Emergency Physician–performed Ultrasound to Diagnose Cholelithiasis: A Systematic Review

Marshall Ross, Michael Brown, MD, MSc, Kyle McLaughlin, MD, CCFP-EM, Paul Atkinson, MB, BCh, BAO, MA, MRCP(UK), FCEM, Jenny Thompson, Susan Powelson, MLS, Steve Clark, MD, CCFP-EM, and Eddy Lang, MD

Abstract

Objectives: The authors sought to determine the diagnostic test characteristics of bedside emergency physician (EP)-performed ultrasound (US) for cholelithiasis in symptomatic emergency department (ED) patients.

Methods: A search was conducted of MEDLINE, EMBASE, the Cochrane Library, bibliographies of previous systematic reviews, and abstracts from major emergency medicine conference proceedings. We included studies that prospectively assessed the diagnostic accuracy of emergency US (EUS) for cholelithiasis, compared to a criterion reference standard of radiology-performed ultrasound (RADUS), computed tomography (CT), magnetic resonance imaging (MRI), or surgical findings. Two authors independently performed relevance screening of titles and abstracts, extracted data, and performed the quality analysis. Disagreements were resolved by conference between the two reviewers. EUS performance was assessed with summary receiver operator characteristics curve (SROC) analysis, with independently pooled sensitivity and specificity values across included studies.

Results: The electronic search yielded 917 titles; eight studies met the inclusion criteria, yielding a sample of 710 subjects. All included studies used appropriate selection criteria and reference standards, but only one study reported uninterpretable or indeterminate results. The pooled estimates for sensitivity and specificity were 89.8% (95% confidence interval [CI] = 86.4% to 92.5%) and 88.0% (95% CI = 83.7% to 91.4%), respectively.

Conclusions: This study suggests that in patients presenting to the ED with pain consistent with biliary colic, a positive EUS scan may be used to arrange for appropriate outpatient follow-up if symptoms have resolved. In patients with a low pretest probability, a negative EUS scan should prompt the clinician to consider an alternative diagnosis.

ACADEMIC EMERGENCY MEDICINE 2011; 18:227–235 © 2011 by the Society for Academic Emergency Medicine

I t is estimated that over 20 million people in the United States have gallbladder disease.[1] Approximately one-third of patients with gallstones develop biliary colic,[2–4] and each year about 1% of these symptomatic patients develop potentially life-threatening complications including acute cholecystitis, pancreatitis, and acute cholangitis.[5] Unfortunately, it is not possible to diagnose gallbladder pathology with a reasonable degree of certainty using history and physical exam alone.[6] The unreliability of clinical findings commonly leads to patients with suspected gallbladder disease undergoing radiology-performed ultrasound (RADUS).

These RADUS are typically performed by a technician outside of the emergency department (ED) and are interpreted by a radiologist. However, inconsistent availability of technicians and high volumes of ED patients often cause delays to RADUS, extending ED wait times, compromising quality of care, and contributing to throughput delays and crowding.[7,8]

These limitations have led to emergency physicians (EPs) exploring the prospect of performing bedside emergency ultrasound (EUS) in the ED to assess for gallbladder pathology.[9–11] EUS enables the operator to visualize the gallbladder. Stones are typically identified as bright, mobile, and dependent structures that cast a shadow. False negatives are most commonly due to stones that are either less than 4 mm in diameter and thus do not consistently cast a shadow or impacted in the gallbladder neck as they are no longer dependent.[12]

Whereas the use of EUS for certain core ED applications, such as screening for abdominal aortic aneurysm, or detecting large pericardial effusions, is supported by strong evidence,[13,14] the use of EUS for many other indications remains controversial due to concerns about the diagnostic accuracy. The goal of this systematic review was to assess the diagnostic accuracy of EUS in detecting gallstones among ED patients who present with symptoms consistent with biliary colic.

METHODS

The clinical question addressed in this systematic review was the following: Among patients presenting to the ED with abdominal pain of suspected biliary origin, what is the diagnostic accuracy of EP-performed ultrasound (US) for the diagnosis of acute cholelithiasis? A systematic review protocol was created to specifically address this question and was reviewed and agreed upon by all co-investigators a priori (available online[15]). The only amendment made to the protocol was to expand our possible criterion reference standard beyond RADUS alone. The PRISMA statement[16] was followed for the purpose of reporting this systematic review.

