sig.
0.791[c]	0.626	0.622	2.613	.000

2.c ANOVA[a]

	Sum of Squares	df	Mean Square	F	Sig.
Regression	903.638	1	903.638	132.383	.000[b]
Residual	539.25	79	6.826		
Total	1442.889	80			

a. Dependent Variable: TS
b. Pearson Correlation
c. Predictors: (Constant), CT

Figure 2: Total Score (TS) and Completion Time (CT)

Citation: H Naseem, A Tam, H Naseem, M Hutton, M Clarke, I Finlay, J Clark. Development and Validation of a Novel Intraoperative Scoring Tool Using Video Analytics to Standardise the Complexity of Gallbladder Surgery. Journal of Surgery and Research. 8 (2025): 423-436.

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

3.a Descriptive Statistics		
	IAS	IAT
Range	10	01:22:22
Minimum	3	00:02:38
Maximum	13	01:25:00
Mean	5.790	00:13:26
Standard Deviation	2.663	00:12:02

3.b Regression Statistics[a]				
R[b]	R Square	Adjusted R Square	Std. Error of the Estimate	Sig.
0.607[c]	0.368	0.36	2.130	.000

3.c ANOVA[a]					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	208.793	1	208.793	45.992	.000[b]
Residual	358.639	79	4.54		
Total	567.432	80			

a. Dependent Variable: IAS
b. Pearson Correlation
c. Predictors: (Constant), IAT

Figure 3: Initial Assessment Score (IAS) and Initial Assessment time (IAT)

4.a Descriptive Statistics		
	TS	IAS
Range	20	10
Minimum	4	3
Maximum	24	13
Mean	8.630	5.79
Standard Deviation	4.247	2.663

4.b Regression Statistics[a]				
R[b]	R Square	Adjusted R Square	Std. Error of the Estimate	Sig.
0.883[c]	0.779	0.776	2.009	.000

4.c ANOVA[a]					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	1124.236	1	1124.236	278.719	.000[b]
Residual	318.653	79	4.034		
Total	1442.889	80			

a. Dependent Variable: TS
b. Pearson Correlation
c. Predictors: (Constant), IAS

Figure 4: Total Score (TS) and Initial Assessment Score (IAS).

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

Initial assessment score and initial assessment time

The preliminary evaluation occurs during the opening of the peritoneal envelope surrounding the gallbladder. At this juncture, the Initial Assessment Score (IAS) is determined based on categories A, B, C, and D from the scoring tool, while the time taken to complete this assessment is referred to as the Initial Assessment Time (IAT). The mean IAS is calculated at 5.790 (range 3-13), and the corresponding IAT averages at 00:13:26 minutes (range of 00:02:38 to 1:25:00 minutes). This data underscores the efficiency and variability in the initial assessment process.

Furthermore, a positive regression correlation of R = 0.607 between the Initial Assessment Score (IAS) and Initial Assessment Time (IAT) has been established. The goodness-of-fit is consistently upheld, and the Significance F, with a p-value of 0.000, is lower than 5%, affirming the overall significance of the proposed model. These findings, as illustrated in figures 3 and 3c, not only highlight the relationship between IAS and IAT but also emphasize the statistical robustness and practical relevance of our proposed scoring system.

Total score and initial assessment score

The data analysis reveals a compelling direct relationship between the initial assessment score (IAS) and the total score (TS), providing valuable predictive insight into the total operative time, right from the outset or during the early stages of the operation. This relationship is represented by a robust positive linear regression, showcasing a strong correlation with an R-value of 0.883 and an R-squared value (R^2) of 0.779 between IAS and TS. The statistical significance, indicated by the exceedingly low p-value of 0.000 (Significance F), well below the 5% threshold, underscores the overall significance of the proposed model. These findings are visually elucidated in figures 4a&4c, reaffirmed by the clear depiction within the Linear Regression Plot as shown in graph-2.

