ORIGINAL SCIENTIFIC REPORT

The Yield of Fever, Inflammatory Markers and Ultrasound in the Diagnosis of Acute Cholecystitis: A Validation of the 2013 Tokyo Guidelines

K. Naidu[1] · E. Beenen[1] · S. Gananadha[1] · C. Mosse[1]

Abstract

Background Each year, 1–4 % of people with known gallstones become symptomatic, either presenting with biliary colic or as acute cholecystitis. The distinction between both diagnoses remains challenging. To aid the proper diagnosis, the revised 2013 Tokyo Guidelines (TG 2013) were proposed with a self-acclaimed diagnostic accuracy of over 90 %. However, this accuracy has not been verified by others so far.

Objective To determine the accuracy of the TG 2013 guidelines in the diagnosis of acute cholecystitis both in its single components of fever, inflammatory markers and US features and of the combined application of the TG 2013 guidelines as a whole.

Methods A 5-year retrospective analysis equal to the TG 2013 validation process of all emergency cholecystectomies for acute cholecystitis or persistent biliary pain with an ultrasound performed during the same admission. Acute cholecystitis at histology was the golden standard.

Results Inclusion criteria were met by 169 patients with a prevalence of acute cholecystitis of 52.7 %. The individual features of fever, gallbladder wall thickening and probe tenderness were not significant in univariate analysis. In multivariate analysis only, neutrophil count was an independent predictor. The combined application of the TG 2013 guidelines led to a better sensitivity of 83.1 % at the cost a reduced specificity of 37.5 % compared to neutrophil count alone. The accuracy was therefore only 60.3 %, which was well below the TG 2013 report.

Conclusion The 2013 Tokyo Guidelines were slightly better in predicting acute cholecystitis but over diagnosed two-thirds of normal gallbladders compared to neutrophil count alone.

Introduction

The overall prevalence of gallstone disease in the Western World is estimated to be between 10 and 15 %, increasing with age. Although the fast majority of people with gallstones remain asymptomatic; per year 1–4 % will develop symptoms of which one-fifth will present with acute

cholecystitis [1–3]. In 2008 within the USA, the diagnosis of acute cholecystitis alone led to 120,000 cholecystectomies performed in that single year [1].

Distinguishing acute cholecystitis from biliary colic in patients presenting to the emergency department with right upper quadrant pain remains a challenging task. Whilst most patients with biliary colic will settle without any further complications, for patients with acute cholecystitis severe complications like gallbladder empyema, perforation, and gangrenous cholecystitis leading to peritonitis and sepsis can lead to significant morbidity and even mortality [1, 4, 5]. Of all registered Australian deaths in 2008, a mortality rate of 0.1 % resulted from acute cholecystitis

✉ E. Beenen
  edwinbeenen@gmail.com

1 Upper GI/HPB Unit, Department of General Surgery, The
   Canberra Hospital, Building 6, Yamba Drive, Garran,
   ACT 2605, Australia

[6]. As such, a rapid and accurate diagnosis is fundamental to determine the appropriate management.

Making the right diagnosis based on clinical findings, biochemistry and imaging has proven to be difficult. This led to a proposal by consensus to standardize the diagnostic criteria: the 2007 Tokyo Guidelines (TG 2007) [5]. To improve diagnostic accuracy further, the Tokyo Guidelines were revised in 2013 (TG 2013) with a self-acclaimed accuracy of well over 90 % [7]. The diagnosis of acute cholecystitis according to TG 2013 guidelines is suspected if RUQ pain and tenderness is associated with systemic features of inflammation. The diagnosis is considered certain if imaging confirms acute cholecystitis [7]. The guidelines allow all imaging modalities to aid in the diagnosis of acute cholecystitis, but recommend abdominal ultrasonography (US) for its presumed reliability and logistical advantages [5].

The accuracy of the TG 2013 guidelines has, however, hardly been tested outside the TG 2013 workgroup and not been validated independently.