Search Techniques

We developed a comprehensive search strategy that included foreign language publications. The MEDLINE search used a combination of MeSH headings and free text terms (Data Supplement S1, available as supporting information in the online version of this paper). Similar search strategies were adapted for EMBASE, OpenSIGLE, and the Cochrane Library. Our gray literature search included manually reviewing abstracts from the following conferences: Canadian Association of Emergency Physicians (2000–2009), American College of Emergency Physicians (ACEP) (2000–2009), Society for Academic Emergency Medicine (2000–2009), Mediterranean Congress of Emergency Medicine (2001–2009), Annual Scientific Meeting of the Australasian College for Emergency Medicine (2001–2009), and all available abstracts from the College of Emergency Medicine, formerly known as the British Association and Faculty of Emergency Medicine (2000–2009). In addition, the references for the ACEP Policy Statement on Emergency Ultrasound Guidelines[9] were screened. We contacted

all major US manufacturers, the Society for Academic Emergency Medicine Ultrasound Interest Group, several independent experts in the field, and the Canadian Emergency Ultrasound Society and requested additional unpublished data sources. The bibliographies of all included studies[17–24] as well as those of previous reviews[25,26] were hand-searched.

Article Selection

Specific inclusion and exclusion criteria were applied to minimize bias in the selection process and ensure that any inferences made based upon the results of the meta-analysis were appropriate and applicable to the ED setting.[27] We included studies with patients who presented to the ED with signs and symptoms of suspected biliary colic including right upper quadrant (RUQ) pain, epigastric pain, or flank pain. We included studies if they used an acceptable criterion reference standard defined as at least one of the following: RADUS, computed tomography (CT), magnetic resonance imaging (MRI) interpreted by a radiologist, or surgical findings. If other reference standards were used, studies were included only if data were obtained comparing results of EUS to at least one of the criterion standards. We included studies assessing the use of EUS for multiple purposes only if data comparing EUS to an appropriate reference standard to detect cholelithiasis were presented or obtained through author contact. We only included studies deemed original research using prospective data collection. We excluded studies if pertinent data could not be obtained through author contact.

Data Collection and Processing

Two reviewers (MR, JT) independently performed a relevance search; each reviewer examined the titles and abstracts of all references identified in the electronic search to determine whether an article was relevant to the research question. The two reviewers then compared their inclusion and exclusion logs, and the level of agreement was calculated using the kappa statistic. The same two reviewers (MR, JT) then performed a full review of all potentially relevant articles applying specific inclusion and exclusion criteria to determine which papers to include in the final analysis.

Data were extracted and collected from included studies using a standardized data collection form. An attempt was made to contact the authors of studies that required data clarification or supplementation. Two reviewers (MR, EL) independently abstracted the data. Disagreements were resolved by conference between the reviewers.

Quality Assessment

Two reviewers with experience in critical appraisal (KM, EL) independently assessed the quality of the studies according to the most relevant items from the QUADAS tool.[28] QUADAS was developed as an evidence-based quality assessment tool to be used in systematic reviews to assess the quality of primary studies of diagnostic accuracy.[28] Although all 14 points of the QUADAS instrument were assessed, seven were considered to be of particular relevance to our clinical

question: 1) was the spectrum of patients representative of the patients who will receive the test in practice, 2) were selection criteria clearly described, 3) is the reference standard likely to correctly classify the target condition, 4) were the reference standard results interpreted without knowledge of the results of the index test, 5) were the index test results interpreted without knowledge of the results of the reference standard, 6) did patients receive the same reference standard regardless of the index test result, and 7) were uninterpretable or intermediate test results reported?

No attempt was made to combine the results into an overall quality score; instead, the results of the assessment were reported in a summary figure. Ultrasound technology used in each study was assessed independently from other methods and results.

Data Analysis

Emergency US test performance was assessed with a traditional summary receiver operator curve (SROC) with a regression model based on unweighted least squares estimation.[29,30] We independently pooled the sensitivity and specificity with 95% confidence intervals (CIs) using the random effects model of DerSimonian and Laird.[31] All calculations were done with the Meta-Disc software.[32] We estimated likelihood ratios (LRs) based on our summary estimates of sensitivity and specificity.[27]

One of the principal causes of heterogeneity between studies of diagnostic test accuracy is a threshold effect. A threshold effect is present when differences in sensitivities and specificities among studies occur due to different cutoffs or test thresholds used to define a positive or negative result.[27,33] EUS exams may have an implicit variation in threshold as a positive or negative result depends on operator interpretation.[27] We assessed for a threshold effect using a validated regression model (see Data Supplement S2 for technical details, available as supporting information in the online version of this paper).

The random effects model of DerSimonian and Laird[31] accounts for sampling variability and unexplained heterogeneity providing suitably conservative estimates of sensitivity and specificity.[27] Formal statistical tests for assessing heterogeneity and publication bias in diagnostic meta-analysis have not been validated.[27,33,34] Therefore, funnel plots and statistical tests such as I^2 were not reported.

