The American Journal of Surgery 217 (2019) 90–97

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

Prospective validation of the Parkland Grading Scale for Cholecystitis

Tarik D. Madni [a], Paul A. Nakonezny [c], Evan Barrios [b], Jonathan B. Imran [a], Audra T. Clark [a], Luis Taveras [a], Holly B. Cunningham [a], Alana Christie [c], Alexander L. Eastman [b], Christian T. Minshall [b], Stephen Luk [b], Joseph P. Minei [b], Herb A. Phelan [b], Michael W. Cripps [b, *]

[a] UT Southwestern Department of Surgery, Dallas, TX, USA
[b] UT Southwestern Division of General and Acute Care Surgery, Dallas, TX, USA
[c] UT Southwestern Department of Clinical Sciences, Division of Biostatistics, Dallas, TX, USA

ARTICLE INFO

Article history:
Received 12 June 2018
Received in revised form
6 August 2018
Accepted 13 August 2018

Keywords:
Gallbladder
Cholecystitis
Score
Grade
Outcomes
Quality improvement

ABSTRACT

Background: The Parkland Grading Scale for Cholecystitis (PGS) was developed as an intraoperative grading scale to stratify gallbladder (GB) disease severity during laparoscopic cholecystectomy (LC). We aimed to prospectively validate this scale as a measure of LC outcomes.
Methods: Eleven surgeons took pictures of and prospectively graded the initial view of 317 GBs using PGS while performing LC (LIVE) between 9/2016 and 3/2017. Three independent surgeon raters retrospectively graded these saved GB images (STORED). The Intraclass Correlation Coefficient (ICC) statistic assessed rater reliability. Fisher's Exact, Jonckheere-Terpstra, or ANOVA tested association between perioperative data and gallbladder grade.
Results: ICC between LIVE and STORED PGS grades demonstrated excellent reliability (ICC = 0.8210). Diagnosis of acute cholecystitis, difficulty of surgery, incidence of partial and open cholecystectomy rates, pre-op WBC, length of operation, and bile leak rates all significantly increased with increasing grade.
Conclusions: PGS is a highly reliable, simple, operative based scale that can accurately predict outcomes after LC.
Table of contents summary: The Parkland Grading Scale for Cholecystitis was found to be a reliable and accurate predictor of laparoscopic cholecystectomy outcomes. Diagnosis of acute cholecystitis, surgical difficulty, incidence of partial and open cholecystectomy rates, pre-op WBC, operation length, and bile leak rates all significantly increased with increasing grade.

Introduction

According to the Agency for Healthcare Research and Quality (AHRQ), cholecystectomy was the fifth most common overall operation performed in the United States and the most common operation paid for by Medicaid.[1] Through the Affordable Care Act, and thus Medicaid expansion, our healthcare environment continues to shift towards high quality, reduced cost, bundled care.[2] As such, surgeons' outcomes will be under closer scrutiny with harsher penalties than previous years. With limited resources, surgeons will be forced to not only operate in the most cost-effective and efficient manner but will also be evaluated via their post-operative outcomes and complications. However, even common procedures can vary from patient to patient due to patient specific factors such as comorbidities as well as from significant anatomical and inflammatory differences appreciated only in the operative environment.

Currently, complication rate reporting for Laparoscopic Cholecystectomy (LC) is uniform regardless of case complexity or level of

* Corresponding author. University of Texas Southwestern Medical Center, Division of General and Acute Care Surgery, 5323 Harry Hines Blvd., Dallas, TX 75390-9158, USA.

E-mail addresses: tarik.madni@utsouthwestern.edu (T.D. Madni), Paul.nakonezny@utsouthwestern.edu (P.A. Nakonezny), evan.barrios@utsouthwestern.edu (E. Barrios), jonathan.imran@utsouthwestern.edu (J.B. Imran), audra.clark@utsouthwestern.edu (A.T. Clark), Luis.taveras@utsouthwestern.edu (L. Taveras), Holly.cunningham@utsouthwestern.edu (H.B. Cunningham), alana.christie@utsouthwestern.edu (A. Christie), alexander.eastman@utsouthwestern.edu (A.L. Eastman), cminshal@hotmail.com (C.T. Minshall), stephen.luk@utsouthwestern.edu (S. Luk), joseph.minei@utsouthwestern.edu (J.P. Minei), herb.phelan@utsouthwestern.edu (H.A. Phelan), michael.cripps@utsouthwestern.edu (M.W. Cripps).

https://doi.org/10.1016/j.amjsurg.2018.08.005
0002-9610/

inflammation. As such, a highly inflamed gallbladder (GB) that may take two hours to remove with a higher chance for post-operative complications is coded the same as a gallbladder that can be removed in 20 min. Treatment for acute cholecystitis in particular is now often treated by separate acute care (ACS) and emergency general surgery (EGS) services given improved outcomes and decreased lengths of stay by such services.[3] Surgeons who opt to take on more difficult cases should thus have their outcomes judged differently compared to those with less complexity. It is imperative that a grading system be promulgated which differentiates between the two.

A large number of pre-operative cholecystitis grading scales have been developed in the past that aim to predict both intra- and post-operative outcomes; however, few of these scales account for intra-operative anatomical differences.[4–22] In addition, other than prediction of difficult cases, few of these scales attempt to be utilized as a manner of post-operative complication comparison. Ultimately, none of these grading scales have been used in clinical practice as they are complex, difficult to remember, and do not allow for effective outcome comparisons. To date, there are no *widespread*, validated grading systems utilized to stratify the intraoperative severity of GB inflammation.

The *Parkland Grading Scale for Cholecystitis* (PGS) was previously developed to stratify the severity of GB disease in response to these pitfalls.[23] This five-tiered, easy to implement, grading system based on anatomy and inflammatory changes was previously demonstrated to be highly reproducible with an Intraclass Correlation Coefficient of 0.804. We aimed to prospectively validate this grading scale as a measure of LC outcomes. The hypothesis was that changes in gallbladder anatomy and inflammation (pursuant to the grading scale) would be related and consistent with peri- and post-operative outcomes. If found to be valid and reliable, this scale of gallbladder anatomy and inflammation can be made to allow for fairer outcome comparisons and complication rate reporting.