We therefore conducted a retrospective review equal to the TG 2013 validation process to determine the accuracy of the TG 2013 guidelines in the diagnosis of acute cholecystitis. In addition, we analyzed the single components of the TG 2007 criteria of fever, inflammatory markers and US findings.

Methods

The Canberra Hospital is a tertiary referral and teaching hospital providing acute surgical care to the Australian Capital Territory and surrounding New South Wales region.

All consecutive patients who had an emergency cholecystectomy at the Canberra Hospital between 1 January 2008 and 31 December 2012 were identified. From these patients, we analyzed adult patients operated with a working diagnosis of acute cholecystitis or persistent biliary colic and who had a pre-operative abdominal US maximum 5 days prior to surgery.

Other indications for an emergency cholecystectomy, e.g. biliary pancreatitis or cholangitis, were excluded as the underlying disease was presumed to elevate the inflammation markers and/or influence the findings on imaging and thus interfere with the TG 2013 criteria. We, arbitrarily, excluded patients who had an interval of more than 5 days between the pre-operative ultrasound and the operation as we assumed the result of histopathology could potentially no longer verify the findings seen on US. Therefore, acute patients who settled down on antibiotics before they could be operated on the emergency list and were rebooked as elective patients as well as all other elective patients were not included in this study nor considered as a control group as almost every patient had an US well before having surgery.

We gathered data regarding patient demographics (sex and age), highest recorded pre-operative in-hospital temperature, inflammatory markers and US findings. Histopathology of the gallbladder specimen was used as the gold standard for the diagnosis of acute cholecystitis [1, 7].

Since fever is not strictly defined in the TG 2007 and TG 2013 guidelines, cut-off points of temperatures above 37.5 and 38.0 °C were tested. The normal haematology and serology was defined as: white cell count (WCC) $<11 \times 10^9$ cells/L; neutrophils (Nc) $<7.5 \times 10^9$ cells/L; and C-Reactive Protein titres (CRP) <5 mg/L.

The US was preferably performed in a fasted patient. The need for an ultrasound was at the discretion by the on-call surgeon and also depended on whether the patient already had proven gallstones at previous imaging. The indications for US were either to prove the existence of gallstones or to confirm the diagnosis of acute cholecystitis. As per the 2013 Tokyo Guidelines, only three features found on US were scored as present or not present: gallbladder wall thickening (≥ 5 mm), the presence of pericholecystic fluid and probe tenderness (US Murphy's sign) [7]. Exact gallbladder wall thickness was derived either directly from the report, or re-measured on imaging if not specifically mentioned. The other two features were only scored as present if clearly mentioned within the radiologist's report.

The histopathology of the cholecystectomy specimen was used as the gold standard [1, 7, 8]. On microscopy, acute cholecystitis was defined by oedema, transmural leucocyte infiltration (especially neutrophils), vascular congestion and/or abscess formation with or without gallbladder wall necrosis. As per TG 2013, a normal gallbladder wall or chronic cholecystitis was scored as normal [7].

We compared the patient demographics, temperature, inflammatory markers and the individual features and final conclusion of the US to the golden standard of histopathology; both individually and combined according to the TG 2013 guidelines.

For statistical analysis, StatView 5.0.1 (SAS Institute, Cary, NC, USA) was used. Statistical difference was computed for continuous data using the Mann–Whitney U test. The Fisher's exact and χ^2 test was used for categorized data. A univariate analysis was performed to identify potential independent parameters. All parameters with statistical significance ($p < 0.05$) or a clear trend ($p < 0.10$) in predicting acute cholecystitis were included in the multivariate analysis. For these parameters, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated.

Ethics approval for the study was obtained from the ACT Health Human Research Ethics Committee.

Results

In the 5-year study period, 448 adult patients had undergone an emergency cholecystectomy. Of these, 145 patients were operated for a different indication (biliary pancreatitis $n = 95$, cholangitis/choledocholithiasis $n = 26$ or miscellaneous diagnosis $n = 24$) and were thus excluded from the study.