RESULTS

Figure 1 summarizes the results of the search and article selection. The relevance screen of 917 titles identified in the MEDLINE, Cochrane, OpenSIGLE, and EMBASE databases showed good agreement between the two reviewers ($\kappa = 0.67$). Expert contact and relevance screening of gray literature and conference proceedings yielded seven potentially relevant articles. After the relevance search was adjudicated, a complete article review was performed on the remaining 33 articles. Upon full-text review, 25 studies did not meet our inclusion criteria ($\kappa = 1$) for various reasons as described in Figure 1. One article retrieved through the
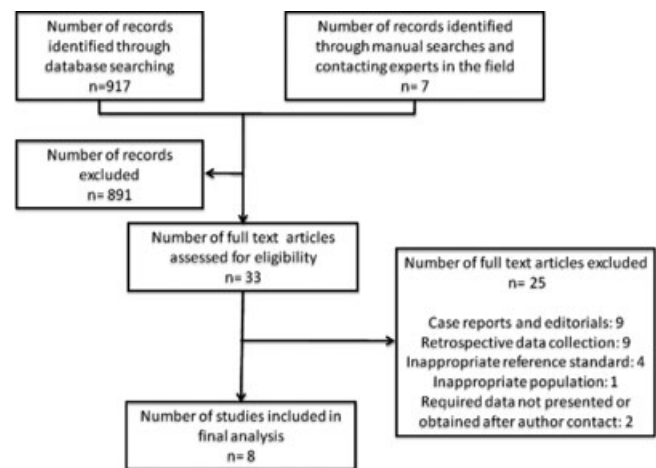

Figure 1. Results of the search strategy.

manual search,[17] one article retrieved through expert contact,[19] and six articles[18,20–24] from the electronic database search were included in the final analysis. Seven articles[17,18,20–24] were written in English and one article was available only in Korean.[19] This article was translated into English for full review. These eight studies yielded a sample of 710 subjects who underwent both EUS and a criterion reference standard. The sensitivity and specificity of each study are presented in Figure 2 and Table 1.

Study Description

Table 2 outlines the characteristics of the eight included studies. The prevalence of gallstones ranged from 46% to 80%, with a median of 60%. In five of the eight studies selected, assessment of the accuracy of EUS to detect gallstones was a primary outcome measure.

The primary focus of the article by Davis et al.[18] was to determine the association between operator confidence and the accuracy of EUS using different criterion standards for each of six different US examinations. The primary goal of the study by Summers et al.[24] was to determine the test characteristics of EUS as well as RADUS for detection of acute cholecystitis, not cholelithiasis.[24] This study compared the outcomes of both EUS and RADUS to clinical follow-up or surgical pathology for 189 patients. The primary focus of the study by Rowland et al.[23] was to determine the accuracy of 9 different EUS exams including a RUQ scan to detect for the presence of gallstones.[23]

There was significant variability in the level of US education and previous experience with RUQ US across studies. Davis et al.,[18] Rowland et al.,[23] and Ha et al.[19] used admittedly naïve operators; Alexander et al.[17] did not report operator experience; and the other four studies included operators with a wide range of experience. In all studies, EUS was performed with a US machine of acceptable quality that is regularly used in emergency practice (Table 2).

Quality Assessment

Table 3 summarizes the seven key quality items. The inclusion criteria required that all eight studies be rated
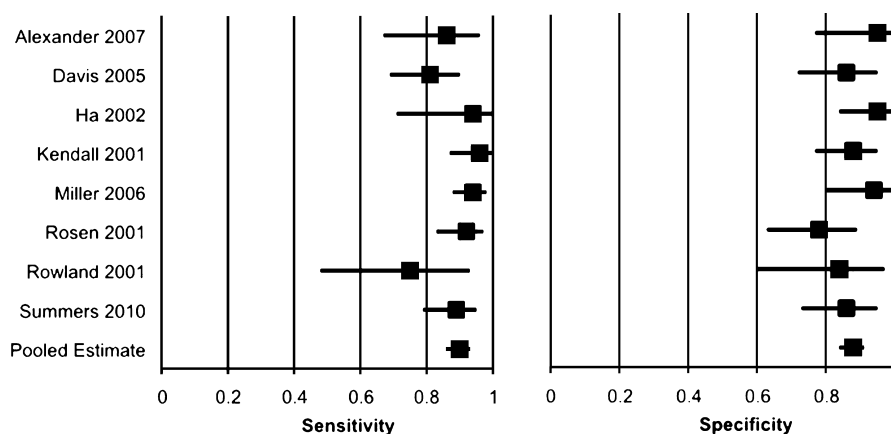
Figure 2. Forest plot of individual study and pooled sensitivity and specificity.