Methods

Design and procedures

This study was approved by the institutional review board at the University of Texas Southwestern Medical Center. The Parkland Memorial Hospital ACS faculty perform both urgent, inpatient LC and elective LC as part of the group's practice pattern. Typically, the urgent and inpatient LCs tend to include the diagnoses of acute/chronic cholecystitis or biliary pancreatitis, while the more elective LCs tend to include biliary colic/symptomatic cholelithiasis or biliary dyskinesia. All of our acute cholecystitis patients receive antibiotics until their operation is performed. During this study period, the antibiotic of choice was piperacillin-tazobactam; however, our group has recently switched to a new regimen of ceftriaxone and metronidazole. Any delay to the operating room due to case load or more urgent cases may result in some patients receiving a few extra doses of antibiotics prior to their operation.

Since October 2015, the standard operating procedure for LC at Parkland Memorial Hospital is to take intraoperative pictures of the GB using the laparoscope once the GB is initially visualized. All patients at Parkland Memorial Hospital who underwent LC or open cholecystectomy by the ACS service between 9/2016 and 3/2017 were eligible to be included in the study. Eleven ACS faculty were asked to self-grade the “initial view” of GBs when performing LC during this 7-month period using the previously developed PGS (LIVE) (Fig. 1).[23] The initial grade was based solely on the objective criteria of the scale and was assigned immediately after the “initial view” was visualized. The “initial view” was defined as follows: After placement of all four laparoscopic ports.

1) If the GB was visualized easily, it was grasped and retracted cephalad giving the “initial view”.
2) If severe inflammation was present which limited mobilization or the ability to visualize the GB, the “initial view” was defined as the view of the inflamed area.

If the faculty surgeon was scrubbed in at the time of the initial view, they were instructed to not change or alter the cholecystitis grade based on further case findings or complexity. These raters then filled out a subjective, post-operative questionnaire regarding case difficulty and peri-operative factors after each operation. This questionnaire was only allowed to be filled out in the post-operative period as to eliminate any potential time bias from surveys filled out at later times. In addition, the faculty were not aware that any correlation between cholecystitis score and case difficulty was to be made to limit any potential scoring bias.

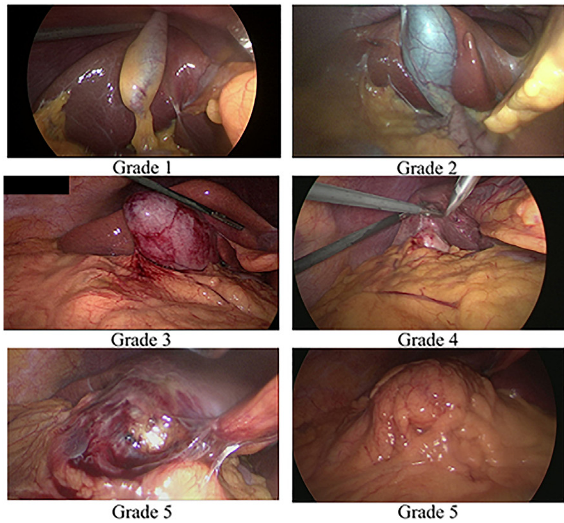
Finally, three surgeon raters who were completely independent of the original study retrospectively reviewed 210 GB “initial views” from the study period that were stored on the electronic medical record (EMR). The 210 images were randomly assigned to the 3 raters (i.e., each of the 3 raters were randomly assigned 70 GB images). Each of the 3 independent raters then rated his or her 70 GB images utilizing PGS (STORED).

Training standard of care

At our institution, the general training pathway for LCs consists of attending physicians taking interns through easier outpatient LCs while the more difficult inpatient LCs are performed by third-year residents supervised by a chief resident as first assistant and an attending physician who is physically present but who may or may not be scrubbed. As the case progresses with the third-year resident functioning as primary surgeon, the chief resident and attending physician monitor progress and safety of the dissection as it proceeds. Should the difficulty of the case exceed the skill set of the third-year resident, the chief resident will typically take over as operating surgeon and the attending physician will move to the role of first assistant. Finally, if the difficulty of the case appears too great, the attending physician will assume the role of primary surgeon. Internally tracked data suggests PGY-3 residents are the primary surgeon in roughly 80% of our LCs, with faculty needing to physically scrub about 15% of LCs.

Measures

Peri-operative factors collected during the questionnaire included: presence of intra-abdominal adhesions (yes/no), normality of anatomy (normal/abnormal), duct closure device utilized (clip/staple), difficulty of surgery measured on a 5-point Likert-Type scale that ranged from 1 (least difficult) and 5 (most difficult). Patient, peri-, and post-operative characteristics collected from the medical record included: age (years), gender, pre-operative laboratory values (white blood cell count (WBC), total bilirubin), aspiration (yes/no), performance of intra-operative cholangiography (yes/no), partial or open cholecystectomy requirements (yes/no), length of operation (minutes), post-operative/pathologic diagnosis (acute cholecystitis, chronic cholecystitis, biliary dyskinesia, biliary pancreatitis, choledocholithiasis, symptomatic cholelithiasis), length of stay (days), post-operative retained stone (yes/no), post-operative wound infection (yes/no), and post-operative bile duct leak within 60 days (yes/no). Our hypothesis was that increases in PGS would correlate with increases in the peri-operative outcomes: conversion rates (open and partial), length and difficulty of operation, and rates of bile duct leaks.



