



Cochrane
Library

Cochrane Database of Systematic Reviews

Ultrasound versus liver function tests for diagnosis of common bile duct stones (Review)

Gurusamy KS, Giljaca V, Takwoingi Y, Higgle D, Poropat G, Štimac D, Davidson BR

Gurusamy KS, Giljaca V, Takwoingi Y, Higgle D, Poropat G, Štimac D, Davidson BR.
Ultrasound versus liver function tests for diagnosis of common bile duct stones.
Cochrane Database of Systematic Reviews 2015, Issue 2. Art. No.: CD011548.
DOI: [10.1002/14651858.CD011548](https://doi.org/10.1002/14651858.CD011548).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	4
BACKGROUND	6
Figure 1.	7
OBJECTIVES	8
METHODS	9
RESULTS	10
Figure 2.	11
Figure 3.	12
Figure 4.	13
Figure 5.	14
Figure 6.	15
DISCUSSION	16
AUTHORS' CONCLUSIONS	17
ACKNOWLEDGEMENTS	17
REFERENCES	19
CHARACTERISTICS OF STUDIES	29
DATA	43
Test 1. Ultrasound.	43
Test 2. Bilirubin (> 22.23 µmol/L).	43
Test 3. Bilirubin (> twice the normal limit).	43
Test 4. Alkaline phosphatase (> 125 IU/L).	43
Test 5. Alkaline phosphatase (> twice the normal limit).	43
ADDITIONAL TABLES	43
APPENDICES	46
CONTRIBUTIONS OF AUTHORS	48
DECLARATIONS OF INTEREST	48
SOURCES OF SUPPORT	48
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	48
NOTES	48
INDEX TERMS	48

[Diagnostic Test Accuracy Review]

Ultrasound versus liver function tests for diagnosis of common bile duct stones

Kurinchi Selvan Gurusamy¹, Vanja Giljaca², Yemisi Takwoingi³, David Higgle⁴, Goran Poropat², Davor Štimac², Brian R Davidson¹

¹Department of Surgery, Royal Free Campus, UCL Medical School, London, UK. ²Department of Gastroenterology, Clinical Hospital Centre Rijeka, Rijeka, Croatia. ³Public Health, Epidemiology and Biostatistics, University of Birmingham, Birmingham, UK. ⁴North Bristol NHS Trust, Bristol, UK

Contact address: Kurinchi Selvan Gurusamy, Department of Surgery, Royal Free Campus, UCL Medical School, Royal Free Hospital, Rowland Hill Street, London, NW3 2PF, UK. k.gurusamy@ucl.ac.uk.

Editorial group: Cochrane Hepato-Biliary Group.

Publication status and date: New, published in Issue 2, 2015.

Citation: Gurusamy KS, Giljaca V, Takwoingi Y, Higgle D, Poropat G, Štimac D, Davidson BR. Ultrasound versus liver function tests for diagnosis of common bile duct stones. *Cochrane Database of Systematic Reviews* 2015, Issue 2. Art. No.: CD011548. DOI: [10.1002/14651858.CD011548](https://doi.org/10.1002/14651858.CD011548).

Copyright © 2015 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Ultrasound and liver function tests (serum bilirubin and serum alkaline phosphatase) are used as screening tests for the diagnosis of common bile duct stones in people suspected of having common bile duct stones. There has been no systematic review of the diagnostic accuracy of ultrasound and liver function tests.

Objectives

To determine and compare the accuracy of ultrasound versus liver function tests for the diagnosis of common bile duct stones.

Search methods

We searched MEDLINE, EMBASE, Science Citation Index Expanded, BIOSIS, and Clinicaltrials.gov to September 2012. We searched the references of included studies to identify further studies and systematic reviews identified from various databases (Database of Abstracts of Reviews of Effects, Health Technology Assessment, Medion, and ARIF (Aggressive Research Intelligence Facility)). We did not restrict studies based on language or publication status, or whether data were collected prospectively or retrospectively.