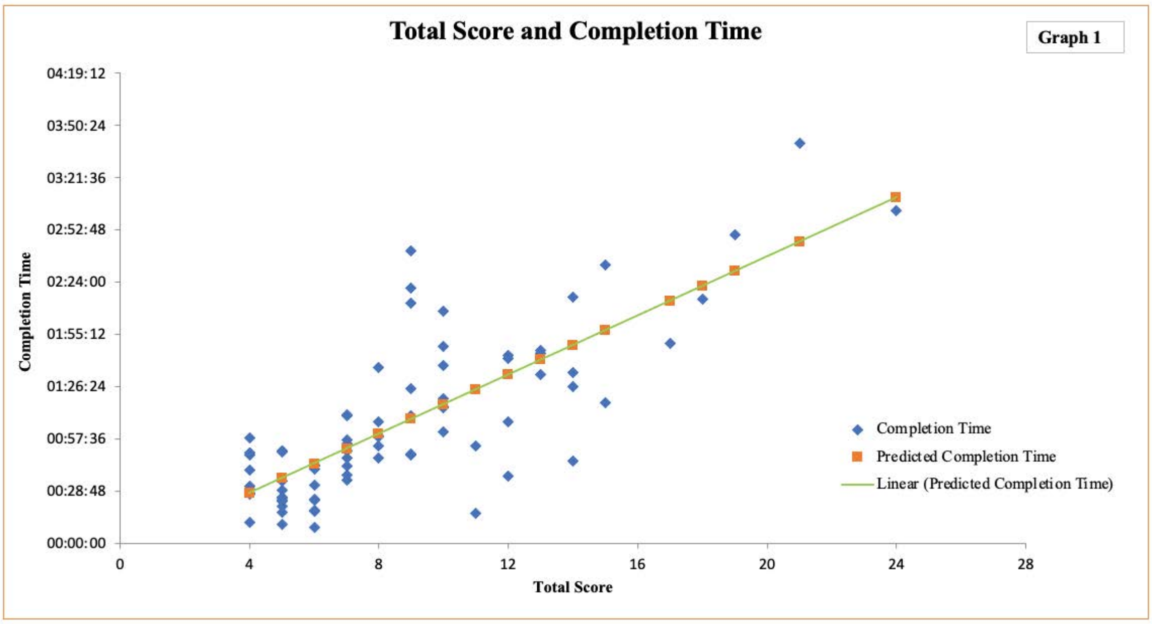
Total score with completion time and initial assessment score

We explored the interplay of three key parameters and identified a robust positive relationship among them. Specifically, we established a strong correlation of R = 0.944 between the initial assessment score, total score, and the completion time, as illustrated in figure 5. The statistical analysis further supports the significance of this relationship, with a p-value of 0.000. The high correlation signifies a substantial and reliable association among these variables.

Total score with completion time and surgical skills

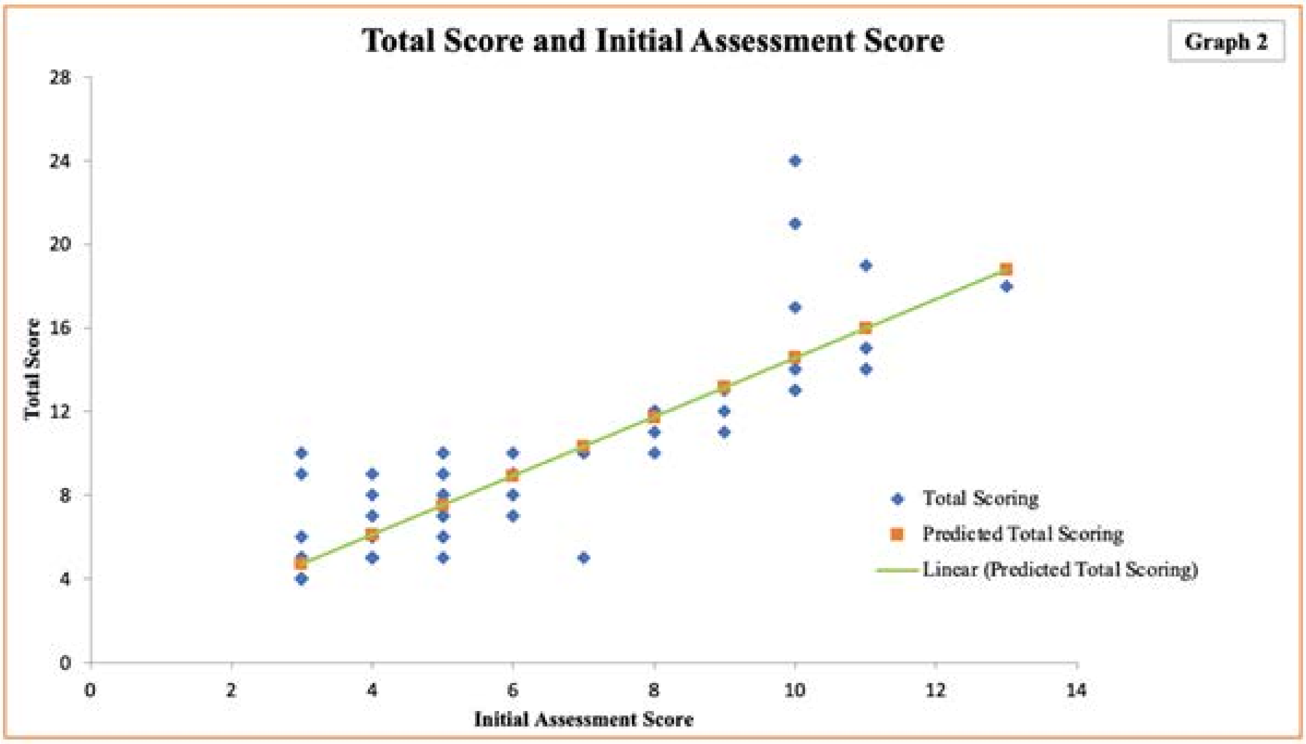
In our original design, the study was structured for a cohort of fifty videos. However, to enhance the statistical power, we doubled the initially calculated sample size resulting in the inclusion of eighty-one videos in our analysis. Notably, our focus was on assessing a subset of the initial fifty videos, concentrating on surgical competency.