The remaining 303 patients were operated with the diagnosis of either acute cholecystitis or non-resolving biliary colic. A further 134 patients were excluded because they did not have an US on the same admission ($n = 113$) or the interval between US and cholecystectomy exceeded 5 days ($n = 21$); leaving 169 patients for our study. All patients had presented through emergency department (ED) with RUQ pain and tenderness (TG 2013 A-criteria).

Of the 169 patients, 64 % were female with a median age of 43 years (range 14.9–87.6). On histopathology, 52.7 % (89/169) had acute cholecystitis (acute cholecystitis $n = 68$, gangrenous cholecystitis $n = 21$); the remaining 80 patients were scored as normal gallbladder (normal $n = 4$, chronic cholecystitis $n = 76$). The median interval between US and cholecystectomy was 1.3 days (range 0.1–5.1 days). An increased interval was not related to a reduced prevalence of acute cholecystitis ($p = 0.132$).

A temperature above 37.5 °C was present in 31.5 % (53/168; temperature for one patient missing) and above 38.0 °C in 17.9 % (30/168) of all patients. Neither a temperature above 37.5 °C nor above 38.0 °C was predictive of acute cholecystitis (see Table 1). Furthermore, the presence of a temperature above 38.0 °C was also not related to the degree of WCC elevation ($p = 0.505$), neutrophil count ($p = 0.498$) or CRP ($p = 0.969$).

On univariate analysis, all acute inflammatory markers (WCC, neutrophils and CRP) were statistically significant in predicting acute cholecystitis, with the neutrophil count having the slightly better predictive values with a sensitivity of 70.0 % and a specificity of 65.8 % (see Table 1). A CRP was unfortunately not determined in 54 patients.

Of the ultrasonography features, the wall thickening and the presence of pericholecystic fluid were predictive for acute cholecystitis (see Table 1). Acute cholecystitis patients showed a slight, statistically significant increase in wall thickness of 0.6 mm. However, due to the large overlap in range, its clinical value was very debatable (Fig. 1). At the TG 2013 cut-off point of 5 mm, wall thickness was not contributing in the diagnosis of acute cholecystitis ($p = 0.418$) nor was probe tenderness
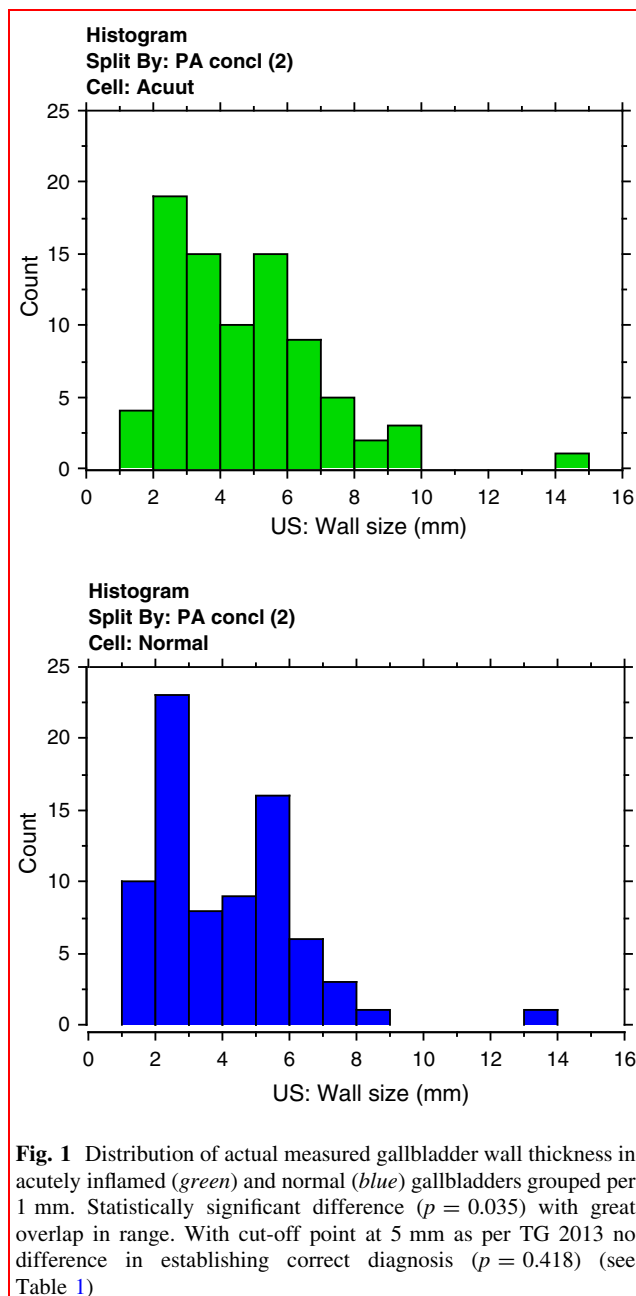
Table 1 Presence and predictive values for 2013 Tokyo Guidelines