Table 1
Individual Study and Pooled Sensitivity and Specificity

Study	TP	FP	TN	FN	n	Sensitivity (95% CI)	Specificity (95% CI)
Alexander 2008[17]	24	1	21	4	50	0.86 (0.67–0.96)	0.95 (0.77–1.00)
Davis 2005[18]	50	6	37	12	105	0.81 (0.69–0.90)	0.86 (0.72–0.95)
Ha 2002[19]	16	2	40	1	59	0.94 (0.71–1.00)	0.95 (0.84–0.99)
Kendall 2001[20]	49	7	51	2	109	0.96 (0.87–1.00)	0.88 (0.77–0.95)
Miller 2006[21]	95	1	25	6	127	0.94 (0.88–0.97)	0.96 (0.80–1.00)
Rosen 2001[22]	60	10	35	5	110	0.92 (0.83–0.97)	0.78 (0.63–0.89)
Rowland 2001[23]	12	3	16	4	35	0.75 (0.48–0.93)	0.84 (0.60–0.97)
Summers 2010[24]	63	6	38	8	115	0.89 (0.79–0.95)	0.86 (0.73–0.95)
Pooled					710	0.90 (0.86–0.93)	0.88 (0.84–0.91)

FN = false negative; FP = false positive; n = number of subjects; TN = true negative; TP = true positive.

''yes'' for having an acceptable reference standard and selection criteria. Seven studies[17,18,20–24] adequately described a representative patient spectrum. Ha et al.[19] did not describe their patient population in sufficient detail. In five studies,[17,19–22] all patients underwent RADUS as the only reference standard. Davis et al.[18] used US, CT, MRI, or surgical findings as a criterion reference. Rowland et al.[23] used RADUS or laparotomy as a criterion reference. Summers et al.[24] provided data for 115 of 189 patients who underwent formal radiology testing; 91 were tested with RADUS and 24 with CT. Only one study reported indeterminate results.[20] In six studies,[17,20–24] the RADUS was interpreted without the knowledge of EUS findings. In two studies,[18,19] this was not adequately reported. In all studies, the EUS was performed and interpreted without knowledge of the results of the RADUS.

Data Analysis
According to the traditional SROC curve analysis,[30] there was no evidence of a significant threshold effect ($\beta = -0.072$, 95% CI = −0.42 to 0.56); in other words, there was no implicit variation in sensitivity and specificity across studies due to operator-dependent differences in what defined a positive or negative test result. The random effects pooled results for sensitivity and specificity were 89.8% (95% CI = 86.4% to 92.5%) and 88.0% (95% CI = 83.7% to 91.4%), respectively.

According to these summary estimates, the positive LR was 7.5 and the negative LR was 0.12. The SROC curve is presented in Figure 3. The small size and number of studies precluded any meaningful subgroup analysis.

DISCUSSION
It is estimated that over 20 million people in the United States have gallbladder disease,[1] and nearly 1 in 10 asymptomatic individuals with gallstones will require treatment within 5 years.[35] Despite the high prevalence of the disease, the clinical diagnosis is still problematic and often relies on a RADUS performed in the radiology department. The use of RADUS is limited by inconsistent availability in smaller community settings. Benefits of using EUS to assess the RUQ include its portability, decreased costs,[36] and ability to decrease the ED length of stay.[8]

Based on the results of this meta-analysis, a positive EUS (LR = 7.5) in patients presenting with a clinical picture consistent with a high probability of biliary colic may be sufficient to make the provisional diagnosis of acutely symptomatic cholelithiasis and arrange for appropriate follow-up if the symptoms resolve in the ED. For example, if we were to assume a pretest probability of 60% (the median prevalence across studies in this review), a positive EUS provides a posttest probability of 92%. At the other end of the spectrum, the EP