Selection criteria

We included studies that provided the number of true positives, false positives, false negatives, and true negatives for ultrasound, serum bilirubin, or serum alkaline phosphatase. We only accepted studies that confirmed the presence of common bile duct stones by extraction of the stones (irrespective of whether this was done by surgical or endoscopic methods) for a positive test result, and absence of common bile duct stones by surgical or endoscopic negative exploration of the common bile duct, or symptom-free follow-up for at least six months for a negative test result as the reference standard in people suspected of having common bile duct stones. We included participants with or without prior diagnosis of cholelithiasis; with or without symptoms and complications of common bile duct stones, with or without prior treatment for common bile duct stones; and before or after cholecystectomy. At least two authors screened abstracts and selected studies for inclusion independently.

Data collection and analysis

Two authors independently collected data from each study. Where meta-analysis was possible, we used the bivariate model to summarise sensitivity and specificity.

Main results

Five studies including 523 participants reported the diagnostic accuracy of ultrasound. One studies (262 participants) compared the accuracy of ultrasound, serum bilirubin and serum alkaline phosphatase in the same participants. All the studies included people with symptoms. One study included only participants without previous cholecystectomy but this information was not available from the remaining studies. All the studies were of poor methodological quality. The sensitivities for ultrasound ranged from 0.32 to 1.00, and the specificities ranged from 0.77 to 0.97. The summary sensitivity was 0.73 (95% CI 0.44 to 0.90) and the specificity was 0.91 (95% CI 0.84 to 0.95). At the median pre-test probability of common bile duct stones of 0.408, the post-test probability (95% CI) associated with positive ultrasound tests was 0.85 (95% CI 0.75 to 0.91), and negative ultrasound tests was 0.17 (95% CI 0.08 to 0.33).

The single study of liver function tests reported diagnostic accuracy at two cut-offs for bilirubin (greater than 22.23 $\mu\text{mol/L}$ and greater than twice the normal limit) and two cut-offs for alkaline phosphatase (greater than 125 IU/L and greater than twice the normal limit). This study also assessed ultrasound and reported higher sensitivities for bilirubin and alkaline phosphatase at both cut-offs but the specificities of the markers were higher at only the greater than twice the normal limit cut-off. The sensitivity for ultrasound was 0.32 (95% CI 0.15 to 0.54), bilirubin (cut-off greater than 22.23 $\mu\text{mol/L}$) was 0.84 (95% CI 0.64 to 0.95), and alkaline phosphatase (cut-off greater than 125 IU/L) was 0.92 (95% CI 0.74 to 0.99). The specificity for ultrasound was 0.95 (95% CI 0.91 to 0.97), bilirubin (cut-off greater than 22.23 $\mu\text{mol/L}$) was 0.91 (95% CI 0.86 to 0.94), and alkaline phosphatase (cut-off greater than 125 IU/L) was 0.79 (95% CI 0.74 to 0.84). No study reported the diagnostic accuracy of a combination of bilirubin and alkaline phosphatase, or combinations with ultrasound.

Authors' conclusions

Many people may have common bile duct stones in spite of having a negative ultrasound or liver function test. Such people may have to be re-tested with other modalities if the clinical suspicion of common bile duct stones is very high because of their symptoms. False-positive results are also possible and further non-invasive testing is recommended to confirm common bile duct stones to avoid the risks of invasive testing.

It should be noted that these results were based on few studies of poor methodological quality and the results for ultrasound varied considerably between studies. Therefore, the results should be interpreted with caution. Further studies of high methodological quality are necessary to determine the diagnostic accuracy of ultrasound and liver function tests.

PLAIN LANGUAGE SUMMARY

Ultrasound and liver function tests for the diagnosis of common bile duct stones

Background

Bile, produced in the liver and stored temporarily in the gallbladder, is released into the small bowel on eating fatty food. The common bile duct is the tube through which bile flows from the gallbladder to the small bowel. Stones in the common bile duct (common bile duct stones), usually formed in the gallbladder before migration into the bile duct, can obstruct the flow of bile leading to jaundice (yellowish discolouration of skin, white of the eyes, and dark urine); infection of the bile (cholangitis); and inflammation of the pancreas (pancreatitis), which can be life threatening. Various diagnostic tests can be performed for the diagnosis of common bile duct stones. Depending upon the availability of resources, these stones are removed endoscopically (usually the case) or may be removed as a part of the operation performed to remove the gallbladder (it is important to remove the gallbladder since the stones continue to form in the gallbladder and can cause recurrent problems). Non-invasive tests such as ultrasound (use of sound waves higher than audible range to differentiate tissues based on how they reflect the sound waves) and blood markers of bile flow obstruction such as serum bilirubin and serum alkaline phosphatase are used to identify people at high risk of having common bile duct stones. Using non-invasive tests means that only those people at high risk can be subjected to further tests. We reviewed the evidence on the accuracy of ultrasound and liver function tests for detection of common bile duct stones. The evidence is current to September 2012.