Our analysis yielded noteworthy results, unveiling a strong positive regression of R = 0.853 between completion time and total score, accounting for surgical competency. Conversely, we identified a negative correlation of -0.351 between total score and surgical skills, as well as -0.123 between completion time and surgical skills, as depicted in figure 6. These findings indicate an inverse relationship between the variables under scrutiny.



Graph 1:

Citation: H Naseem, A Tam, H Naseem, M Hutton, M Clarke, I Finlay, J Clark. Development and Validation of a Novel Intraoperative Scoring Tool Using Video Analytics to Standardise the Complexity of Gallbladder Surgery. Journal of Surgery and Research. 8 (2025): 423-436.



Graph 2:

5.a Descriptive Statistics		
	Mean	Standard Deviation
Total Score	8.630	4.247
Initial Assessment Score	5.790	2.663
Completion Time	01:05:20	00:43:48

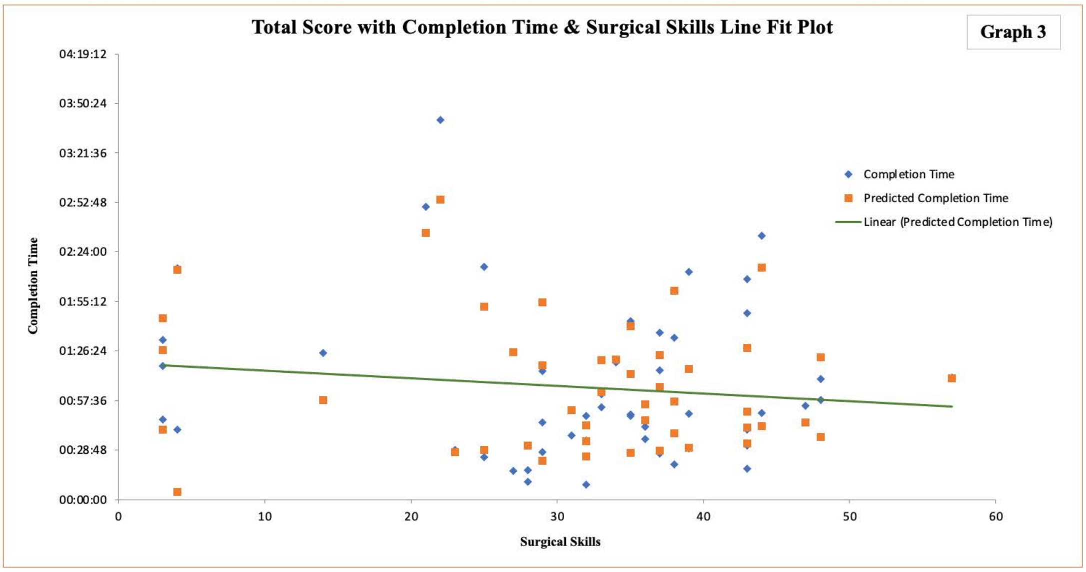
5.b Regression Statistics[a]				
R[b]	R Square	Adjusted R Square	Std. Error of the Estimate	Sig.
0.944[c]	0.89	0.887	1.425	.000

5.c ANOVA[a]					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	1284.592	1	642.296	316.488	.000[b]
Residual	158.297	79	2.029		
Total	1442.889	80			

a Dependent Variable: TS
b. Pearson Correlation
c. Predictors: (Constant), IAS, CT

Figure 5: Total Score (TS) with Completion Time (CT) and Initial Assessment Score (IAS).

Specifically, our data indicates that as surgical skills advance, there is a corresponding reduction in completion time, contingent upon the absence of procedural complications. It is essential to highlight that the study exclusively involved procedures performed by experienced consultants possessing comparable skill sets. As anticipated, this homogeneity resulted in a gradual slope shown in graph-3.