Table 2
Characteristics of Included Studies

Study	Outcomes	Center	Population	Bedside EUS	Reference Standard
Alexander 2008[17]	1°—Accuracy of EUS to detect gallstones 2°—Accuracy of EUS to detect GBWT, PCF, and CBDD	Australian tertiary care center with ~47,000 annual ED visits	Convenience sample of patients over 18 who presented to the ED with RUQ pain Mean age: 53 yr (range = 20–85 yr) Sex: 70% female Prevalence: 56% n = 50	Three EPs performed 47 of 50 scans. Machine: GE Logiq Pro 200 Probe: Phased array	RADUS
Davis 2005[18]	1°—Association between operator confidence and accuracy of EUS 2°—Accuracy of six EUS exams, including RUQ EUS to detect the presence of gallstones	American urban university–affiliated hospital receiving ~40,000 annual visits.	Convenience sample of patients receiving EUS for 1 yr after initiation of ED US training program Mean age: NR Sex: NR Prevalence: 59% n = 105	56 scans (53%) were performed by attending physicians and 49 by residents; none had previous experience Machine: Seimens Sonosite 180 Probe: NR	RADUS, CT, MRI, or surgical findings
Ha 2002[19]	1°–Agreement of EUS and RADUS for detecting gallstones, GBWT, PCF, CBDD gallbladder distention, and choledocholithiasis	Korean university–affiliated hospital receiving ~35,000 annual ED visits	Convenience sample of patients receiving RADUS for RUQ or epigastric pain Mean age: 54.5 yr (SD ±13.1 yr) Sex: 53% female Prevalence: 29% n = 59	All scans were performed by two PGY-3 EM residents with limited US experience Machine: Seimens Sonosite 180 Probe: 3.5-MHz curved	RADUS
Kendall 2001[20]	1°—Accuracy of EUS to detect gallstones, GBWT, CBDD, PCF, sludge, and air in gallbladder wall 2°—Accuracy of EUS vs. RADUS for detecting a SMS	American urban university–affiliated hospital with ~55,000 annual ED visits	Convenience sample of patients receiving RADUS for RUQ pain, epigastric pain, or jaundice. Mean age: 39 yr (range = 16–88 yr) Sex: 79% female Prevalence: 46% n = 109	Operators with 25 or fewer previous scans performed 51% (n = 57) of scans. 14 scans were performed by an operator with over 100 previous scans. Machine: Toshiba Capasee (n = 30), Toshiba 140A (n = 79) Probe: NR	RADUS for all findings excluding SMS SMS was compared to surgical findings
Miller 2006[21]	1°— Accuracy of EUS to detect gallstones, SMS, GBWT, and CBDD	American urban university–affiliated hospital with ~140,000 annual ED visits	Convenience sample of patients over 18 yr with RUQ or epigastric pain with biliary disease in the differential Mean age: 39.7 yr (SD ±14.6 yr) Sex: 79% female Prevalence: 80% n = 132	One experienced faculty member performed 43% of scans (n = 57); remainder performed by naive residents with basic training Machine: Treason 200 Probe: adjustable 3–5-MHz	RADUS
Rosen 2001[22]	1°—Accuracy of EUS to detect gallstones and cholecystitis 2°—Accuracy of EUS vs. RADUS for detection of acute cholecystitis	American urban university–affiliated hospital with ~66,000 annual ED visits	Convenience sample of patients over 18 receiving RADUS for RUQ or epigastric pain and suspicion of biliary colic Mean age: 49 yr (no range given) Sex: 72% female Prevalence: 60% n = 116	15 EPs enrolled patients. 30 scans (26%) were performed by operators with less than 25 previous RUQ scans Machine: Aloka Echo Camera SSD-500 or Siemens Sonoline Prima Probe: 3.5 MHz	RADUS for gallstones Cholecystitis results were compared with surgical findings

Table 2
(*Continued*)

Study	Outcomes	Center	Population	Bedside EUS	Reference Standard
Rowland 2001[23]	1°—Accuracy of nine different EUS exams, including RUQ scans to detect gallstones 2°—Correlation between operator confidence and image quality with accuracy	Australian urban tertiary care center with ~55,000 annual ED visits	Convenience sample of adults presenting to the ED with a number of specific conditions including RUQ pain and suspicion of gallstone disease Mean age: NR Sex: NR Prevalence: 46% $n = 35$	Inexperienced operators performed scans following a focused 3-day US course Machine: Aloka SD 1100 Probe: 3.5 or 5 MHz	RADUS or surgical findings
Summers 2010[24]	1°—Accuracy of EUS as well as RADUS for detection of acute cholecystitis 2° Accuracy of EUS for gallstones, GBWT, PCF, and SMS to detect cholecystitis	American urban university–affiliated hospital with ~36,000 annual ED visits	Convenience sample of patients over 18 yr presenting to the ED with suspected cholecystitis Median age: 36 yr (range = 18–87 yr) Sex: 73% female Prevalence: 62% $n = 91$ received RADUS $n = 24$ received CT $n = 115$ included in our analysis	43 EPs with a wide range of experience Machine: Sonosite Micromax or Toshiba Xario Probe: Phased array, large footprint curvilinear array, or microconvex array	Surgical findings or clinical follow-up*

CBDD = common bile duct dilation; EUS = emergency physician performed bedside emergency ultrasound; GBWT = gallbladder wall thickness; MRI = magnetic resonance imaging; NR = not reported; PCF = pericholecystic fluid; RADUS = radiology performed ultrasound; RUQ = right upper quadrant; SMS = sonographic Murphy's sign.
*Data comparing EUS to RADUS or CT for detecting gallstones were obtained through author contact and used in our analysis.

