American Journal of Emergency Medicine 37 (2019) 61–66

Contents lists available at ScienceDirect

American Journal of Emergency Medicine

journal homepage: www.elsevier.com/locate/ajem

Original Contribution

Derivation and validation of a practical Bedside Score for the diagnosis of cholecystitis☆

D. Dante Yeh [a,*], Yuchiao Chang [b], Maryam Bita Tabrizi [c], Liyang Yu [b], Catrina Cropano [d], Peter Fagenholz [d], David R. King [d], Haytham M.A. Kaafarani [d], Marc de Moya [d], George Velmahos [d]

[a] Ryder Trauma Center, DeWitt Daughtry Family Department of Surgery, University of Miami Miller School of Medicine, United States
[b] Massachusetts General Hospital, Department of Medicine, United States
[c] North Shore Medical Center, Department of Surgery, United States
[d] Massachusetts General Hospital, Department of Surgery, Division of Trauma, Emergency Surgery, and Surgical Critical Care, United States

ARTICLE INFO

Article history:
Received 23 August 2017
Received in revised form 20 April 2018
Accepted 23 April 2018

Keywords:
Cholecystitis
Ultrasound
Murphy's sign
Cholelithiasis
Gallbladder

ABSTRACT

Objective: We sought to develop a practical Bedside Score for the diagnosis of cholecystitis and test its accuracy against the Tokyo Guidelines (TG13).
Methods: We conducted a retrospective study of 438 patients undergoing urban, academic Emergency Department (ED) evaluation of RUQ pain. Symptoms, physical signs, ultrasound signs, and labs were scoring system candidates. A random split-sample approach was used to develop and validate a new clinical score. Multivariable regression analysis using development data was conducted to identify predictors of cholecystitis. Cutoff values were chosen to ensure positive/negative predictive values (PPV, NPV) of at least 0.95. The score was externally validated in 80 patients at a different hospital undergoing RUQ pain evaluation.
Results: 230 patients (53%) had cholecystitis. Five variables predicted cholecystitis and were included in the scores: gallstones, gallbladder thickening, clinical or ultrasonographic Murphy's sign, RUQ tenderness, and post-prandial symptoms. A clinical prediction score was developed. When dichotomized at 4, overall accuracy for acute cholecystitis was 90% for the development cohort, 82% and 86% for the internal and external validation cohorts; TG13 accuracy was 62%–79%.
Conclusions: A clinical prediction score for cholecystitis demonstrates accuracy equivalent to TG13. Use of this score may streamline work-up by decreasing the need for comprehensive ultrasound evaluation and CRP measurement and may shorten ED length of stay.

1. Introduction

1.1. Background

Cholelithiasis is estimated to affect approximately 20 million Americans, or roughly 7% of men and 11% of women [1,2]. While the majority of patients with cholelithiasis are asymptomatic, biliary colic develops at a rate of 1%–2% annually [3,4]. If untreated, cholecystitis eventually develops in about 20% of symptomatic patients [2]. Hepatobiliary scintigraphy (HIDA scan) has been reported to be superior to ultrasonography for the diagnosis of acute cholecystitis [5], but is rarely used as the first diagnostic test for suspected acute cholecystitis because of the expense, time required, and limited availability during nights and weekends. Thus, ultrasound (US) is the most common first imaging study performed in the evaluation of possible acute cholecystitis. However, right upper quadrant (RUQ) pain is a common symptom of many other diseases such as hepatitis, peptic ulcer disease, liver abscess, and myocardial infarction [6], and the accuracy of US alone has been debated [7-12], with a sensitivity ranging from as low as 48% to as high as 94% [13,14]. High-resolution computed tomography (CT) has also been reported to have high accuracy in diagnosing cholecystitis [15].

1.2. Importance

Clinical diagnosis of acute cholecystitis is sometimes difficult, and no single sign or symptom has sufficient positive predictive value (PPV) or negative predictive value (NPV) to confidently rule in or rule out the disease [16,17]. Hence, diagnosis is usually achieved by assessing a constellation of signs and symptoms such as RUQ pain or tenderness, post-prandial nausea or vomiting, anorexia, fever, leukocytosis, and Murphy's sign (inspiratory arrest during palpation of the gallbladder). Preoperative diagnosis using defined criteria have been shown to be

☆ None of the authors report any conflicts of interest.

* Corresponding author at: Ryder Trauma Center, DeWitt Daughtry Family Department of Surgery, University of Miami Miller School of Medicine, PO Box 016960 D40, Miami, FL 33101, United States.
E-mail address: dxy154@miami.edu (D.D. Yeh).

https://doi.org/10.1016/j.ajem.2018.04.051
0735-6757/

inaccurate in almost one half of cases of acute cholecystitis confirmed by intraoperative findings and expert pathological diagnosis [18].