Graph 3:

6.a Pearson Correlation			
	TS	CT	SS
Total Score	1	0.815	-0.351
Completion Time	0.815	1	-0.123
Surgical Skills	-0.351	-0.123	1

6.b Regression Statistics[a]				
R[b]	R Square	Adjusted R Square	Std. Error of the Estimate	Sig.
0.853[c]	0.728	0.717	2.276	.000

6.c ANOVA[a]					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	652.847	2	326.424	295.312	.000[b]
Residual	243.473	47	5.18		
Total	896.32	49			

a Dependent Variable: TS
b. Pearson Correlation
c. Predictors: (Constant), CT, SS

Figure 6: Total Score (TS) with Completion Time (CT) and Surgical Skills (SS).

Total score with surgical skills, completions time, initial assessment scoring, and initial assessment time

Lastly, our investigation delved into the intricate relationship between Total Score and other key variables, namely surgical skills, completion time, initial assessment score, and initial assessment time. The outcomes reveal a robust positive regression with an R-value of 0.965 across all five variables. The model's goodness-of-fit is solidified by a residual sum of squares (SS) value of 60.983, significantly lower than the total residual SS value of 896.32. This affirms the accuracy of the proposed model. Moreover, the statistical significance, as indicated by the F-test (p-value = 0.000), is well below the 5% threshold, confirming the overall significance of our model as shown in figure 7.

Validation

The scoring tool was internally validated by three experienced Royal Cornwall Hospital Trust surgeons blinded to the identity of operating surgeons and themselves not being involved in operations. Intraclass Correlation Coefficient of 0.967 showed a high degree of agreement with average measures for 0.834 to 0.996 in a 95% confidence interval for the Total Score (Figure 8). This statistical assessment further confirms the consistency and reliability of the scoring tool's outcomes.

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

7.a Regression Statistics[a]				
R[b]	R Square	Adjusted R Square	Std. Error of the Estimate	Sig.
0.965[c]	0.932	0.926	1.164	.000

7.b ANOVA[a]					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	835.337	4	208.834	154.1	.000[b]
Residual	60.983	45	1.355		
Total	896.32	49			

a. Dependent Variable: TS

b. Pearson Correlation

c. Predictors: (Constant), IAT, SS, IAS, CT

Figure 7: Total Score (TS) with Surgical Skills (SS), Completion Time (CT) and Initial Assessment.

Total Score - Intraclass Correlation Coefficient between the Validators							
	Intraclass Correlation[a]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.907[b]	0.626	0.989	30.372	4	8	.000
Average Measures	.967[c]	0.834	0.996	30.372	4	8	.000

Initial Assessment Score - Intraclass Correlation Coefficient between the Validators							
	Intraclass Correlation[a]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.888[b]	0.566	0.987	24.8	4	8	.000
Average Measures	.960[c]	0.796	0.996	24.8	4	8	.000

a Type C intraclass correlation coefficients using a consistency definition. The between-measure variance is excluded from the denominator variance.

b The estimator is the same, whether the interaction effect is present or not.

c This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Figure 8: Inter- Rater Reliability

Discussion

Laparoscopic cholecystectomy is one of the most common operations performed worldwide by surgeons in elective and emergency settings. It has also become the index operation of training for upcoming trainee surgeons [3,20]. Recent technological strides have empowered researchers to capture intraoperative intricacies. In response to these advancements, we have devised an intraoperative scoring tool that is both quantifiable and time-sensitive for laparoscopic cholecystectomies, remaining unaffected by preoperative factors. This tool not only predicts the estimated operative time at the outset but also represents a concerted effort to optimize the scheduling and management of theatre lists, alleviating waiting times and enhancing overall efficiency. By adhering to European working time directives, it simultaneously provides essential services and training opportunities, particularly in healthcare systems under strain [20-22]. Moreover, in the future, after obtaining more robust data, incorporating, and integrating this tool with artificial intelligence gives some objectivity to prevailing somewhat inaccurate and often erroneous estimates by surgeons. Galileo Galilei once said, "Measure what is measurable and make measurable what is not so".

Theatre allocation is one of many factors seen in poor theatre utilisation. However, Mizumoto et al. [23], in their randomised control trial, have shown that a surgeon-led model can significantly improve theatre change over time. Such tools can help put surgeons into the driving seat and organise theatre management. In 2020, the Royal Cornwall Hospital was facing a concerning situation, with 4% of patients awaiting elective surgery for over 52 weeks, marking a notable increase from the 2% reported in August 2019 [24].