Test	Acute cholecystitis ($n = 90$)	Normal gallbladder ($n = 79$)	Sens	Spec	PPV	NPV	Univariate analysis, p value	Multivariate analysis, p value
B-criteria: signs of inflammation								
Temperature >37.5 °C, presence	36.0 % (32/89) *	26.6 % (21/79)	36.0 %	73.4 %	60.4 %	50.4 %	0.244	X
Temperature >38.0 °C, presence	21.3 % (19/89) *	13.9 % (11/79)	21.3 %	86.1 %	63.3 %	49.3 %	0.232	X
White cell count ($\times 10^9$ cells/L), median (range)	12.7 (1.0–29.8)	8.3 (2.1–18.8)	63.3 %	65.8 %	67.9 %	61.2 %	0.0002	0.263
Neutrophils count ($\times 10^9$ cells/L), median (range)	10.1 (0.5–28.3)	6.0 (1.2–16.5)	70.0 %	65.8 %	70.0 %	65.8 %	<0.0001	0.014
C-reactive protein (mg/L), median (range)	52 (0.8–416) Δ	12 (0.2–265)	86.7 %	32.7 %	58.4 %	69.2 %	0.015	0.085
C-criteria: ultrasound features								
Wall thickening, absolute, median (range) in mm	4.1 (1.6–14) ∓	3.5 (1.2–13)					0.035	0.804
Wall thickening, cut-off >5 mm, presence	42.2 % (35/83) ∓	35.1 % (27/77)	42.2 %	64.9 %	56.5 %	51.0 %	0.418	X
US Murphy's sign (probe tenderness), presence	55.6 % (50/90)	54.4 % (43/79)	55.6 %	45.6 %	53.8 %	47.4 %	>0.999	X
Pericholecystic fluid, presence	32.2 % (29/90)	15.2 % (12/79)	32.2 %	84.8 %	70.7 %	52.3 %	0.012	0.324

B-criteria: fever and inflammation markers (WCC, neutrophils, CRP) and C-criteria: imaging (US only in this study)

* Temperature missing for one patient with acute cholecystitis. Δ CRP missing for 54 patients. ∓ Actual wall thickness missing for nine patients

Sens sensitivity, *Spec* specificity, *PPV* positive predictive value, *NPV* negative predictive value (cut-off values: WCC <11 × 10^9 cells/L; neutrophils <7.5 × 10^9 cells/L; CRP <5 mg/L)

Fig. 1 Distribution of actual measured gallbladder wall thickness in acutely inflamed (*green*) and normal (*blue*) gallbladders grouped per 1 mm. Statistically significant difference ($p = 0.035$) with great overlap in range. With cut-off point at 5 mm as per TG 2013 no difference in establishing correct diagnosis ($p = 0.418$) (see Table 1)

($p > 0.999$). For nine patients, the actual wall thickness could not be retrieved.

At multivariate analysis, all inflammation markers and significant US features were included. Only neutrophil count was found to be an independent predictor of acute cholecystitis ($p = 0.014$), although the sensitivity and specificity were poorly performing with 70.0 and 65.8 %, respectively. The CRP showed a trend towards significance ($p = 0.085$); whilst all other TG 2007 criteria lost significance (see Table 1).