In 2007, the Tokyo Guidelines (TG07) were published as the world's first guidelines to include diagnostic criteria for acute cholecystitis [19,20]. These guidelines were derived from best-available evidence and expert consensus and required the combination of local signs (Murphy's sign or RUQ mass/pain/tenderness), systemic signs (fever, elevated C-reactive protein (CRP), or leukocytosis), and imaging findings characteristic of acute cholecystitis (enlarged gallbladder, thickening of the gallbladder wall of 5 mm or greater, pericholecystic fluid, ultrasonographic Murphy's sign, gallbladder stones, and debris echo). A validation study of the TG07 reported sensitivity of 92.1% and specificity of 93.3% [21]. However, that validation was performed retrospectively in an exclusively Japanese population. Recent modifications to the TG07 led to development of TG13, with a reported improvement in specificity to 96.9% and an overall accuracy of up to 94% [21]. (Table 1) However, full implementation of the TG13 requires routine measurement of CRP, which may not be a part of usual clinical practice for patients presenting with symptoms suggestive of cholecystitis. Additionally, although it is not specifically mentioned in the TG13, it is implied that formal comprehensive ultrasonographic examination is required to accurately assess for all imaging findings characteristic of acute cholecystitis.

1.3. Goals of this investigation

We sought to validate the TG13 criteria in our cohort of patients presenting to the Emergency Department (ED) with RUQ pain and possible cholecystitis. Furthermore, we also sought to develop a practical score which could be applied at the bedside by clinicians to aid in the diagnosis of cholecystitis.

2. Methods

2.1. Study design and setting

In the development phase, we performed a retrospective study at a tertiary-care urban, academic hospital. For external validation, a separate cohort of consecutive patients presenting to a different hospital (suburban community hospital setting) and undergoing RUQ ultrasound in the ED for possible acute cholecystitis was analyzed. This study was approved by the Institutional Review Board at both hospitals.

2.2. Selection of participants

Over the study period (July 2010 to January 2014), patients undergoing cholecystectomy by the Acute Care Surgery service for presumed acute cholecystitis (as determined by the attending surgeon) were eligible for inclusion. The preoperative criteria for presumed acute cholecystitis were a clinical symptoms including RUQ abdominal tenderness, fever, leukocytosis, nausea/vomiting, post-prandial symptoms, and Murphy's sign combined with a radiologic imaging study (US or CT) showing features characteristic of acute inflammation such as gallbladder (GB) distention, GB wall thickening >3 mm, pericholecystic fluid, impacted stone at the neck of the GB, and a sonographic Murphy's sign [19,22]. These criteria for diagnosing acute cholecystitis comprise the constellation of signs and symptoms typically associated with this condition. Our division does not have a formal definition and the decision to operate was based on the discretion of the attending surgeon. The presence of cholelithiasis was not considered mandatory for the diagnosis of presumed acute cholecystitis. The TG13 criteria were applied using all available information obtained in the usual course of clinical care. Exclusion criteria included: transfer from another hospital or direct hospital admission (i.e. bypassing the ED), prior cholecystotomy tube, and prior endoscopic retrograde cholangiopancreatography (ERCP) with sphincterotomy. The control group was comprised of a cohort of 99 patients (May 2010 to August 2010) undergoing RUQ US evaluation for RUQ abdominal pain, but ultimately discharged from the ED or admitted to the hospital with a non-biliary diagnosis. The external validation cohort consisted of all consecutive patients undergoing Emergency Department RUQ ultrasound for RUQ pain at another hospital between January 2016 and March 2016. Both derivation and external validation cohorts were identified retrospectively. Data collected included age, gender, clinical findings, radiographic, operative, and histopathology reports.

Table 1
Summary of the Tokyo Guidelines 2013 (TG13).

A. Local signs of inflammation
(1) Murphy's sign
(2) RUQ mass/pain/tenderness
B. Systemic signs of inflammation
(1) Fever
(2) Elevated CRP
(3) Elevated WBC
C. Imaging findings characteristic of acute cholecystitis.

CRP = C-reactive protein; RUQ = right upper quadrant; WBC = white blood cell.
Suspected diagnosis: at least one item in A and at least one item in B.
Definite diagnosis: at least one item from A, B and C.