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

Various studies have shown that direct operating theatre cost has been calculated and ranges from £14 - £40 per minute in NHS and $30 - $100 per minute in the USA. Surgical care accounted for about one-third of healthcare costs and aggregated half of the inpatient hospital costs [8,25-29]. One of the factors for poor theatre utilisation was over or under-booking of theatre lists and unexpected cancellations costing the different healthcare systems millions.

Our proposed scoring tool aims to reduce these alarming statistics, where the predictive completion time can assist with better scheduling of theatre times by accommodating emergency surgeries or reallocating theatres when they become available, reducing the extensive waiting lists.

The landscape of intraoperative intricacy grading tools for laparoscopic cholecystectomy is currently limited. Notably, the Tokyo Guidelines [9] predominantly consider preoperative factors, while AAST [12] and PGS [11] focus on intraoperative factors assessed preoperatively through imaging. Elkbuli et al. [29] underscore the drawbacks of existing tools, emphasising the misalignment of preoperative findings with intraoperative exploration. They highlight that PGS is promising in capturing gallbladder appearances but fail to note its deficiency in considering gallbladder size, a crucial factor related to complications. In response, our proposed scoring tool, detailed in Category B (Figure 1), addresses this gap by incorporating gallbladder size and wall characteristics.

The CholeS [14] study aims to predict the operative time for elective laparoscopic cholecystectomy, focusing on cases exceeding 90 minutes. Simultaneously, it endeavours to enhance theatre scheduling for elective lists [27]. El Sharkawy et al. [15] introduced the Cholecystectomy as A Day Case (CAAD) Score in 2019, derived from data in the CholeS study's subset of 3662 patients nearly half of the original study ((49.3%). The CAAD Score, designed to predict day-case feasibility with a score of 5 or less out of 15, aligns with the overarching goal of optimizing theatre utilization and cost-effectiveness. While both CholeS and CAAD scores focus solely on elective laparoscopic cholecystectomies, our proposed tool offers a more versatile application, catering to both elective and emergency cases, encompassing early and late laparoscopic cholecystectomies.

Our scoring tool presents a comprehensive approach to laparoscopic cholecystectomy assessment, addressing limitations in existing grading systems. Its applicability to a broader range of cases, including both elective and emergency scenarios, positions it as a more versatile and accurate tool for optimizing theatre utilization and cost-effectiveness in the realm of laparoscopic cholecystectomy.

Pawan et al. [30] have suggested that LC is considered complex if the dissection of Calot's triangle and gallbladder bed takes more than 20 minutes each, along with other factors such as gallbladder tearing, bile spillage, and stones. Nasser et al. [31] also tried to navigate factors affecting dissection to achieve a difficult Calot's triangle dissection and critical view of safety. Alfred Cuschieri also proposed a difficulty scale for cholecystectomy to stratify it [3]. In our proposed scoring tool, the scores of three and four in Categories C and E reflect a fibrosed and complicated Calot's triangle and its dissection. It equates to approximately 23 to 30 minutes each by time per score calculations, see figure 1. Our study aligns with the findings of Cuschieri, Pawan, and Nasser. However, a crucial distinction is our reliance on data-driven time assessments, contrasting Pawan et al.'s methodology based on clinical experience.

In a recent study by Ewen A. Griffiths et al. [10], the Nassar operative scale was validated using data from the CholeS Study's two cohorts. Preoperative variables like patient demographics and ultrasound imaging were utilized, with surgeons grading operative data on a scale of 1 to 5. The investigation focused on several preoperative factors, indicating worse outcomes with increasing age, gender, CBD dilation, and conditions like CBD stones or cholecystitis. Moreover, the study noted that higher BMI and thick-walled gallbladders were associated with significantly longer operative durations.

Our developed scoring tool assigns a score of four to thick-walled gallbladders in Category B due to the potential complexity they present, roughly equivalent to 30 minutes according to our scoring system (Figure 1). This tool evaluates disease severity, encompassing factors such as gallbladder size, wall thickness, adhesions, inflammation degree, and Calot's triangle, using an Initial Assessment Score. The average initial assessment time calculated in our study was 13:26 minutes (Figure 3). This phase, occurring early in the procedure when the peritoneal envelope is opened, enables surgeons to gauge complexity, foresee challenges, and anticipate subsequent steps.

The Initial Assessment Score aids in predicting the Total Score by evaluating Calot's triangle dissection difficulty and the potential need for additional procedures from the outset. TS correlates with the estimated Completion Time by multiplying the time per score factor. This dynamic tool accounts for intraoperative complications, adjusting Total Score due to events like excessive bleeding necessitating haemostatic control, thereby impacting the Completion Time. Our study's data supports this correlation, displaying a robust agreement between Total Score, Completion Time, and Initial Assessment Score as evidenced by a multivariate correlation of 0.944 (Figure 5).