The ability of the TG 2013 guidelines to identify acute cholecystitis was statistically significant (p value <0.001). However, a definitive diagnosis of acute cholecystitis

(presence of TG 2013 criteria A + B + C) was only found in 80 of 169 patients of which 59 patients (73.8 %) had histological confirmation of acute cholecystitis. A suspected diagnosis of acute cholecystitis (presence of A- and B-criteria only) was present in 44 of 169 patients, of which only 15 patients (34.1 %) had histological confirmation of acute cholecystitis. In 45 patients, criteria for either suspicious or definitive diagnosis of acute cholecystitis were not fulfilled according to TG 2103 guidelines. Of the 25 patients with acute cholecystitis on US but without raised inflammation markers 11 patients (44 %) had histological acute cholecystitis, and of the 20 patients with neither raised inflammation markers nor acute cholecystitis on US, four patients (20 %) still had acute cholecystitis confirmed on histology (Table 2).

The TG 2013 correctly predicted 83.1 % (74/89) of all histology confirmed acute cholecystitis, but over diagnosed in 62.5 % (50/80) of the normal gallbladders (Table 2) leading to an accuracy rate of 60.3 %.

Discussion

The diagnosis of acute cholecystitis has in the past been based on history and physical examination in conjunction with biochemical investigations [1]; however, the accuracy was low [3]. This led to the development by consensus of the 2007 Tokyo Guidelines, in an attempt to establish diagnostic and reporting criteria [5]. Although an increased diagnostic accuracy was reported after the introduction of the guidelines, accuracy was not increased to a satisfactory level; leading to the subsequent revision of the Tokyo Guidelines in 2013 [7]. In the TG 2013 guidelines, a suspicion of acute cholecystitis relies on the presence of fever or raised inflammation markers (B-criteria). To confirm the diagnosis of acute cholecystitis confirmation by imaging (C-criteria) needs to be sought [7]. Although multiple imaging modalities (US, CT-scan, MRI and HIDA scan) are allowed within the guidelines; US is recommended [7].

With this retrospective study, we found that the presence of fever, inflammatory markers and US imaging performed poorly in distinguishing acute cholecystitis from biliary colic in the emergency patient. In our series, only the elevated neutrophil count was an independent predictor with a slightly lower sensitivity but far better specificity compared to the compiled TG 2013 criteria, with no predictive value at all with the presence of fever. In literature, the role of fever and inflammatory markers has mainly been described as a predictive factor for conversion to open cholecystectomy [9] or as a predictor of gangrenous cholecystitis [10]. Its role as an aid in the actual diagnosis of acute cholecystitis is, however, poorly described and only the study by Juvonen et al. [11] reported an 18 %

Table 2 Predictive value of the 2013 Tokyo Guidelines for acute cholecystitis

Clinical diagnosis by TG 2013 guidelines	B-criteria	C-criteria	Number (n)	AC on histology	AC histology (n = 89)	Normal histology (n = 80)
Confirmed AC	Present	Present	80	73.8 % (59/80)	Correct diagnosis (sensitivity)	Overdiagnosis
Suspected AC	Present	Absent	44	34.1 % (15/44)	83.1 % (74/89)	62.5 % (50/80)
No AC	Absent	Present	25	44.0 % (11/25)	Missed diagnosis	Correct diagnosis (specificity)
	Absent	Absent	20	20.0 % (4/20)	16.9 % (15/89)	37.5 % (30/80)

All patients presented with RUQ pain and tenderness (A-criteria)

TG 2013 = 2013 Tokyo Guidelines, AC = acute cholecystitis, B-criteria = presence of systemic signs of inflammation (fever, abnormal white cell count, raised C- reactive protein), C-criteria = signs of acute cholecystitis on imaging; which was ultrasound only in our series

increase in diagnostic accuracy from 79 to 97 % when an increased CRP level was added to the US findings. Our series found that the neutrophil count was the most reliable and only independent predictor of acute cholecystitis, although its predictive values in the context of a prevalence of acute cholecystitis of 52 % would still have left many patients under or over diagnosed.