2.3. Outcomes

For patients who underwent cholecystectomy, the final pathologist's impression was used to establish the diagnosis of cholecystitis. For all others, the electronic medical record was reviewed to ensure that the patient did not undergo cholecystectomy for cholecystitis within 30 days after the index ED visit.

2.4. Statistical analysis

We used a random split-sample approach to develop the new clinical prediction score for cholecystitis: 2/3 of the sample was used to develop the scoring system and the remaining 1/3 for internal validation. We pre-selected 13 candidate predictor variables from patient demographics (gender), clinical symptoms (nausea/vomiting, post-prandial symptoms, RUQ abdominal tenderness, Murphy's sign, fever), labs (WBC > 11), and features from ultrasound radiologic imaging study (GB wall thickening >3 mm, pericholecystic fluid, sonographic Murphy's sign, impacted stone at the neck of the GB, GB distention, cholelithiasis) to use for model development in the derivation cohort. To increase model stability, we combined the clinical or sonographic Murphy's sign into one variable, which reduced our candidate predictors from 13 to 12. We excluded predictors not significant at 0.1 level after the initial univariate analysis using chi-square tests. To avoid the problems with the traditional variable selection methods [23], we constructed 1000 bootstrap samples based on the development sample. For each sample, we used a stepwise logistic regression model to determine predictors significant at 0.05 level. Variables consistently chosen in >50% of the bootstrap samples were included in the final model. A prediction score was then created based on the estimates from the final multivariable regression model. To simplify the scoring, we gave three points to predictors with OR ≥ 30, two points to predictors with OR 10–29, and 1 point to variables with OR < 10. Subsequently, the scoring system was tested in the internal and external validation samples. Traditional measures of model fit were calculated for both the development sample and the validation samples, including c statistic for discrimination and the Hosmer-Lemeshow goodness-of-fit statistic for calibration. Cutoff values were chosen to group patients into low, moderate, and high risk categories to ensure the low risk category has a negative predictive value of at least 0.95 and the high risk category has a positive predictive value of at least 0.95 based on the data from the derivation cohort. Additionally, test characteristics

(accuracy, positive predictive value [PPV], and negative predictive value [NPV]) were calculated for the dichotomized score to compare the performance with individual signs and the TG13 criteria using McNemar's tests. Because the original Tokyo Guidelines 2013 document did not clarify the exact threshold to define "elevated WBC" in the systemic signs of inflammation section, we used the definition of "evidence of inflammatory response" by the same working group in a related document: WBC < 4 or >10 [21,24]. The threshold for significance was set at $p < 0.05$. All statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC).

3. Results

The study sample (derivation and internal validation) included 230 cholecystitis patients (53%) and 208 controls. In the two-thirds derivation sample, there were 153 cholecystitis patients and 138 controls. The remaining one-third of the patients was used for internal validation (77 with cholecystitis and 70 without cholecystitis). In the external validation group, the incidence of acute cholecystitis was 18%, as confirmed by surgical pathology. None of the remaining patients were subsequently diagnosed with acute cholecystitis after ED discharge.

Table 2
Candidate Signs and Symptoms as Predictors for Cholecystitis. The original cohort was split into "development" and "internal validation" cohorts.