Cholecystitis Severity Grade	Description of Severity
1	Normal appearing gallbladder ("robin's egg blue") • No adhesions present • Completely normal gallbladder
2	Minor adhesions at neck, otherwise normal gallbladder • Adhesions restricted to the neck or lower of the gallbladder
3	Presence of ANY of the following: • Hyperemia, pericholecystic fluid, adhesions to the body, distended gallbladder
4	Presence of ANY of the following: • Adhesions obscuring majority of gallbladder • Grade I-III with abnormal liver anatomy, intrahepatic gallbladder, or impacted stone (Mirrizi)
5	Presence of ANY of the following: • Perforation, necrosis, inability to visualize the gallbladder due to adhesions

Fig. 1. The Parkland grading scale for cholecystitis.

Statistical analysis

Demographic characteristics of the patients were described using the sample mean and standard deviation for continuous variables and the frequency and percentage for categorical variables. The Fischer's Exact Test was used to evaluate the association of grade with the binary peri-operative clinical variables, Jonckheere-Terpstra test was used for doubly-ordered categorical variables, and one-way ANOVA was used to assess the association of grade with continuous peri-operative clinical measures.

We used the Welch's one-way ANOVA, which is a robust test of equality of means, to evaluate the relationship of PGS with the continuous outcomes of difficulty of surgery and, in a separate model, length of surgery. The Tukey-Kramer *post-hoc* test was used to evaluate all pairwise comparisons among the 5 grades.

Since our previous work found PGS reliable through retrospective review of still, intra-operative images,[23] we aimed to confirm the utility of PGS during intra-operative use. Thus, we assessed the reliability (or consistency) of the cholecystitis grading scale ratings of the faculty surgeon who performed the operation as well as the three independent raters with both each other and that of the difficulty of surgery ratings using an ICC statistic.

The alpha-level for all tests was set at 0.05 (two-tailed) and, to address multiple testing, p-values were adjusted using the False Discovery Rate (FDR) procedure. Statistical analyses were carried out using SAS software, version 9.4 (SAS Institute, Inc, Cary NC).

Results

A total of 317 LCs were graded utilizing PGS. All cholecystectomies during the study period were attempted laparoscopically, none were started via open technique. Table 1 demonstrates the number of GBs graded, as well as the average assigned cholecystitis severity grade and difficulty score per rater. Fig. 2 visually demonstrates the relationship (observed cumulative frequency distribution) between **a)** each of the 11 original raters and assigned surgical difficulty score, **b)** each of the 11 original raters and assigned cholecystitis severity grade. There were 60 grade 1 GBs (19%), 90 Grade 2 GBs (28%), 102 Grade 3 GBs (32%), 28 Grade 4 GBs (9%), and 37 Grade 5 GBs (12%). Further patient as well as peri- and post-operative characteristics are illustrated in Table 2. Sixteen (5%) of patients required partial cholecystectomies, and 9 (3%) of patient required a conversion to open surgery. Forty-three (14%) patients demonstrated either abnormal biliary or arterial anatomy. There were no intra-operative adverse events. Regarding post-operative complications, 5/317 patients (1.58%) had a bile duct leak, 3/317 (0.95%) had wound infections, and 4/317 (1.26%) had retained stones. Diagnosis of acute cholecystitis, difficulty of surgery, partial cholecystectomy rate, open conversion rate, pre-op WBC, length of operation, and presence of a post-operative bile duct leak all significantly increased with increasing grade. (Table 3). Rates of post-operative wound infections and retained stones were not found to increase with increasing PGS grade.

The PGS (via the faculty surgeons who performed and rated the difficulty of the operation) was found to be significantly associated with greater mean difficulty of surgery scores ($R^2 = 0.57$) as well as greater mean length of surgery ($R^2 = 0.19$) with increasing grades in the Welch's 1-way ANOVA (Table 4). The Tukey-Kramer test for all pairwise comparisons revealed that each Grade (1–5) was significantly different from each other (at $p < 0.05$) on difficulty of surgery. The relationship between assigned cholecystitis grade and post-operative difficulty score is further visually represented in Fig. 2c. Regarding length of surgery, however, the Tukey-Kramer test for all pairwise comparisons (at $p < 0.05$) revealed that Grade 1 was significantly different from Grades 3,4,5; Grade 2 was significantly different from Grades 4,5; Grade 3 was significantly different from Grades 1,5; Grade 4 was significantly different from Grades 1,2; and Grade 5 was significantly different from Grades 1,2,3. In a sensitivity analysis, we replicated the ANOVA model described above using ratings of the PGS from the 3 independent raters and found that the basic pattern of findings, as reported in Table 4, persisted (results not reported).

The ICC showed excellent reliability of cholecystitis grading scale ratings with that of the difficulty of surgery ratings by both the surgeon who rated the difficulty of the operation (ICC = 0.857, 95% CI: 0.822 to 0.886, p = 0.0001) as well as the three independent raters (ICC = 0.7389, 95% CI: 0.6574 to 0.8011, p = 0.0001). Finally, the ICC between cholecystitis severity grade rated by the three independent raters and the faculty surgeons who performed the operation demonstrated excellent reliability (ICC = 0.8210, 95% CI: 0.7652 to 0.8636, p = 0.0001).

Table 1
Total gallbladders graded and average scores per rater.

Rater	Total Gallbladders Graded N (%)	Average Cholecystitis Grade (M±SD)	Average Difficulty Score (M±SD)
1	60 (19%)	2.5 ± 1.2	2.2 ± 1.2
2	4 (1%)	1.8 ± 1.0	2.3 ± 1.3
3	41 (13%)	2.8 ± 1.0	2.7 ± 1.0
4	61 (19%)	3.0 ± 1.3	2.7 ± 1.4
5	32 (10%)	2.7 ± 1.2	2.7 ± 1.1
6	35 (11%)	2.1 ± 1.2	2.2 ± 1.2
7	56 (18%)	2.7 ± 1.2	2.7 ± 1.4
8	18 (6%)	3.2 ± 1.3	2.7 ± 1.1
9	1 (<1%)	1.0 ± 0.0	1.0 ± 0.0
10	7 (2%)	2.1 ± 1.5	2.1 ± 1.5
11	2 (1%)	4.0 ± 1.4	3.5 ± 0.7
Total	**317 (100%)**		

Note: LC = Laparoscopic Cholecystectomy, M = Mean, SD = Standard Deviation.