The US of the gallbladder is well recognized for its ability to identify gallstones with sensitivities and specificities reported of 98 and 95 %, respectively [1]. However, the US's ability to recognize acute cholecystitis is more debatable with sensitivities and specificities ranging from 27.2 to 93 % and 46 to 89 %, respectively [3, 8]. A recent 2012 meta-analysis estimated the sensitivity and specificity of US in identifying acute calculous cholecystitis at 81 and 83 % [12]. In our series, sensitivity and specificity were below that with 78.9 and 55.7 % based on the report's final conclusion.

A further point is that the US criteria to identify acute cholecystitis are not established with changing criteria between different papers [3, 5, 7, 8, 13–16]. In the meta-analysis, 14 different criteria protocols were found amongst 26 studies [12].

Wall thickness is generally presumed to be predictive [8], but became insignificant in our multivariate analysis and has also been an unreliable feature in other studies [13, 16]. When we applied the 5-mm cut-off as per TG 2013 no statistical significant difference was found ($p = 0.418$). In literature, hyperaemia is proposed to be the distinguishing feature in thickened gallbladder walls between acute inflammation and secondary thickening due to other causes, including chronic cholecystitis [16–18]; but this feature is not included in the TG 2013. The feature of probe tenderness (US Murphy's sign) has been recommended [7, 16–18], but we found it to be absolutely unhelpful ($p > 0.999$) as has been reported as well in the literature [13–15]. Our series is of course a retrospective analysis where reporting was not according to a study protocol. Although it could be argued that reporting might

have been less thorough; we feel it is more likely to represent the everyday practice in many hospitals.

In the recent meta-analysis, a HIDA scan was found to be far more reliable in predicting acute cholecystitis with a sensitivity and specificity of 95 and 92 % [1, 12]. However, the HIDA scan is not as easily available, provides no imaging of structures outside the biliary tract and is more time consuming and invasive in nature compared to an US [7, 12]. The role of CT-scan and MRI in the diagnosis of acute cholecystitis had been underreported in the literature to make any recommendations [12]. We only used US as the imaging study with a median interval of 1.3 days between study and operation. In the validation study of the TG 2013 guidelines, neither the imaging modality nor the result of the imaging is mentioned or analyzed. Rather their recommendations regarding imaging came from their review of the literature [7].

The combined application of the TG 2013 criteria did not increase the diagnostic accuracy (60.3 %) compared to the neutrophil count alone (67.9 %) and tended to over diagnose acute cholecystitis in patients with histological normal gallbladders. Just like the TG 2013 publication, our series was a retrospective analysis with a nearly equal prevalence of acute cholecystitis of just over 50 % [7]. We did not manage to reach the high accuracy as reported by TG 2013 of over 90 %. As the TG 2013 validation does not report on the imaging modalities used, nor any analysis is given for its separate components [7]; we cannot comment on how this difference in accuracy could be explained.

Conclusion

We found neutrophil count to be the only independent predictor of acute cholecystitis, although its predictive values were not satisfactory. Ultrasound findings did not make an independent contribution in the diagnosis of acute cholecystitis. The application of the 2013 Tokyo

Guidelines as a whole did lead to a slightly better identification of patients with acute cholecystitis, but also led to the over diagnoses of nearly two-thirds of the patients with chronic cholecystitis or normal gallbladders leaving a mediocre overall accuracy of only 60.3 %.

In our opinion, a broader range of clinical parameters and imaging features should be analyzed in a prospective protocol rather than consensus opinion in order to establish a reliable and uniformly agreed upon guideline.