	Development			Internal validation			External validation		
	N	Cholecystitis (%)	p value	N	Cholecystitis (%)	p value	N	Cholecystitis (%)	p value
All	291	153 (53)		147	77 (52)		80	14 (18)	
Gender			0.11			0.42			0.69
Female	164	93 (57)		89	49 (55)		55	9 (16)	
Male	127	60 (47)		58	28 (48)		25	5 (20)	
Nausea/vomiting			0.006			0.003			0.29
No	99	41 (41)		40	13 (33)		30	7 (23)	
Yes	192	112 (58)		107	64 (60)		50	7 (14)	
Post-prandial symptoms			<0.0001			0.061			0.0001
No	203	89 (44)		107	51 (48)		57	4 (7)	
Yes	88	64 (73)		40	26 (65)		23	10 (43)	
RUQ tenderness			<0.0001			<0.0001			0.20
No	71	9 (13)		40	9 (23)		29	3 (10)	
Yes	220	144 (65)		107	68 (64)		51	11 (22)	
Murphy's sign			<0.0001			<0.0001			0.004
No	185	61 (33)		103	36 (35)		70	9 (13)	
Yes	106	92 (87)		44	41 (93)		10	5 (50)	
WBC > 11			<0.0001			<0.0001			0.073
No	175	71 (41)		87	33 (38)		51	6 (12)	
Yes	116	82 (71)		60	44 (73)		29	8 (28)	
Fever			0.84			0.53			0.96
No	248	131 (53)		132	68 (52)		74	13 (18)	
Yes	43	22 (51)		15	9 (60)		6	1 (17)	
US GB wall thickening			<0.0001			<0.0001			0.030
No	203	74 (36)		112	47 (42)		67	9 (13)	
Yes	88	79 (90)		35	30 (86)		13	5 (38)	
US pericholecystic fluid			<0.0001			0.002			0.079
No	236	102 (43)		130	62 (48)		76	12 (16)	
Yes	55	51 (93)		17	15 (88)		4	2 (50)	
US Murphy's sign			<0.0001			<0.0001			0.0001
No	222	85 (38)		115	47 (41)		75	10 (13)	
Yes	69	68 (99)		32	30 (94)		5	4 (80)	
US distended GB			<0.0001			0.001			0.22
No	240	113 (47)		126	59 (47)		78	13 (17)	
Yes	51	40 (78)		21	18 (86)		2	1 (50)	
US incarcerated gallstone			<0.0001			0.002			<0.0001
No	250	114 (46)		130	62 (48)		76	10 (13)	
Yes	41	39 (95)		17	15 (88)		4	4 (100)	
Cholelithiasis			<0.0001			<0.0001			<0.0001
No	113	11 (10)		57	7 (12)		53	0 (0)	
Yes	178	142 (80)		90	70 (78)		27	14 (52)	
Murphy's sign (clinical or US)			<0.0001			<0.0001			0.002
No	162	39 (24)		93	27 (29)		67	7 (10)	
Yes	129	114 (88)		54	50 (93)		13	7 (54)	
TG local signs			<0.0001			<0.0001			0.074
No	68	7 (10)		38	7 (18)		28	2 (7)	
Yes	223	146 (65)		109	70 (64)		52	12 (23)	
TG systemic signs			0.0008			0.001			0.63
No	116	50 (43)		68	26 (38)		39	6 (15)	
Yes	175	103 (59)		79	51 (65)		41	8 (20)	
TG imaging findings			<0.0001			<0.0001			0.0006
No	168	41 (24)		99	35 (35)		62	6 (10)	
Yes	122	112 (92)		48	42 (88)		18	8 (44)	
TG suspected diagnosis			<0.0001			<0.0001			0.20
No	151	54 (36)		86	32 (37)		52	7 (13)	
Yes	140	99 (71)		61	45 (74)		28	7 (25)	
TG definite diagnosis			<0.0001			<0.0001			0.18
No	207	74 (36)		120	53 (44)		71	11 (15)	
Yes	84	79 (94)		27	24 (89)		9	3 (33)	

GB = gallbladder; RUQ = right upper quadrant; TG = Tokyo Guidelines; US = ultrasound; WBC = white blood cell.

Table 3
Bedside acute cholecystitis score.

	Points
Post-prandial symptoms	1
RUQ tenderness	1
Murphy's sign	2
Gallbladder thickening	2
Gallstones	3
Total score	9

The candidate predictors initially assessed for predictive value along with the Tokyo Guideline components are shown in Table 2. Ten predictors were found to be significant at 0.1 level based on the development sample. In the bootstrap samples, five predictors were found to be consistently significant at 0.05 level in the multiple logistic regression models. Their importance, according to the regression coefficients, was determined as follows: presence of gallstones, Murphy's sign (clinical or sonographic), GB thickening, clinical RUQ tenderness, and post-prandial symptoms. Based on the magnitude of the independent effect, we devised a practical diagnostic score to aid in the diagnosis of cholecystitis. Variables with odds ratio < 10 were given 1 point, odds ratios 10–29 were given 2 points, and odds ratios ≥ 30 were given 3 points (Table 3). The score had a c statistic of 0.96 from the development sample and decreased to 0.91 for the internal validation sample but performed quite well in the external validation sample ($c = 0.96$). The Hosmer-Lemeshow goodness-of-fit statistic did not show evidence of poor calibration.

Table 4 shows the prevalence rate of cholecystitis in the original derivation cohort, internal validation cohort, and external cohort based on the score risk category. In the development cohort, the prevalence rate was <5% for the low risk category and around 95% for the high risk category. When validated in the internal cohort, the prevalence rate was higher than expected (12.1%) for the low risk category; however the prevalence rate was similar to the development cohort for the high risk category (91.4%). When validated in the external cohort, the prevalence rate was <2% for the low risk category and at least 75% for high risk category. When dichotomized at 4, the PPV deteriorated in the external cohort while the NPV improved. The overall accuracy was 90% in the development cohort, 82% in the internal validation cohort, and 86% in the external validation cohort.