Study characteristics

We identified five studies including 523 participants that reported the diagnostic test accuracy of ultrasound. One of these studies, involving 262 participants, also reported the diagnostic test accuracy of serum bilirubin and serum alkaline phosphatase. All the studies included people with symptoms. One study included only participants who had not undergone previous cholecystectomy (removal of gallbladder). This information was not available from the remaining studies.

Key results

Based on an average sensitivity of 73% for ultrasound, we would expect that on average 73 out of 100 people with common bile duct stones will be detected while the remaining 27 people will be missed and will not receive appropriate treatment. The average number of people with common bile duct stones detected using ultrasound may vary between 44 and 90 out of 100 people. Based on an average specificity of 91% for ultrasound, we would expect that on average 91 out of 100 people without common bile duct stones would be identified as not having common bile duct stones; 9 out of 100 would be false positives and not receive appropriate treatment. The average number of false positives could vary between 5 and 16 out of 100 people.

Evidence from one study suggested that using a level of serum alkaline phosphatase higher than 125 units to distinguish between people who have and people who do not have common bile duct stones gave better diagnostic accuracy than using a level twice the normal limit (which usually ranges between 0 and 40). The study also showed better accuracy for serum alkaline phosphatase compared to serum bilirubin.

The sensitivity of serum alkaline phosphatase at the 125 units cut-off was 92%, which means that 92 out of 100 people with common bile duct stones would be detected but 8 out of 100 people will be missed. The number detected could vary between 74 and 99 out of 100 people. Based on the specificity of 79%, 79 out of 100 people without common bile duct stones will be correctly identified as not having common bile duct stones while the remaining 21 people will be false positives. The number of false positives could vary between 16 and 26 out of 100 people. This suggests that further non-invasive tests may be useful to diagnose common bile duct stones prior to the use of invasive tests.

Quality of evidence

All the studies were of low methodological quality, which may undermine the validity of our findings.

Future research

Further studies of high methodological quality are necessary.

SUMMARY OF FINDINGS

Summary of findings 1. Performance of ultrasound and liver function tests for diagnosis of common bile stones

Population	People suspected of having common bile duct stones.				
Settings	Secondary and tertiary care setting in different parts of the world.				
Index tests	Ultrasound and liver function tests (bilirubin and alkaline phosphatase). For bilirubin, the cut-offs used to define test positivity were > 22.23 µmol/L and > twice the normal limit. For alkaline phosphatase, the cut-offs were > 125 IU/L and > twice the normal limit.				
Reference standard	Endoscopic or surgical extraction of stones in people with a positive index test result or clinical follow-up (minimum 6 months) in people with a negative index test result.				
Target condition	Common bile duct stones.				
Number of studies	5 studies (162 cases, 523 participants) evaluated ultrasound. 1 of these studies (25 cases, 262 participants) also evaluated bilirubin and alkaline phosphatase.				
Methodological quality concerns	All the studies were of poor methodological quality; most studies were at high risk of bias or gave high concern about applicability across all domains of quality assessment, or both.				
Test (cut-off)	Summary sensitivity (95% CI)¹	Summary specificity (95% CI)¹	Pre-test probability²	Positive post-test probability (95% CI)³	Negative post-test probability (95% CI)⁴
Ultrasound	0.73 (0.44 to 0.90)	0.91 (0.84 to 0.95)	0.095	0.45 (0.31 to 0.60)	0.03 (0.01 to 0.07)
			0.408	0.85 (0.75 to 0.91)	0.17 (0.08 to 0.33)
			0.658	0.94 (0.89 to 0.97)	0.37 (0.20 to 0.58)
Bilirubin (> 22.23 µmol/L)	0.84 (0.64 to 0.94)	0.91 (0.86 to 0.94)	0.095	0.49 (0.38 to 0.59)	0.02 (0.01 to 0.04)
Bilirubin (> twice the normal limit)	0.42 (0.22 to 0.63)	0.97 (0.95 to 0.99)	0.095	0.63 (0.41 to 0.81)	0.06 (0.04 to 0.08)
Alkaline phosphatase (> 125 IU/L)	0.92 (0.74 to 0.99)	0.79 (0.74 to 0.84)	0.095	0.32 (0.26 to 0.38)	0.01 (0.00 to 0.04)
Alkaline phosphatase (> twice the normal limit)	0.38 (0.19 to 0.59)	0.97 (0.95 to 0.99)	0.095	0.61 (0.38 to 0.80)	0.06 (0.05 to 0.08)