References

1. Strasberg SM (2008) Clinical practice. Acute calculous cholecystitis. N Engl J Med 358(26):2804–2811
2. Halldestam I, Enell EL, Kullman E, Borch K (2004) Development of symptoms and complications in individuals with asymptomatic gallstones. Br J Surg 91(6):734–738
3. Al-Azawi D, Mc Mahon D, Rajpal PK (2007) The diagnosis of acute cholecystitis in patients undergoing early laparoscopic cholecystectomy in a community hospital. Surg Laparosc Endosc Percutan Tech. 17(1):19–21
4. Everhart JE, Khare M, Hill M, Maurer KR (1999) Prevalence and ethnic differences in gallbladder disease in the United States. Gastroenterology 117(3):632–639
5. Hirota M, Takada T, Kawarada Y, Nimura Y, Miura F, Hirata K, Mayumi T et al (2007) Diagnostic criteria and severity assessment of acute cholecystitis: Tokyo Guidelines. J Hepatobiliary Pancreat Surg 14(1):78–82
6. 3303.0—Causes of Death, Australia, 2008. Released 31-10-2010 by Australian Bureau of Statistics. www.abs.gov.au
7. Yokoe M, Takada T, Strasberg SM, Solomkin JS, Mayumi T, Gomi H, Pitt HA et al (2012) Tokyo Guidelines Revision Committee New diagnostic criteria and severity assessment of acute cholecystitis in revised Tokyo Guidelines. J Hepatobiliary Pancreat Surg 19(5):578–585
8. Kalimi R, Gecelter GR, Caplin D, Brickman M, Tronco GT, Love C, Yao J et al (2001) Diagnosis of acute cholecystitis: sensitivity of sonography, cholescintigraphy, and combined sonography-cholescintigraphy. J Am Coll Surg 193(6):609–613
9. Wevers KP, van Westreenen HL, Patijn GA (2013) Laparoscopic cholecystectomy in acute cholecystitis: C-reactive protein level combined with age predicts conversion. Surg Laparosc Endosc Percutan Tech. 23(2):163–166
10. Mok KW, Reddy R, Wood F, Turner P, Ward JB, Pursnani KG, Date RS (2014) Is C-reactive protein a useful adjunct in selecting patients for emergency cholecystectomy by predicting severe/gangrenous cholecystitis? Int J Surg. 12(7):649–653
11. Juvonen T, Kiviniemi H, Niemelä O, Kairaluoma MI (1992) Diagnostic accuracy of ultrasonography and C reactive protein concentration in acute cholecystitis: a prospective clinical study. Eur J Surg 158(6–7):365–369
12. Kiewiet JJ, Leeuwenburgh MM, Bipat S, Bossuyt PM, Stoker J, Boermeester MA (2012) A systematic review and meta-analysis of diagnostic performance of imaging in acute cholecystitis. Radiology 264(3):708–720
13. Bingener J, Schwesinger WH, Chopra S, Richards ML, Sirinek KR (2004) Does the correlation of acute cholecystitis on ultrasound and at surgery reflect a mirror image? Am J Surg 188(6):703–707
14. Bree RL (1995) Further observations on the usefulness of the sonographic Murphy sign in the evaluation of suspected acute cholecystitis. J Clin Ultrasound 23(3):169–172
15. Kendall JL, Shimp RJ (2001) Performance and interpretation of focused right upper quadrant ultrasound by emergency physicians. J Emerg Med 21(1):7–13
16. Uggowitzer M, Kugler C, Schramayer G, Kammerhuber F, Gröll R, Hausegger KA, Ratschek M, Quehenberger F (1997) Sonography of acute cholecystitis: comparison of color and power Doppler sonography in detecting a hypervascularized gallbladder wall. AJR Am J Roentgenol 168(3):707–712
17. Draghi F, Ferrozzi G, Calliada F, Solcia M, Madonia L, Campani R (2000) Power Doppler ultrasound of gallbladder wall vascularization in inflammation: clinical implications. Eur Radiol 10(10):1587–1590
18. Sariego J, Matsumoto T, Kerstein M (1992) Significance of wall thickness in symptomatic gallbladder disease. Arch Surg 127(10):1216–1218