The performance of the TG13 is also included in Table 4. When dichotomized at a score of <4 vs. ≥4, the new score was significantly more accurate than the TG13 definite cholecystitis in the derivation cohort (90% vs. 73%, $p < 0.0001$), and internal validation cohort (82% vs. 62%, $p = 0.0001$), but the differences did not reach statistical significance in the external validation cohort (86% vs. 79%, $p = 0.20$). Combining the development cohort and the internal validation cohort, overall accuracy for acute cholecystitis was 87.7% in the original cohort. By comparison, the accuracy of the TG13 diagnostic criteria for suspected and definite diagnosis of acute cholecystitis was 67.4% and 69.2%, respectively. For positive predictive value, the 95% confidence interval from the new score overlapped with the 95% confidence interval from the TG13 diagnostic criteria for definite diagnosis of acute cholecystitis. On the other hand, the negative predictive value from the new score (96%, 95%CI: 91%–99%) was higher than the TG13 diagnostic criteria for definite diagnosis of acute cholecystitis (64%, 95%CI: 57%–71%) in our cohorts.

4. Discussion

In this study, we report that the 2013 Tokyo Guidelines were not accurate in diagnosing cholecystitis in our patient cohort presenting to the ED with RUQ pain. While local signs and imaging findings were

Table 4
Performance characteristics.

	D Cohort	IV Cohort	EV Cohort
Bedside Score			
Cholecystitis prevalence rate (%)			
Low: <4	5/120 (4.2)	7/58 (12.1)	1/57 (1.8)
Moderate: 4–5	23/40 (57.5)	17/31 (54.8)	4/11 (36.4)
High: ≥6	125/131 (95.4)	53/58 (91.4)	9/12 (75.0)
PPV for ≥4 (95%CI)	87% (81%–91%)	79% (69%–87%)	57% (34%–77%)
NPV for <4 (95%CI)	96% (91%–99%)	88% (77%–95%)	98% (91%–100%)
Accuracy (95%CI)	90% (86%–94%)	82% (75%–88%)	86% (77%–93%)
TG13 local signs			
PPV	65% (59%–72%)	64% (54%–73%)	23% (13%–37%)
NPV	90% (80%–96%)	82% (66%–92%)	93% (76%–99%)
Accuracy	71% (66%–76%)	69% (61%–76%)	48% (36%–59%)
TG13 systemic signs			
PPV	59% (51%–66%)	65% (53%–75%)	20% (9%–35%)
NPV	57% (47%–66%)	62% (49%–73%)	85% (69%–94%)
Accuracy	58% (52%–64%)	63% (55%–71%)	51% (40%–63%)
TG13 imaging signs			
PPV	59% (51%–66%)	65% (53%–75%)	20% (9%–35%)
NPV	57% (47%–66%)	62% (49%–73%)	85% (69%–94%)
Accuracy	82% (77%–86%)	72% (64%–79%)	80% (70%–88%)
TG13 suspected cholecystitis			
PPV	71% (62%–78%)	74% (61%–84%)	25% (11%–45%)
NPV	64% (56%–72%)	63% (52%–73%)	87% (74%–94%)
Accuracy	67% (62%–73%)	67% (59%–75%)	65% (54%–75%)
TG13 definite cholecystitis			
PPV	94% (87%–98%)	89% (71%–98%)	33% (7%–70%)
NPV	64% (57%–71%)	56% (46%–65%)	85% (74%–92%)
Accuracy	73% (67%–78%)	62% (54%–70%)	79% (68%–87%)

D = development; EV = external validation; IV = internal validation.
[a] Comparing accuracy of each score in the EV Cohort to TG13 definite cholecystitis.

common in those with cholecystitis (94%), systemic signs were present in about two thirds of patients (67%) and thus TG13 Suspected Diagnosis could only be applied to 63% and TG13 Definite Diagnosis to only 45% of those with cholecystitis. Like other investigators, we found that systemic symptoms such as fever and leukocytosis/leukopenia occurred infrequently in our patient population [25,26], and this was the main reason for the poor sensitivity of the TG13. TG13 local symptoms alone, however, cannot be used to diagnosis cholecystitis, as more than half of patients without cholecystitis in our study also had TG13 local symptoms. Thus, it is poorly specific.

CT scanning is increasingly utilized for the evaluation of abdominal pain, especially in cases with atypical symptoms where the diagnosis is obscure. Despite a well-recognized decreased ability to detect non-calcified gallstones, the sensitivity of CT to detect signs of cholecystitis (such as fat stranding, pericholecystic fluid, and gallbladder wall thickening) is reportedly superior to US [9,27,28]. Although some patients also underwent CT scanning as part of the evaluation of RUQ pain, we chose not to include CT findings in our score, so as not to encourage potentially unnecessary irradiation.