Discussion

The purpose of this study was to validate the PGS as an accurate measure of LC outcomes to pave the way for a more simple, accurate measure of outcome comparison. Here we present a simple, five-tiered scoring system that is found to correlate significantly with both post-operative outcomes, such as biliary leak rate, as well as peri-operative factors such as open conversion rate, length of operation, and case difficulty. In addition, an ICC of 0.8210 between the prospective (LIVE) graders and the retrospective (STORED) reviewers confirms the reliability of this scale when utilized in an intra-operative fashion. Overall, these results suggest the potential utility of PGS as an easy to use system for outcome comparison which can be calculated quickly, and reliably, during the intra-operative period.

As our healthcare environment shifts, the need for an accurate manner of outcome comparisons is imperative. Now that surgeons' complications are being published to be viewed by both insurers and the public, the ability to match one operation to another is all the more important. Given this need, our results suggest that a grade 5 gallbladder should be assessed differently than a grade 1, or any other grade, given the increased operating room (OR) utilization and risk for complications (Table 3). For example, the literature has shown that biliary injury during a LC can increase costs by 4.6–26 fold compared to an uncomplicated procedure.[24] These costs include not only the repair for such an injury, but also the additional inpatient care and treatments the patient may require. As these costs can be associated with penalties to the surgeon and the institution, a grade 5 GB with associated biliary injury *should* be viewed less harshly, and thus impose fewer potential penalties, than a grade 1 GB. In our subset, we did not note any direct biliary injury intra-operatively; however, had five (1.5%) biliary leaks. The additional costs of imaging, intervention, and care of these patients cannot be discounted.

Since our aim was to generate a method of outcome comparison, the reader may question why a five-tiered scoring system was chosen versus utilization of the already existing Modifier 22 label. In its original development, we decided on an odd-numbered grading system but felt three grades of severity (i.e., normal, moderate, severe) would be too general (and potentially limiting) of distinctions. Our hypothesis was that a five-tiered grading system would afford a more precise, stepwise range for surgeons to differentiate one gallbladder from another in terms of severity.[23] However, looking at Table 3, the reader will notice most of our partial cholecystectomy rates, open conversion rates, and post-operative biliary leaks occur in Grade 5 GBs seemingly supporting a two category (i.e., visualized perforation/necrosis versus all other GBs) classification. While Modifier 22 may work in such a two-category system, there are more results to compare between LCs than these three binary intra- and post-operative outcomes. For example, Table 4 illustrates a significant 10–21% increase in operative time between adjacent grades and a difficulty score increasing at a similar rate to the peri-operative gallbladder grade. As such, we believe that our basic findings seen in Table 4 that greater mean difficulty of surgery scores as well as greater mean length of surgery with increasing grades demonstrate the need for more than just two grades. While some may argue surgical difficulty and case length can be surgeon dependent in nature, any surgeon should hope that their longer and harder LCs be judged differently than their easier cohorts. PGS attempts to create a scale that standardizes such differences in difficulty. In addition, LC complications are not as common as more complex operations, and only 12/317 (3.8%) of our patients had post-operative complications (5 bile duct leaks, 4 retained stones, and 3 wound infections). None of our patients

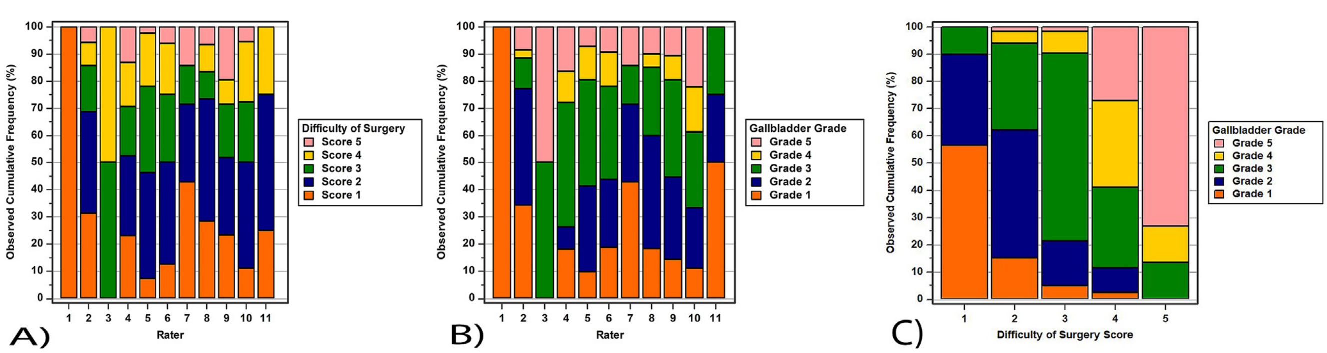


Fig. 2. Relationship (Observed Cumulative Frequency Distribution) Between a) Rater and Surgical Difficulty. b) Rater and Cholecystitis Severity Grade. c) Cholecystitis Grade and Surgical Difficulty.

Table 2
Patient, peri-operative, and post-operative characteristics.