Table 3
Results of Quality Assessment

	Study							
	Alexander 2008[17]	Davis 2005[18]	Ha 2002[19]	Kendall 2001[20]	Miller 2006[21]	Rosen 2001[22]	Rowland 2001[23]	Summers 2010[24]
1. Representative spectrum?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Selection criteria described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. Quality reference standard?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4. Blinding of index test?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5. Blinding of reference test?	Yes	Unsure	Unsure	Yes	Yes	Yes	Yes	Yes
6. Did everyone receive the same reference standard?	Yes	No	Yes	Yes	Yes	Yes	No	No
7. Uninterpretable/intermediate results reported?	No	No	No	Yes	No	No	No	No

should consider alternative diagnoses in patients with a negative EUS (LR = 0.12) and low clinical suspicion for biliary colic. However, negative EUS results should be interpreted with caution in patients with a high clinical pretest probability.

The presence of gallstones may be an incidental finding in patients with RUQ pain of a different origin. Although we discovered that EUS has a specificity of 88.0%, a positive EUS exam for gallstones cannot in itself make the diagnosis of cholelithiasis.

The presence of gallstones is the primary sonographic criteria in diagnosing acute cholecystitis,[37,38] and approximately 90% to 95% of patients with acute cholecystitis have gallstones.[38,39] Furthermore, acalculous cholecystitis is most prevalent in the ICU and is usually associated with specific risk factors such as burns, severe multisystem trauma, or total parenteral nutrition.[39] However, to establish the sonographic diagnosis of acute cholecystitis, secondary signs such as pericholecystic fluid, thickened gallbladder wall, or a sonographic Murphy's sign should be elicited.[25] Determining the accuracy of EUS to detect secondary signs of cholecystitis is beyond the scope of this review and remains an area for further investigation.
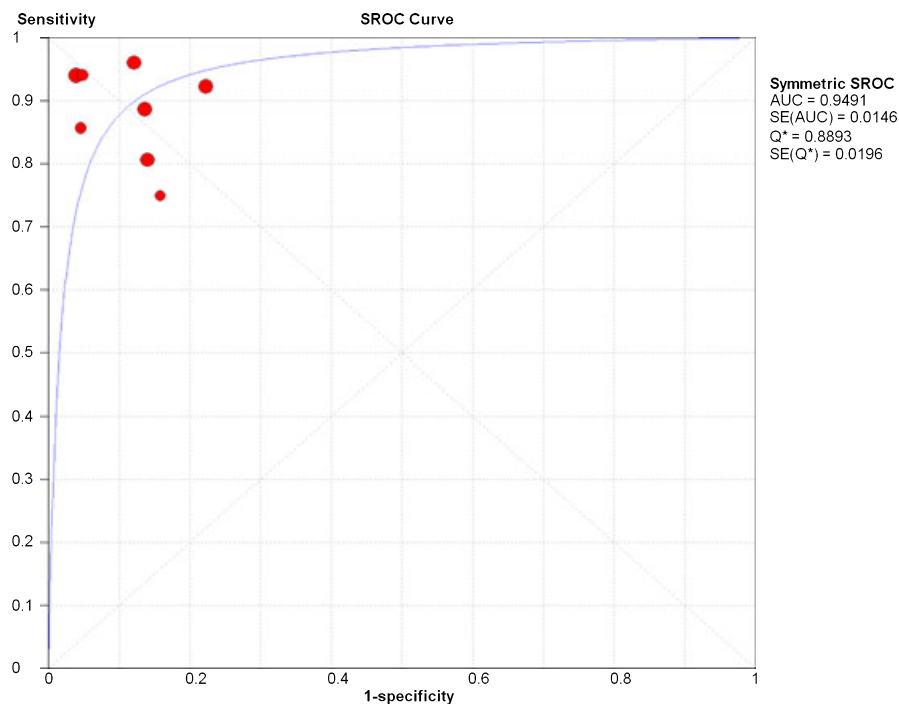
Figure 3. Summary receiver operator characteristics (SROC) curve. Black circles represent individual studies.

LIMITATIONS

Over 80% of patients included in our analysis underwent RADUS as a criterion reference standard. To some extent, the SROC curve represents agreement data as opposed to true diagnostic test performance. A systematic review found the sensitivity and specificity of RADUS to be 97 and 95%, respectively.[40] However, after accounting for a partial verification bias, these values are reported as 84 and 99%. The authors admit that the best estimate probably lies between the adjusted and unadjusted values.

The reference standards used in each study could be challenged, as it is unknown how many patients with cholelithiasis were missed by both EUS and the criterion reference. It is possible that the number of false-negative studies was therefore underestimated.

A differential verification bias may be present, as data from three articles used more than one reference standard. If the results of the index test in any way influenced the selection of reference standard, we would expect an overestimation of accuracy.

That all eight studies used convenience sampling suggests a potential for selection bias. For example, corpulent subjects may have been selectively excluded from participation. If publication bias was present, our estimates for test sensitivity and specificity may be inflated.