This study introduces the Time Score, a novel metric designed to predict procedure completion time in laparoscopic cholecystectomy. Our findings reveal a calculated time per score, providing a practical and viable predictor for

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

surgical durations. For instance, a score of 4 corresponds to approximately 29:36 minutes, offering valuable insights for surgical planning. In our opinion, the Time Score can be customized for both hospital-based surgical teams and individual surgeons in private practices. We envision extending this project to multiple centres to gather extensive data, ensuring external validation and establishing universality for the Time Score. Such information and comprehensive data can be beneficial for its incorporation into Artificial Intelligence in future. Integration of artificial intelligence studies has been reported. Lauren G et al. [32], highlighted that "64% of studies have reported improvements in clinical decisions with automated decision support", providing real-time suggestions. Choudhury A et al. [33] arrived at a similar conclusion in their systemic review; however, it requires further validation.

This scoring tool facilitates real-time decision-making in the operating room, allowing for efficient utilization of resources by adapting surgical schedules and anaesthesia times to address unexpected challenges. Its dynamic nature empowers surgical teams to enhance patient care by navigating uncertainties during laparoscopic cholecystectomy procedures. Surgeons often face challenging decisions under time constraints, and our tool aims to alleviate such pressures by reducing difficult choices.

Additionally, the aspect of surgeons' competency in performing the procedure was considered using the GOALS tool, ensuring a standardised evaluation. As explained by Melina C et al. [34], "The Global Operative Assessment of Laparoscopic Skills" (GOALS) consists of a 5-item global rating scale for competence and case difficulty was used. In a study by O'Connell L et al. [35], the crucial surgical steps for laparoscopic cholecystectomy were also considered to assess surgical skills [23,34,36]. Sage Difficulty Index and Cuschieri scale for cholecystectomy were also studied to develop the severity level within the proposed scoring system [3,37]. In our study, we chose all surgeons to be consultants with the premise that they would be of similar competencies and skill sets to reduce bias. Graph-3, where the line is almost levelled, accentuates the consistency of competency among consultants, affirming our assumption. A subset of the initial fifty videos, constituting sixty-two per cent of the total, assessed surgical skill. The study's power was later increased to eighty-one for enhanced statistical strength. While we believe this adjustment may not heavily impact the current graph. We ruminate that it holds importance for future studies evaluating trainees with varying experience levels.

Three experienced surgeons at the Royal Cornwall Hospital validated the proposed scoring tool with a high degree of agreement, averaging between 0.960 and 0.996 in a 95% confidence interval (Figure 8). These robust results affirm the tool's consistency in predicting cholecystitis severity and operative time. The tool, unique as a single-centre approach, captures intra-operative aspects of cholecystitis, anatomical variations, and a hostile abdomen with consideration for time. It introduces additional procedures not found in previous scoring systems, collectively influencing operative times for enhanced practicality. Notably, the tool excludes preoperative factors like obesity, CBD stones, or prior emergency admissions, simplifying its usability. Despite this exclusion, the author recognizes the importance of pre-operative factors, routinely considered by surgeons in their standard practice reviews and planning.

The current study has some limitations due to the desire to simplify surgeons' involvement and ensure a broad, ethically agreed international input. Developing a scoring system needs two stages, development and validation, and a much more comprehensive validation is ideal. A small data set from a single centre may have introduced a bias in the study with several cases as assumed failed to record as the hospital comprises only one theatre with a recording facility. Additionally, reluctance amongst the surgeons to participate and allow their theatre recordings to be assessed was also a hurdle to the achievability of this research. A relatively small data set also included some outliers, a maximum Cook's distance of 0.847, which appeared to exert undue influence on the regression line. Finally, COVID-19 has been a significant challenge where surgical practices have been suspended. The scoring tool's strength lies in its dynamic nature but requires manual adjustments during the operation, which could be addressed through future AI integration. We are fully committed to enhancing the validity of this study by gathering data from multiple institutions, ensuring its robustness, and thoroughly investigating its practical utility.

The proposed Humair & Clark scoring system demonstrates robust potential as a practical predictor of cholecystitis severity and laparoscopic cholecystectomy complexity, notably correlated with operation duration. Its primary strength lies in its real-time prediction capability, enhancing theatre efficiency by facilitating precise scheduling and anaesthesia planning. This tool's adaptability serves a threefold purpose: Firstly, its capacity to predict final scores and completion time from the outset of initial assessment, streamlines theatre utilization and aids in forecasting anaesthesia requirements, thereby optimizing operational efficiency. Secondly, its quantitative assessment feature is instrumental in tracking clinical residents' progress and allocating appropriate cases based on their skill level, thereby enhancing training program efficacy. Finally, the tool's potential integration with Artificial Intelligence fortifies its future relevance, promising advancements in surgical skill quantification, simulation-based training, refined scheduling, surgeon certification, and tool development and assessment. This envisaged integration holds promise for enhancing the tool's adaptability and efficacy.