	N (%)
Gender	
Female	244 (77.0)
Male	73 (23.0)
Mean age, years (range)	40.1 ± 13.6 (18–84)
Initial Grade	
1	60 (18.9)
2	90 (28.4)
3	102 (32.2)
4	28 (8.8)
5	37 (11.7)
Diagnosis	
Acute cholecystitis	166 (52.4)
Chronic cholecystitis	14 (4.4)
Biliary dyskinesia	1 (0.3)
Biliary pancreatitis	27 (8.5)
Choledocholithiasis	26 (8.2)
Symptomatic cholelithiasis	83 (26.2)
Intra-abdominal adhesions	
No	163 (51.4)
Yes	154 (48.6)
Gallbladder aspirated?	
No	223 (70.3)
Yes	71 (22.4)
Decompressed with manipulation	23 (7.3)
Partial cholecystectomy	16 (5.0)
Converted to open surgery	9 (2.8)
Duct closure	
Clip	283 (89.3)
Staple	13 (4.1)
None**	21 (6.6)
Abnormal anatomy	43 (13.6)
Intraoperative Cholangiogram	16 (5.0)
Reason for IOC	
Anatomy	4 (25.0)
Intraductal stone	1 (6.3)
Routine	11 (68.8)
Difficulty of surgery	
1	69 (21.8)
2	113 (35.6)
3	61 (19.2)
4	44 (13.9)
5	30 (9.5)
	Mean ± Std. Dev.
Pre-op total bilirubin	0.5 (0.3–0.8)
Pre-op WBC count	9.3 (6.7–12.2)
Length of stay post-op, days	0.4 ± 1.2
Length of operation, minutes	77.9 ± 30.6

**WBC: White blood count, IOC: Intra-operative cholangiogram.*
**No duct closure references to situations in partial and open cholecystectomies where duct closure was not specified, and the remaining cystic remnant may have been fulgurated.

had any intra-operative adverse events. As such, we believe a larger, multi-institutional trial would allow evaluation of differences in such complications among grades.

If utilizing this operative grading scale, the surgeon may question how to act on the information provided. Given grade 5 GBs are longer, harder, have higher conversion rates, and more complications, could it be argued the case should be aborted? We do not necessarily believe so. Rather, we consider two main uses for this scale: 1) To allow for fairer outcome comparisons in complication rate reporting 2) To allow for a potential, pre-emptive change in operative strategy. For example, a senior resident who visualizes a grade 5 GB may be more actively involved or take over the case if being performed by a junior resident. Or, perhaps a faculty surgeon may opt to convert the case to an open operation sooner knowing the higher future complication and conversion rates. These are just a few examples on how operative strategy can change based on a standardized scale supported by evidence.

In addition to outcome comparisons and surgical decision making, residents could also benefit from the incorporation of such a grading scale into case logging practices. As the Resident Review Committee (RRC) increased the minimally invasive surgery (MIS) case requirements from 34 to 60 (basic laparoscopy) and 0 to 25 (complex laparoscopy), many residents and thus training programs may have problems reaching these requirements.[25] In the current American Board of Surgery case log system, residents are forced to log a LC as CPT 47562 regardless of how long or how difficult the case may be. Our findings suggest that it would be incorrect to log both a Grade 1 and a Grade 5 gallbladder as "basic laparoscopy" when the latter takes almost twice the operative time and has a significantly higher case difficulty. It is our opinion that these higher-grade cases could and should be logged as "complex" as they have longer operative times, more difficulty, and higher risks for complications.

In the past, a number of predictive scoring systems for LC have been proposed, the majority of which have dealt with pre-operative findings; however, none of these scoring systems are in current widespread use.[4–22] In contrast to a pre-operative grading scale, an intra-operative scoring system is thought to allow a more meaningful comparison of outcomes or potentially trigger a change in surgical strategy (i.e., conversion to open) to mitigate risk of a potential complication.[20] A few intra-operative scales have also been created in the past; however, most are complex, have not been validated, and have not been widelyadopted.[20,22]

In an effort to standardize comparisons in research and quality of care specifically in emergency general surgery (EGS) patients, the AAST proposed a uniform, anatomic grading scale to measure the severity of disease for 16 common EGS diseases including acute cholecystitis.[26,27] Different from our operative-only findings scale, the AAST scale gives 1 through 5 grade definitions for four categories (clinical, imaging, operative, and pathologic) of each disease. The highest assigned grade becomes the true, overall grade of the disease.[21] Thus, a major limitation of the AAST scale is it requires extensive data collection and the user must wait days to weeks for pathological results prior to its determination. This is in comparison to our operative only scale, where a true grade can be determined immediately during the intra-operative period. However, the AAST anatomic severity scale has been retrospectively tested and validated in other multiple other EGS diseases to date: small bowel obstruciton,[28] skin and soft tissue infections,[29] pancreatitis,[30] appendicitis,[31] diverticulitis,[32] and recently cholecystitis.[33] The validation of this scaling system has not been without limitations.

Specifically focusing on biliary disease, the AAST acute cholecystitis scale was validated as a measure of LC outcomes at a single center institution.[33] In this study, Vera et al. (2018) found excellent reliability between raters (kappa coefficient of 1.0) and the AAST cholecystitis grade was found to correlate with post-operative morbidity. Several limitations were, however, noted in this manuscript. First, the AAST cholecystitis validation was a retrospective review, and thus requires a prospective evaluation to further validate its findings. PGS was validated in a prospective fashion. Second, the AAST scale starts (Grade 1) at the diagnosis of acute cholecystitis, and thus excludes the comparison of potential other EGS/ACS GB disease (biliary colic/symptomatic cholelithiasis, biliary pancreatitis, etc.). The PGS can be applicable to any operative gallbladder. Third, the majority of the patient population had low AAST grade disease in the validation study: Grade 1 (69.5%), Grade 2 (23.8%), Grade 3 (5.7%), Grade 4 (0.0%), Grade 5 (1.0%).[33] Given this dispersion, Vera et al. (2018) found outcomes such as adverse events, ICU use, and length of stay do not begin to reach significance until Grade 3 or higher. As such, and acknowledged by the authors, the AAST scale does not offer true granularity between

Table 3
Association of perioperative gallbladder grade with other surgical parameters.