The accuracy of the Murphy's sign has been debated, with some authors reporting high sensitivity and predictive value [26] and others reporting high specificity, but low sensitivity [29]. Likewise, the sonographic Murphy's sign has been reported to have widely varying accuracy [30-32]. At present, the strongest conclusion that can be drawn is that the presence of a Murphy's sign is strongly suggestive (but not diagnostic) of acute cholecystitis, but its absence does not exclude the disease. For example, in cases of gangrenous cholecystitis, Murphy's sign is frequently absent, possibly due to denervation of the gallbladder wall secondary to ischemic changes [6].

It has previously been reported that ultrasonography can be performed by emergency physicians (i.e. bedside ultrasound) to evaluate for cholelithiasis and acute cholecystitis [33]. In a prospective study of bedside abdominal ultrasonography, Rosen et al. assessed agreement between bedside (i.e. limited) US and comprehensive ultrasound (performed in the Radiology department) and reported high accuracy for the detection of cholelithiasis (92% sensitivity, 78% specificity) and acceptable accuracy for the detection of cholecystitis (91% sensitivity, 66% specificity) [34]. With our clinical prediction score, history and physical exam elements garner a maximum score of 4, placing the patient in the moderate risk category (Table 4). It is conceivable that utilization of this score (using limited bedside ultrasound rather than comprehensive ultrasound assessment) may result in decreased resource utilization, faster diagnosis and treatment, and shorter ED length of stay. Before widespread adoption, utilization of this score combined with usual care (clinical judgment and unstructured work-up) must be compared to usual care alone to determine the added benefit. Ideally, this would proceed in a prospective fashion, with clinicians calculating the bedside cholecystitis score prior to or blinded to the results of comprehensive ultrasound assessment.

Using a constellation of signs and symptoms, we devised a new practical scoring system, similar to the Alvarado appendicitis score, which was equivalent in accuracy to the TG13 for the diagnosis of cholecystitis. This score, if validated, may be used to aid in the diagnosis of cholecystitis. Clinical scoring systems are widely used in the ED for other disease such as heparin-induced thrombocytopenia [35], deep vein thrombosis [36], and lower extremity injuries [37]. For diseases without a pathognomonic sign, symptom, or laboratory value, clinical prediction rules (CPR) may assist inexperienced practitioners by focusing their attention to key findings and assigning relative weights [38]. By linking treatment decisions to probability estimates, a CPR is then referred to as a clinical decision rule (CDR) and can potentially increase quality of care through more efficient or timely resource utilization. For example, implementation of the Ottawa Ankle Rules led to a relative reduction in ankle and foot radiography (28% and 14%, respectively), shorter Emergency Department length of stay (80 min vs 116 min, $p < 0.0001$), and lower total medical costs ($62 vs $173, $p < 0.001$) with no difference in satisfaction with emergency physician care or subsequent radiography [39]. It is important to emphasize that CPRs and CDRs are meant to supplement (not replace) clinical judgment and to aid in the training of novices. In this study, we present the derivation, internal validation, and preliminary external validation process of developing a novel CPR.