Ultrasound is an operator-dependent test. The performance of RUQ US is difficult to standardize and is not necessarily uniform. The level of operator experience and training is considered an important determinant of test accuracy and was highly variable within the included studies (Table 2). Although we planned a subgroup analysis based on operator experience, we were unable to address this question due to the limited number and size of the included studies. None of the eight included studies sufficiently described their scanning technique. The variation of technical ability between operators may be a significant source of clinical heterogeneity and may have distorted our pooled estimates of sensitivity and specificity.

CONCLUSIONS

This systematic review suggests that emergency bedside ultrasound is a useful adjunct for the diagnosis or exclusion of cholelithiasis in the ED. Based on the results of eight studies ($n = 710$) of variable quality, the estimates for emergency bedside ultrasound sensitivity and specificity for cholelithiasis are 89.8 and 88.0%, respectively. In patients presenting to the ED with a high pretest probability of symptomatic cholelithiasis, a positive emergency bedside ultrasound scan (likelihood ratio = 7.5) may be used to confirm the diagnosis and arrange for appropriate outpatient follow-up if symptoms have resolved. In patients with a low pretest probability, a negative emergency bedside ultrasound scan (likelihood ratio = 0.12) should prompt the clinician to consider an alternative diagnosis or further diagnostic testing.

References

1. Everhart JE, Khare M, Hill M, Maurer KR. Prevalence and ethnic differences in gallbladder disease in the United States. Gastroenterology. 1999; 117:632–9.
2. Attili AF, De Santis A, Capri R. The natural history of gallstones: the GREPCO experience. The GREPCO Group. Hepatology. 1995; 21:655–60.

3. Wenckert A, Robertson B. The natural course of gallstone disease: eleven year review of 781 nonoperated cases. Gastroenterology. 1966; 50:376–81.

4. Barbara L, Sama C, Morselli Labate AM. A population study on the prevalence of gallstone disease: the Sirmione Study. Hepatology. 1987; 7:913–7.

5. Friedman GS, Raviola CA, Fireman B. Prognosis of gallstones with mild or no symptoms: 25 years of follow-up in a health maintenance organization. J Clin Epidemiol. 1989; 42:127–36.

6. Trowbridge RL, Rutkowski NK, Shojania KG. Does this patient have acute cholecystitis? JAMA. 2003; 289:80–6.

7. Kell MR, Aherne NJ, Coffey C, Power CP, Kirwan WO, Redmond HP. Emergency surgeon- performed hepatobiliary ultrasonography. Br J Surg. 2002; 89:1402–4.

8. Blaivas M, Harwood RA, Lambert MJ. Decreasing length of stay with emergency ultrasound examination of the gallbladder. Acad Emerg Med. 1999; 6:1020–3.

9. American College of Emergency Physicians. Emergency ultrasound guidelines. Ann Emerg Med. 2009; 53:550–70.

10. Mateer J, Plummer D, Heller M, et al. Model curriculum for physician training in emergency ultrasound. Ann Emerg Med. 1994; 23:95–102.

11. Bektas F, Eken C, Soyuncu S, Kusoglu L, Cete Y. Contribution of goal-directed ultrasonography to clinical decision-making for emergency physicians. Emerg Med J. 2009; 26:169–72.

12. Roy S. Hepatobiliary. In: Ma OJ, Mateer JR (eds). Emergency Ultrasound. New York, NY: McGraw-Hill Professionals, 2002.

13. Dent B, Kendall RJ, Boyle AA, Atkinson PR. Emergency ultrasound of the abdominal aorta by UK emergency physicians: a prospective cohort study. Emerg Med J. 2007; 24:547–9.

14. Mandavia DP, Hoffner RJ, Mahaney K, Henderson SO. Bedside echocardiography by emergency physicians. Ann Emerg Med. 2001; 38:377–82.

15. Ross M, McLaughlin K, Lang E, Thompson J, Brown M, Atkinson P, Clark S. Protocol for Emergency Physician Performed Ultrasonography to Diagnose Cholelithiasis: A Systematic Review. Available from: http://docs.google.com/Doc?docid=0AZyT1p1lltlDZG dzZDl2c3RfMWZ3cDN3cWc2&hl=en. Accessed April 10, 2010.

16. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009; 339:b2535.

17. Alexander DN, Ragg M, Stella J. Emergency department ultrasound for the investigation of right upper quadrant abdominal pain. Emerg Med Australasia. 2008; 20(Suppl 1):A21.

18. Davis DP, Campbell CJ, Poste JC, Ma G. The association between operator confidence and accuracy of ultrasonography performed by novice emergency physicians. J Emerg Med. 2005; 23:259–64.

19. Ha YR, Kim H, Yoo S, Chung SP, Kim SH, Yoo IS. Accuracy of emergency ultrasonography for biliary parameters by physicians with limited training. J Korean Soc Emerg Med. 2002; 13:407–10.