Conclusions

We propose a Naseem & Clark scoring system which shows promise as a predictor of cholecystitis severity and laparoscopic cholecystectomy complexity, linked to operation duration. Its real-time prediction capability enhances theatre efficiency by aiding in almost precise scheduling and anaesthesia planning. The tool's adaptability serves three main purposes: predicting final scores and completion time from initial assessment to optimise theatre usage. Secondly, quantitatively assessing clinical residents' progress to enhance training program efficacy. Finally, potential integration with Artificial Intelligence for future advancements in surgical skill quantification, simulation-based training, refined scheduling and surgeon certification.

Declarations & Disclosures

Ethical Approval

We were advised by a local ethical approver in the Research Development and Innovation department at Royal Cornwall Hospital Treliske that our research does not require any ethical approval due to its nature and anonymity. Our study and data do not reveal the patient identity or patients' static data.

Declaration of Conflicts of interest /competing interests

Humair Naseem, Adam Tam, Hasnat Naseem, Michael Hutton, Michael Clarke, Ian Finlay and James Clark declare that there is no conflict of interest regarding the publication of this article.

Financial and Funding Disclosures

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors. There were no sources of funding.

Disclosure and Declaration

Humair Naseem, Adam Tam, Hasnat Naseem, Michael Hutton, Michael Clarke, Ian Finlay and James Clark have no conflicts of interest or financial ties to disclose.

References

1. Douglas M Heuman. Gallstones (cholelithiasis). Practice Essentials, Background, Pathophysiology 18 (2023): 12
2. Soper NJ, Stockmann PT, Dunnegan DL, et al. Laparoscopic Cholecystectomy The New "Gold Standard"? Archives of Surgery 127 (1992): 917.
3. Tafazal H, Spreadborough P, Zakai D, et al. Laparoscopic cholecystectomy: A prospective cohort study assessing the impact of grade of operating surgeon on operative time and 30-day morbidity. Annals of the Royal College of Surgeons of England 100 (2018): 178-184.
4. Rance C, Jones A. Gallstone disease. InnovAiT 9 (2016): 11-17.
5. Glaysher MA, May-Miller P, Carter NC, et al. Specialist-led urgent cholecystectomy for acute gallstone disease. Surgical endoscopy 37 (2023): 1038-1043.
6. Lee DJ, Ding J, Guzzo TJ. Improving Operating Room Efficiency. Current urology reports 20 (2019): 28.
7. Gillies MA, Wijeysundera DN, Harrison EM. Counting the cost of cancelled surgery: a system wide approach is needed. BJA: British Journal of Anaesthesia 121(2018): 691-694.
8. Lawrentschuk N, Hewitt PM, Pritchard MG. Elective laparoscopic cholecystectomy: Implications of prolonged waiting times for surgery. ANZ journal of surgery 73 (2003): 890-893.
9. Miura, Fumihiko, Okamoto, et al. Tokyo Guidelines 2018: initial management of acute biliary infection and flowchart for acute cholangitis. Journal of Hepato-Biliary-Pancreatic Sciences 25 (2017): 109
10. Griffiths EA, Hodson J, Vohra RS, et al. Utilisation of an operative difficulty grading scale for laparoscopic cholecystectomy. Surg Endosc 33 (2019): 110-121.
11. Abdul Razack GS, Avinash K, Manjunath BD, et al. Pre-operative evaluation with parkland grading system in assessing difficult laparoscopic cholecystectomy and expectant operative and post-operative complications. Int J Surg Sci 3 (2019): 20-25.
12. Moore EE, Cogbill TH, Jurkovich GJ, et al. Organ injury scaling: spleen and liver (1994 revision). The Journal of trauma 38 (1995): 323-324.
13. Elkbuli A, Meneses E, Kinslow K, et al. Current grading of gall bladder cholecystitis and management guidelines: Is it sufficient?. Annals of medicine and surgery 60 (2012): 304-307.
14. Bharamgoudar R, Sonsale A, Hodson J, et al. The development and validation of a scoring tool to predict the operative duration of elective laparoscopic cholecystectomy. Surgical endoscopy 32 (2018): 3149-3157.
15. El-Sharkawy AM, Tewari N, Vohra RS, et al. The Cholecystectomy as a day case (CAAD) Score: A Validated Score of Preoperative Predictors of Successful Day-Case Cholecystectomy Using the CholeS Data Set. World J Surg 43 (2019): 1928-1934.
16. Tanzi L, Piazzolla P, Vezzetti E. Intraoperative surgery room management: A deep learning perspective. International Journal of Medical Robotics and Computer Assisted Surgery 16 (2020): 1-12.