	Perioperative Gallbladder Grade					p-value
	1 (n = 60)	2 (n = 90)	3 (n = 102)	4 (n = 28)	5 (n = 37)	(FDR)
N (%)						
Diagnosis						0.0001
Acute cholecystitis	22 (36.7)	35 (38.9)	58 (56.9)	20 (71.4)	31 (83.8)	(0.0002)
Other diagnosis	38 (63.3)	55 (61.1)	44 (43.1)	8 (28.6)	6 (16.2)	
Difficulty of surgery*						0.0001
1	39 (65.0)	23 (25.6)	7 (6.9)	0 (0.0)	0 (0.0)	(0.0002)
2	17 (28.3)	53 (58.9)	36 (35.3)	5 (17.9)	2 (5.4)	
3	3 (5.0)	10 (11.1)	42 (41.2)	5 (17.9)	1 (2.7)	
4	1 (1.7)	4 (4.4)	13 (12.7)	14 (50.0)	12 (32.4)	
5	0 (0.0)	0 (0.0)	4 (3.9)	4 (14.3)	22 (59.5)	
Abnormal anatomy						0.1402
No	56 (93.3)	81 (90.0)	82 (80.4)	24 (85.7)	31 (83.8)	(0.1558)
Yes	4 (6.7)	9 (10.0)	20 (19.6)	4 (14.3)	6 (16.2)	
Partial cholecystectomy						0.0001
No	60 (100.0)	90 (100.0)	98 (96.1)	28 (100.0)	25 (67.6)	(0.0002)
Yes	0 (0.0)	0 (0.0)	4 (3.9)	0 (0.0)	12 (32.4)	
Converted to open						0.0001
No	60 (100.0)	90 (100.0)	102 (100.0)	27 (96.4)	29 (78.4)	(0.0002)
Yes	0 (0.0)	0 (0.0)	0 (0.0)	1 (3.6)	8 (21.6)	
Bile duct leak post-op						0.0013
No	60 (100.0)	89 (98.9)	102 (100.0)	28 (100.0)	32 (88.9)	(0.0022)
Yes	0 (0.0)	1 (1.1)	0 (0.0)	0 (0.0)	4 (11.1)	
Readmission						0.1761
No	57 (95.0)	79 (87.8)	97 (95.1)	27 (96.4)	32 (86.5)	(0.1761)
Yes	3 (5.0)	11 (12.2)	5 (4.9)	1 (3.6)	5 (13.5)	
Mean ± Std. Dev.						
Pre-op total bilirubin	0.6 ± 0.6	0.7 ± 0.7	0.6 ± 0.6	0.6 ± 0.3	1.1 ± 1.2	0.0257 (0.0321)
Pre-op WBC count	8.8 ± 3.0	9.0 ± 3.2	9.6 ± 3.5	10.7 ± 3.8	12.3 ± 4.4	0.0001 (0.0002)
Length post-op stay, days	0.30 ± 0.81	0.14 ± 0.46	0.28 ± 1.18	0.21 ± 0.57	1.54 ± 2.33	0.008 (0.0114)

Fisher's Exact Test was used to test the association of grade with binary (categorical) variables and 1-way ANOVA was used to assess the association of grade with continuous measures. To address multiple testing.
p-values were adjusted using the False Discovery Rate (FDR) procedure.
*Jonckheere-Terpstra test for doubly-ordered categorical data was used.

lower severity of disease, with 90% of the disease demonstrated in Grades 1 and 2.[33] In comparison, our PGS has a more normal distribution amongst grades. Given the limitations of the AAST scale, our future work will be to directly compare our cholecystitis operative grading scale against the AAST scale for anatomic severity.

We acknowledge a few limitations within our study. First, this study was derived from a single institution. While 11 different acute care surgeons participated in the survey, they are all members of the same General and Acute Care Surgery division, and thus they may adopt similar practices over time. Such a bias could have affected responses/grades that could be different at other institutions around the country. Second, difficulty score is largely subjective based on the prior experiences and expertise of the faculty surgeon. As such a LC which is scored as a difficulty score of five to a junior faculty member, may in fact have a difficulty level of a three or four to a more senior attending. However, we want to affirm the fact that difficulty was just one peri-operative variable collected in our analysis. We noted that increasing grade correlates with increased rates of acute cholecystitis, conversion rates, length of operation, and rates of complication. As such, we believe this scale, even without the difficulty correlation, is useful in outcome comparisons. Third, it is possible a rater who assigns a higher score on the cholecystitis grading scale during the "initial view" may assign a higher difficulty score post operatively as a self-fulfilling bias. We note, however, that this direct relationship is consistent with our *a prior* hypothesis. Excellent reliability is seen between cholecystitis grade and the difficulty score when assigned by the same surgeon (ICC 0.8520), and visually this can be seen in Fig. 2c. Nonetheless, in the study design itself, we attempted to minimize

Table 4
Association of perioperative gallbladder grade with difficulty and length of surgery.

Outcome	Perioperative Gallbladder Grade Mean ± SD (95% CI)					p-value
	1 (n = 60)	2 (n = 90)	3 (n = 102)	4 (n = 28)	5 (n = 37)	
Length of Surgery, minutes*	63.31 ± 22.46 (57.51–69.11)	69.75 ± 22.52 (65.03–74.47)	79.76 ± 28.04 (74.25–85.27)	89.10 ± 28.79 (77.94–100.27)	108.13 ± 41.67 (94.24–122.03)	0.0001
Difficulty of Surgery Score**	1.43 ± 0.67 (1.26–1.61)	1.94 ± 0.74 (1.79–2.10)	2.71 ± 0.91 (2.54–2.90)	3.61 ± 0.95 (3.24–3.98)	4.46 ± 0.80 (4.19–4.73)	0.0001

Note. The means presented in this table are the sample means; SD = Standard Deviation; Length of surgery was measured in minutes; Difficulty of surgery was measured on a 5-point Likert-Type scale that ranged from 1 (least difficult) to 5 (most difficult).
The Welch's one-way ANOVA was used to evaluate the omnibus relationship of the Parkland Grading Scale for Cholecystitis with the continuous outcomes of difficulty of surgery ($R^2 = 0.57$) and, in a separate model, length of surgery ($R^2 = 0.19$). Reported p-values were adjusted using the False Discovery Rate procedure.
*The Tukey-Kramer test for all pairwise comparisons (at $p < 0.05$) revealed that Grade 1 was significantly different from Grades 3,4,5; Grade 2 was significantly different from Grades 4,5; Grade 3 was significantly different from Grades 1,5; Grade 4 was significantly different from Grades 1,2; and Grade 5 was significantly different from Grades 1,2,3.
**The Tukey-Kramer test for all pairwise comparisons revealed that each Grade (1–5) was significantly different from each other (at $p < 0.05$).