20. Kendall JL, Shrimp RJ. Performance and interpretation of focused right upper quadrant ultrasound by emergency physicians. J Emerg Med. 2001; 21:7–13.

21. Miller AH, Pepe PE, Brockman CR, Delaney KA. ED ultrasound in hepatobiliary disease. J Emerg Med. 2006; 30:69–74.

22. Rosen CL, Brown DF, Chang Y, et al. Ultrasonography by emergency physicians in patients with suspected cholecystitis. Am J Emerg Med. 2001; 19:32–6.

23. Rowland JL, Kuhn M, Bonnin RL, Davey MJ, Langlois SL. Accuracy of emergency department bedside ultrasonography. Emerg Med. 2001; 13:305–13.

24. Summers SM, Scruggs W, Menchine MD, et al. A prospective evaluation of emergency department bedside ultrasonography for the detection of acute cholecystitis. Ann Emerg Med. 2010; 56:114–22.

25. Shah K, Wolf RE. Hepatobiliary ultrasound. Emerg Med Clin North Am. 2004; 22:661–73.

26. Hintz KK, Jones JS. BET 3: emergency physician bedside ultrasound for the diagnosis of cholelithiasis. Emerg Med J. 2009; 26:667–9.

27. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM; Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. Ann Intern Med. 2008; 149:889–97.

28. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003; 3:25.

29. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med. 1993; 12:1293–316.

30. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol. 1995; 48:119–30.

31. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7:177–88.

32. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Method. 2006; 6:31–43.

33. Tatsioni A, Zarin DA, Aronson N, et al. Challenges in systematic reviews of diagnostic technologies. Ann Intern Med. 2005; 142(12 pt 2):1048–55.

34. Terrin N, Schmid CH, Lau J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. J Clin Epidemiol. 2005; 58:894–901.

35. Halldestam I, Enell EL, Kullman E, Borch K. Development of symptoms and complications in individuals with asymptomatic gallstones. Br J Surg. 2004; 91:734–8.

36. Durston W, Carl ML, Guerra W, et al. Comparison of quality and cost-effectiveness in the evaluation of symptomatic cholelithiasis with different approaches to ultrasound availability in the ED. Am J Emerg Med. 2001; 19:260–9.

37. Cooperberg PL, Gibney RG. Imaging of the gall-bladder 1997. Radiology. 1987; 163:605–13.
38. Ralls PW, Colletti PM, Lapin SA, et al. Real-time sonography in suspected acute cholecystitis: a prospective evaluation of primary and secondary signs. Radiology. 1985; 155:767–71.
39. Hanbidge AE, Buckler PM, O'Malley ME, Wilson SR. Imaging evaluation for acute pain in the right upper quadrant. Radiographics. 2004; 24:1117–35.
40. Shea JA, Berlin JA, Escarce JJ, et al. Revised estimates of diagnostic test sensitivity and specificity in suspected biliary tract disease. Arch Intern Med. 1994; 154:2573–81.

Call for *Academic Emergency Medicine* (*AEM*) Reviewers

Deadline: March 17, 2011

Members of the Society for Academic Emergency Medicine (SAEM) are invited to submit nominations to serve as peer-reviewers for *Academic Emergency Medicine (AEM)*. As an indicator of familiarity with the peer-review process, the medical literature, and the research process in general, peer-reviewers are expected to have published at least two peer-reviewed papers in the medical literature as first or second author. Some of these papers should be original research work. Other scholarly work or experience will be considered as evidence of expertise (i.e., informatics experience demonstrated by network/database/desktop development).

AEM peer-reviewers are invited to review manuscripts based on their area(s) of expertise. Once a reviewer has accepted an invitation to review a manuscript, the reviewer is expected to complete the review within 14 days of receipt of the electronic version of the manuscript.

To provide feedback, reviewers receive the consensus review from each manuscript that they review. In addition, each review is evaluated by the Associate Editor in the areas of timeliness, assessment of manuscript strengths and weaknesses, constructive suggestions, summarizing major issues and concerns, and overall quality of the review. Scores are compiled in the *AEM* database. Each year the Editor-in-Chief designates "outstanding reviewers" for public acknowledgement of excellent contributions to the peer-review process. Reviewers who consistently fail to respond to requests to review, who are unavailable to perform reviews, or who submit late or incomplete reviews may be dropped from the peer-reviewer database at any time, at the discretion of the Editor-in-Chief.

Individuals interested in being considered for appointment as an *AEM* peer-reviewer must send a letter of interest including areas of expertise as defined on the reviewer topic survey and a current CV. The reviewer topic survey can be found at: www.saem.org/inform/resurvey.html. Most peer-reviewer appointments are for three years. **Applications must be submitted electronically to: aem@saem.org by March 17, 2011.**