H Naseem, et al., J Surg Res 2025
DOI:10.26502/jsr.10020468

17. Koulas SG, Tsimoyiannis J, Koutsourelakis I. Laparoscopic cholecystectomy performed by surgical trainees. Journal of the Society of Laparoendoscopic Surgeons 10 (2006): 484-487.

18. Hashimoto DA, Rosman G, Volkov MS, et al. Artificial Intelligence for Intraoperative Video Analysis: Machine Learning's Role in Surgical Education. Journal of the American College of Surgeons 225 (2017): S171.

19. Van Der Ven WH, Veelo DP, Wijnberge M, et al. One of the first validations of an artificial intelligence algorithm for clinical use: The impact on intraoperative hypotension prediction and clinical decision-making. Surgery 169 (2021): 1300-1303.

20. Wyman MG, Huynh R, Owers C. The European Working Time Directive: Will Modern Surgical Training in the United Kingdom Be Sufficient?. Cureus 14 (2022): e21797.

21. Breen KJ, Hogan AM, Mealy K. The detrimental impact of the implementation of the European working time directive (EWTD) on surgical senior house officer (SHO) operative experience. Irish journal of medical science 182 (2013): 383-387.

22. Schimmack S, Hinz U, Wagner A, et al. Maximizing time from the constraining European Working Time Directive (EWTD): The Heidelberg New Working Time Model. Health Economics Review 4 (2014): 56.

23. Mizumoto R, Cristaudo AT, Hendahewa R. A surgeon-led model to improve operating theatre change-over time and overall efficiency: A randomised controlled trial. International journal of surgery (London, England) 30 (2016): 83-89.

24. Ferguson E. Royal Cornwall Hospitals waiting times record high due to Covid. Falmouth Packet 23 (2020).

25. Cerfolio RJ, Ferrari-Light D, Ren-Fielding C, et al. (2019). Improving Operating Room Turnover Time in a New York City Academic Hospital via Lean. The Annals of thoracic surgery 107 (2019): 1011-1016.

26. Childers CP, Maggard-Gibbons M. Understanding Costs of Care in the Operating Room. JAMA Surgery 153 (2018): e176233.

27. Brodsky JB. Cost Savings in the Operating Room. Anesthesiology 11 (2008): 61-72

28. Moody AE, Gurnea TP, Shul CP, et al. True Cost of Operating Room Time: Implications for an Orthopaedic Trauma Service. Journal of orthopaedic trauma 34 (2020): 271-275.

29. Elkbuli A, Meneses E, Kinslow K, et al. Huge gangrenous gallbladder presenting as gastro-esophageal reflux disease successfully treated by laparoscopic cholecystectomy: Case report and literature review. International Journal of Surgery Case Reports 11 (2020): 31.

30. Lal P, Agarwal PN, Malik VK, et al. A difficult laparoscopic cholecystectomy that requires conversion to procedure can be predicted by preoperative ultrasonography. JSLS: Journal of the Society of Laparoendoscopic Surgeons 6 (2020): 59-63.

31. Nassar AHM, Ng HJ, Wysocki AP, et al. Achieving the critical view of safety in the difficult laparoscopic cholecystectomy: a prospective study of predictors of failure. Surgical endoscopy 35 (2021): 6039-6047.

32. Gordon L, Grantcharov TP, Rudzicz F. Explainable Artificial Intelligence for Safe Intraoperative Decision Support. JAMA Surgery 154 (2019): 1064.

33. Choudhury A, Asan O. Role of Artificial Intelligence in Patient Safety Outcomes: Systematic Literature Review. JMIR medical informatics 8 (2020): e18599.

34. Vassiliou MC, Feldman LS, Andrew CG, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. American Journal of Surgery 190 (2005): 107-113.

35. O'Connell L, McKevitt K, Khan W, et al. Impact of targeted trainer feedback via video review on trainee performance of laparoscopic cholecystectomy. The surgeon: Journal of the Royal Colleges of Surgeons of Edinburgh and Ireland 19 (2021): e107-e111.

36. Moorthy K, Munz Y, Sarker SK, et al. Objective assessment of technical skills in surgery. BMJ Clinical research 327 (2003): 1032-1037.

37. Everson, H. (Ed.) (Vols. 1-4). Sage Publications, Inc., (2018) pp: 522 523.