4.1. Limitations

There are several limitations of our study which must be acknowledged. Firstly, the development (and internal validation) cohort consisted of patients all treated at a single urban academic hospital, and thus the results may not be generalized to other settings. The second stage of validation occurred at only a single external site and had a relatively small sample. Therefore, additional external validation using large sample sizes will help further improve accuracy by informing refinement of the relative scoring weight components and adjustment of cut score thresholds if necessary. Secondly, we do not routinely measure C-reactive protein in all patients presenting with RUQ pain and therefore our calculation of the Tokyo Guidelines 2013 score may be flawed. This is a limitation of the retrospective study design and the local practice. Approximately 43% (116 of 291) of all the patients had “negative” TG13 systemic signs, and it is possible that some patients who were afebrile and had a normal WBC would have had elevated CRP and thus would have scored “positive” for cholecystitis according to TG13. However, this limitation also illustrates the limitation of using the TG13 score in actual clinical practice. TG13 criteria require the combination of local signs, systemic signs, and imaging signs to meet the criteria for definite cholecystitis. To further explore the extent to which missing the CRP may have affected the accuracy of the TG13, we searched for all cases where an elevated CRP (had it been drawn) would have changed the TG13 diagnosis from suspected to definite cholecystitis, i.e. positive local signs, negative systemic signs, and positive imaging signs. It is in these cases that a missing CRP may have resulted in a falsely negative TG13 assessment. In the entire cohort (derivation, internal validation, and external validation), there were only 3 cases identified, one in each subgroup. Changing all three cases from negative to positive TG13 definite cholecystitis did not significantly affect our results and conclusions. Thirdly, while some patients in the control group (i.e. non-cholecystitis) were admitted and confirmed to have non-biliary sources of their RUQ pain, there were some patients who were discharged from the ED. While we can confirm that they did not return to our hospital (or any other hospital within our network) within 30 days, we cannot rule out the possibility that they presented to an outside hospital with cholecystitis. Unfortunately, we are not aware of any high-quality studies reporting the incidence of return to a (different) Emergency Department after a misdiagnosis of cholecystitis. However, it is reassuring to note that we did not discover even one instance of a missed diagnosis of acute cholecystitis which returned to our hospital system. Our control group was drawn from a much narrower time frame than from our cholecystitis cohort for convenience, as the majority of patients undergoing RUQ ultrasound did not have cholecystitis. While this may theoretically affect the development of the cholecystitis score, we do not believe the clinical characteristics and demographics of the patients changed significantly over a period of four years, and thus we think it is unlikely to have significantly influenced the derivation and validation of the score. Fourthly, the score was constructed from a derivation cohort. Although we validated the score internally using a split-sample technique and externally in a separate patient cohort, additional external validation, preferably at multiple institutions by different investigators in different patient populations with various ultrasound machines, is required before widespread adoption. Importantly, CRP must be drawn on every patient to be certain that the TG13 is being properly applied as conceived. Thus, the accuracy of the Bedside Score compared to TG13 and usual care must be placed into the context of the cost (for labs and comprehensive ultrasound assessment) and time spent in the Emergency Department. Despite these limitations,

we believe our study is valuable because it demonstrates that cholecystitis may be diagnosed in the Emergency Department with accuracy similar to TG13 using a combination of history, physical exam, and limited ultrasound without requiring laboratory testing. The high negative predictive value in the external validation cohort (98%) for a Bedside Score < 4 suggests that its value may lie in ruling out the disease. Thus, this score is potentially most useful in resource-scarce settings without comprehensive laboratory or ultrasound support.

5. Conclusions

In summary, we have developed a new practical clinical prediction score to aid in the diagnosis of cholecystitis in patients presenting to the ED with RUQ pain using limited ultrasound and without requiring laboratory testing. Our Bedside Score demonstrates accuracy equivalent to the TG13 and should be further externally validated prior to widespread adoption.

Financial support

None.