this potential bias by making the grading scale based on objective anatomical and inflammatory changes to be assigned during the initial view of the gallbladder fossa before the dissection begins. Operative difficulty was graded separately during the post-operative period with no referral or prior stated relationship to the cholecystitis operative grade. In addition, post-hoc we performed a retrospective review of GB "initial view" images stored on our EMR, and these images were given a cholecystitis severity grade by three independent surgeon raters. We found that the independent ratings of the GB images by the 3 raters also agree with the assigned difficulty score of the operation (ICC = 0.7389). In summary, we found that the two independent sets of raters (the operative surgeons as well as the three post-hoc raters) not only both agree on the ratings of the GB images (ICC = 0.8210) but both sets of independent ratings also agree with the assigned difficulty score of the operation. Finally, we want to emphasize this is the validation of an operative classification only. This scale was not developed to be used for those who receive less invasive management (i.e., cholecystostomy tubes).

Future work will focus on improving these aforementioned limitations. First, we require a multi-center trial to assess for both external validation as well as comparison against the recently validated AAST scale for cholecystitis.[33] A multi-center trial would allow us to see if this grading scale can be effectively taught and utilized by other centers around region or nation, as well as effectively power a study to potential detect complication rate differences among five individual grades. Second, we would like to expand our outcome assessment. Rarer complication such as common bile duct injury or death will require a much larger patient sample. Multi-center validation will allow us to power a study for such an assessment.

Conclusion

This study successfully validates our previously developed PGS in a prospective fashion. Increasing grade is significantly associated with increased difficulty of surgery, conversion rates, length of the operation, and incidence of post-operative bile duct leak. We believe we have created an operative grading scale in which higher scores can predict longer, more difficult LCs and higher complication rates. Compared to other scoring systems, our grading scale has now been both prospectively validated and found to have a high inter-rater reliability amongst multiple reviewers in both a prospective and retrospective review. Such a grading scale should and can be used as a step towards changes in outcome comparisons. Specifically, the simplicity of such an intraoperative grading scale validated for peri-operative outcome comparison can be used as a guide to both private and government insurers to potentially broaden the global acceptable risk per surgeon. As such, surgeons and centers who are exposed to more complex, longer operations will no longer be penalized for taking on the extra risks of associated complications.

Meeting submitted and conflicts of interest

This original work was presented by podium at the 31st Annual Meeting of the Eastern Association for the Surgery of Trauma, January 9–13, 2018 in Orlando, Florida and has not been published elsewhere. The authors have no conflicts of interest to disclose.

Authorship contribution

Tarik D Madni MD, MBA – Literature search, study design, data interpretation, writing.

Paul A Nakonezny PhD—Statistical analysis, data interpretation, writing, critical revision.

Evan Barrios BS – Literature search, data collection, study design, writing.

Audra T Clark MD – Literature search, data interpretation, critical revision.

Jonathan B Imran MD –Literature search, data interpretation, critical revision.

Luis Taveras MD – Data collection, critical revision.

Holly B Cunningham MD – Data collection, critical revision.

Herb A. Phelan MD, MSCS - Study design, data interpretation, critical revision.

Joseph P. Minei, MD, MBA – Literature search, study design, data collection, critical revision.

Michael W. Cripps, MD, MSCS - Study design, data interpretation, writing.

Stephen Luk, MD, MD - Study design, data collection, data interpretation, critical revision.

Christian T. Minshall, MD, PhD – Literature search, study design, data collection, critical revision.

Alexander L. Eastman, MD, MPH - Literature search, study design, data collection, critical revision.

Alana Christie, MS - Study design, data analysis, data interpretation, writing, critical revision.

Acknowledgements

Inna Donovan, NP – data collection.

Karen Garofalo, PA – data collection.

Cynthia Chernyakhovsky, PA – data collection.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.amjsurg.2018.08.005.

References

1. Agency for Healthcare Research and Quality - Healthcare Cost and Utilization Project (H-CUP). *Most Frequent Operating Room Procedures Performed in U.S. Hospitals, 2003-2012*; 2014, 2017 . Accessed October 23, 2017.
2. Frazee RC, Elliott VG, Larsen W, et al. Can laparoscopic cholecystectomy be performed with a positive margin at medicaid reimbursement rates? *J Am Coll Surg*. 2014;218(4):546–551.
3. Davis PT, Palilonis MA, Miller PR, Hildreth AN. Time to operating room for cholecystitis decreases with a mature emergency general surgery service. *Am Surg*. 2017;83(1):58–63.
4. Agrawal N, Singh S, Khichy S. Preoperative prediction of difficult laparoscopic cholecystectomy: a scoring method. *Niger J Surg: Off. Publ. Niger. Surg. Res. Soc.* 2015;21(2):130–133.
5. Bourgouin S, Mancini J, Monchal T, et al. How to predict difficult laparoscopic cholecystectomy? Proposal for a simple preoperative scoring system. *Am J Surg*. 2016;212(5):873–881.
6. Donkervoort SC, Kortram K, Dijksman LM, et al. Anticipation of complications after laparoscopic cholecystectomy: prediction of individual outcome. *Surg Endosc*. 2016;30(12):5388–5394.
7. Giger UF, Michel JM, Opitz I, et al. Risk factors for perioperative complications in patients undergoing laparoscopic cholecystectomy: analysis of 22,953 consecutive cases from the Swiss Association of Laparoscopic and Thoracoscopic Surgery database. *J Am Coll Surg*. 2006;203(5):723–728.
8. Gupta N, Ranjan G, Arora MP, et al. Validation of a scoring system to predict difficult laparoscopic cholecystectomy. *Int J Surg*. 2013;11(9):1002–1006.
9. Kim MS, Kwon HJ, Park HW, et al. Preoperative prediction model for conversion of laparoscopic to open cholecystectomy in patient with acute cholecystitis: based on clinical, laboratory, and CT parameters. *J Comput Assist Tomogr*. 2014;38(5):727–732.
10. Lee NW, Collins J, Britt R, Britt LD. Evaluation of preoperative risk factors for converting laparoscopic to open cholecystectomy. *Am Surg*. 2012;78(8): 831–833.
11. Lipman JM, Claridge JA, Haridas M, et al. Preoperative findings predict conversion from laparoscopic to open cholecystectomy. *Surgery*. 2007;142(4): 556–563. discussion 563-555.
12. Lowndes B, Thiels CA, Habermann EB, et al. Impact of patient factors on operative duration during laparoscopic cholecystectomy: evaluation from the

National Surgical Quality Improvement Program database. *Am J Surg*. 2016;212(2):289–296.

13. Massoumi RL, Trevino CM, Webb TP. Postoperative complications of laparoscopic cholecystectomy for acute cholecystitis: a comparison to the ACS-NSQIP risk calculator and the Tokyo guidelines. *World J Surg*. 2017;41(4):935–939.
14. Nidoni R, Udachan TV, Sasnur P, et al. Predicting difficult laparoscopic cholecystectomy based on clinicoradiological assessment. *J Clin Diagn Res: J Clin Diagn Res*. 2015;9(12):Pc09–Pc12.
15. Randhawa JS, Pujahari AK. Preoperative prediction of difficult lap chole: a scoring method. *Indian J Surg*. 2009;71(4):198–201.
16. Rosen M, Brody F, Ponsky J. Predictive factors for conversion of laparoscopic cholecystectomy. *Am J Surg*. 2002;184(3):254–258.
17. Schrenk P, Woisetschlager R, Rieger R, Wayand WU. A diagnostic score to predict the difficulty of a laparoscopic cholecystectomy from preoperative variables. *Surg Endosc*. 1998;12(2):148–150.
18. Soltes M, Radonak J. A risk score to predict the difficulty of elective laparoscopic cholecystectomy. *Wideochirurgia i inne techniki maloinwazyjne = Videosurgery and other miniinvasive techniques*. 2014;9(4):608–612.
19. Sippey M, Grzybowski M, Manwaring ML, et al. Acute cholecystitis: risk factors for conversion to an open procedure. *J Surg Res*. 2015;199(2):357–361.
20. Sugrue M, Sahebally SM, Ansaloni L, Zielinski MD. Grading operative findings at laparoscopic cholecystectomy- a new scoring system. *World J Emerg Surg: WJES*. 2015;10:14.
21. Tominaga GT, Staudenmayer KL, Shafi S, et al. The American Association for the Surgery of Trauma grading scale for 16 emergency general surgery conditions: disease-specific criteria characterizing anatomic severity grading. *J. Trauma Acute Care Surg.* 2016;81(3):593–602.
22. Vivek MA, Augustine AJ, Rao R. A comprehensive predictive scoring method for difficult laparoscopic cholecystectomy. *J Minimal Access Surg*. 2014;10(2): 62–67.
23. Madni TD, Leshikar DE, Minshall CT, et al. The Parkland grading scale for cholecystitis. *Am J Surg*. 2018 Apr;215(4):625–630.
24. Savader SJ, Lillemoe KD, Prescott CA, et al. Laparoscopic cholecystectomy-related bile duct injuries: a health and financial disaster. *Ann Surg*. 1997;225(3):268–273.
25. Richards MK, McAteer JP, Drake FT, et al. A national review of the frequency of minimally invasive surgery among general surgery residents: assessment of ACGME case logs during 2 decades of general surgery resident training. *JAMA surgery*. 2015;150(2):169–172.
26. Shafi S, Aboutanos M, Brown CV, et al. Measuring anatomic severity of disease in emergency general surgery. *J. Trauma Acute Care Surg.* 2014;76(3):884–887.
27. Crandall ML, Agarwal S, Muskat P, et al. Application of a uniform anatomic grading system to measure disease severity in eight emergency general surgical illnesses. *J. Trauma Acute Care Surg.* 2014;77(5):705–708.
28. Baghdadi YMK, Morris DS, Choudhry AJ, et al. Validation of the anatomic severity score developed by the American Association for the Surgery of Trauma in small bowel obstruction. *J Surg Res*. 2016;204(2):428–434.
29. Ray-Zack MD, Hernandez MC, Younis M, et al. Validation of the American Association for the Surgery of Trauma emergency general surgery grade for skin and soft tissue infection. *J. Trauma Acute Care Surg.* 2018;84(6):939–945.
30. Younis M, Hernandez M, Ray-Zack M, et al. Validation of AAST EGS grade for acute pancreatitis. *J Gastrointest Surg: Off. J. Soc. Surg. Aliment. Tract*. 2018;22(3): 430–437.
31. Hernandez MC, Aho JM, Habermann EB, et al. Increased anatomic severity predicts outcomes: validation of the American association for the surgery of trauma's emergency general surgery score in appendicitis. *J. Trauma Acute Care Surg.* 2017;82(1):73–79.
32. Shafi S, Priest EL, Crandall ML, et al. Multicenter validation of American Association for the Surgery of Trauma grading system for acute colonic diverticulitis and its use for emergency general surgery quality improvement program. *J. Trauma Acute Care Surg.* 2016;80(3):405–410. discussion 410-401.
33. Vera K, Pei KY, Schuster KM, Davis KA. Validation of a new American Association for the Surgery of Trauma (AAST) anatomic severity grading system for acute cholecystitis. *J. Trauma Acute Care Surg.* 2018;84(4):650–654.