References

[1] Duncan CB, Riall TS. Evidence-based current surgical practice: calculous gallbladder disease. J Gastrointest Surg 2012;16(11):2011–25.
[2] Strasberg SM. Clinical practice. Acute calculous cholecystitis. N Engl J Med 2008;358(26):2804–11.
[3] Friedman GD. Natural history of asymptomatic and symptomatic gallstones. Am J Surg 1993;165(4):399–404.
[4] Elwood DR. Cholecystitis. Surg Clin North Am 2008;88(6):1241–52 [viii].
[5] Kiewiet JJ, et al. A systematic review and meta-analysis of diagnostic performance of imaging in acute cholecystitis. Radiology 2012;264(3):708–20.
[6] Hanbidge AE, et al. From the RSNA refresher courses: imaging evaluation for acute pain in the right upper quadrant. Radiographics 2004;24(4):1117–35.
[7] Alobaidi M, et al. Current trends in imaging evaluation of acute cholecystitis. Emerg Radiol 2004;10(5):256–8.
[8] Fink-Bennett D, et al. The sensitivity of hepatobiliary imaging and real-time ultrasonography in the detection of acute cholecystitis. Arch Surg 1985;120(8):904–6.
[9] De Vargas Macciucca M, et al. Ultrasonographic and spiral CT evaluation of simple and complicated acute cholecystitis: diagnostic protocol assessment based on personal experience and review of the literature. Radiol Med 2006;111(2):167–80.
[10] Chatziioannou SN, et al. Hepatobiliary scintigraphy is superior to abdominal ultrasonography in suspected acute cholecystitis. Surgery 2000;127(6):609–13.
[11] Golea A, Badea R, Suteu T. Role of ultrasonography for acute cholecystic conditions in the emergency room. Med Ultrason 2010;12(4):271–9.
[12] Bingener J, et al. Does the correlation of acute cholecystitis on ultrasound and at surgery reflect a mirror image? Am J Surg 2004;188(6):703–7.
[13] Shea JA, et al. Revised estimates of diagnostic test sensitivity and specificity in suspected biliary tract disease. Arch Intern Med 1994;154(22):2573–81.
[14] Kalimi R, et al. Diagnosis of acute cholecystitis: sensitivity of sonography, cholescintigraphy, and combined sonography-cholescintigraphy. J Am Coll Surg 2001;193(6):609–13.
[15] Fagenholz PJ, et al. Computed tomography is more sensitive than ultrasound for the diagnosis of acute cholecystitis. Surg Infect 2015;16(5):509–12.
[16] Berger MY, et al. Abdominal symptoms: do they predict gallstones? A systematic review. Scand J Gastroenterol 2000;35(1):70–6.
[17] Trowbridge RL, Rutkowski NK, Shojania KG. Does this patient have acute cholecystitis? JAMA 2003;289(1):80–6.
[18] Fitzgibbons Jr RJ, et al. Acute cholecystitis. Does the clinical diagnosis correlate with the pathological diagnosis? Surg Endosc 1996;10(12):1180–4.
[19] Hirota M, et al. Diagnostic criteria and severity assessment of acute cholecystitis: Tokyo Guidelines. J Hepato-Biliary-Pancreat Surg 2007;14(1):78–82.
[20] Takada T, et al. Background: Tokyo Guidelines for the management of acute cholangitis and cholecystitis. J Hepato-Biliary-Pancreat Surg 2007;14(1):1–10.
[21] Yokoe M, et al. New diagnostic criteria and severity assessment of acute cholecystitis in revised Tokyo Guidelines. J Hepatobiliary Pancreat Sci 2012;19(5):578–85.
[22] Raghavendra BN, et al. Acute cholecystitis: sonographic-pathologic analysis. AJR Am J Roentgenol 1981;137(2):327–32.
[23] Harrell F. Regression Modeling Strategies. New York: Springer-Verlag; 2001.
[24] Kiriyama S, et al. New diagnostic criteria and severity assessment of acute cholangitis in revised Tokyo Guidelines. J Hepatobiliary Pancreat Sci 2012;19(5):548–56.
[25] Gruber PJ, et al. Presence of fever and leukocytosis in acute cholecystitis. Ann Emerg Med 1996;28(3):273–7.
[26] Singer AJ, et al. Correlation among clinical, laboratory, and hepatobiliary scanning findings in patients with suspected acute cholecystitis. Ann Emerg Med 1996;28(3):267–72.
[27] Fidler J, Paulson EK, Layfield L. CT evaluation of acute cholecystitis: findings and usefulness in diagnosis. AJR Am J Roentgenol 1996;166(5):1085–8.
[28] Bennett GL, Balthazar EJ. Ultrasound and CT evaluation of emergent gallbladder pathology. Radiol Clin N Am 2003;41(6):1203–16.
[29] Yokoe M, et al. TG13 diagnostic criteria and severity grading of acute cholecystitis (with videos). J Hepatobiliary Pancreat Sci 2013;20(1):35–46.
[30] Bree RL. Further observations on the usefulness of the sonographic Murphy sign in the evaluation of suspected acute cholecystitis. J Clin Ultrasound 1995;23(3):169–72.
[31] Soyer P, et al. Color velocity imaging and power Doppler sonography of the gallbladder wall: a new look at sonographic diagnosis of acute cholecystitis. AJR Am J Roentgenol 1998;171(1):183–8.
[32] Ralls PW, et al. Prospective evaluation of the sonographic Murphy sign in suspected acute cholecystitis. J Clin Ultrasound 1982;10(3):113–5.
[33] Summers SM, et al. A prospective evaluation of emergency department bedside ultrasonography for the detection of acute cholecystitis. Ann Emerg Med 2010;56(2):114–22.
[34] Rosen CL, et al. Ultrasonography by emergency physicians in patients with suspected cholecystitis. Am J Emerg Med 2001;19(1):32–6.
[35] Lo GK, et al. Evaluation of pretest clinical score (4 T's) for the diagnosis of heparin-induced thrombocytopenia in two clinical settings. J Thromb Haemost 2006;4(4):759–65.
[36] Wells PS, et al. Value of assessment of pretest probability of deep-vein thrombosis in clinical management. Lancet 1997;350(9094):1795–8.
[37] Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med 2006;144(3):201–9.
[38] Wasson JH, et al. Clinical prediction rules. Applications and methodological standards. N Engl J Med 1985;313(13):793–9.
[39] Stiell IG, et al. Implementation of the Ottawa ankle rules. JAMA 1994;271(11):827–32.