Interpretation of results:

For ultrasound, at a pre-test probability of 9.5%, out of 100 people with a positive result, common bile duct stones would be present in 45 people, at a pre-test probability of 40.8%, they would be present in 85 people, and at a pre-test probability of 65.8%, they would be present in 94 people.

For ultrasound, at a pre-test probability of 9.5%, out of 100 people with a negative result, common bile duct stones would be present in 3 people, at a pre-test probability of 40.8%, they would be present in 17 people, and at a pre-test probability of 65.8%, they would be present in 37 people.

For bilirubin, at a pre-test probability of 9.5%, out of 100 people with a positive result at a cut-off above 22.23 $\mu\text{mol/L}$, common bile duct stones would be present in 49 people, and out of 100 people with a negative result, common bile duct stones would be present in 2 people.

For alkaline phosphatase, at a pre-test probability of 9.5%, out of 100 people with a positive result a cut-off above 125 IU/L, common bile duct stones would be present in 32 people, and out of 100 people with a negative result, common bile duct stones would be present in 2 people.

Comparison of the diagnostic accuracy of ultrasound and liver function tests: because only 1 study reported the diagnostic accuracy of liver function tests, it was not possible to compare test accuracy in a meta-analysis. The study compared ultrasound, bilirubin, and alkaline phosphatase and reported higher sensitivities for bilirubin and alkaline phosphatase than ultrasound at both cut-offs but the specificities of the markers were higher at only the cut-off that was twice the normal limit.

Conclusions: a negative ultrasound or liver function result may lead to many people with common bile duct stones being missed and further testing is needed in people with a high suspicion of the condition. However, the strength of evidence on the accuracy of the tests was very weak because it was based on few and methodologically flawed studies.

¹Summary sensitivity and specificity for ultrasound only. Only 1 study evaluated bilirubin and alkaline phosphatase, so there are no pooled estimates.

²The pre-test probability (proportion with common bile duct stones out of the total number of participants) was computed for each included study. For ultrasound, these numbers represented the minimum, median, and maximum values from the 5 studies. For bilirubin and alkaline phosphatase, the minimum pre-test probability was also the pre-test probability in the 1 study that evaluated the tests.

³Post-test probability of common bile duct stones in people with positive index test results.

⁴Post-test probability of common bile duct stones in people with negative index test results.

BACKGROUND

Biliary stones are conglomerates of precipitated bile salts that form in the gallbladder or the common bile duct. The common bile duct carries bile from the liver to the duodenum (first part of the small intestine). The term 'gallstones' generally refer to the stones in the gallbladder while 'common bile duct stones' refer to stones in the common bile duct. Common bile duct stones may form inside the common bile duct (primary common bile duct stones), or they may form in the gallbladder and migrate to the common bile duct (secondary common bile duct stones) (Williams 2008). A significant proportion of people presenting with common bile duct stones may be asymptomatic (Sarli 2000). In some people, the stones pass silently into the duodenum, and in other people, the stones cause clinical symptoms such as biliary colic, jaundice, cholangitis, or pancreatitis (Caddy 2006). The prevalence of gallstone disease in the general population is about 6% to 15% with a higher prevalence in females (Barbara 1987; Loria 1994). Only 2% to 4% of people with gallstones become symptomatic with biliary colic (pain), acute cholecystitis (inflammation), obstructive jaundice, or gallstone pancreatitis in one year (Attili 1995; Halldestam 2004), and removal of gallbladder is recommended in people with symptomatic gallstones (Gurusamy 2010). Among people who undergo laparoscopic cholecystectomy (removal of gallbladder) for symptomatic gallstones, 3% to 22% also have concomitant common bile duct stones (Arnold 1970; Lill 2010; Yousefpour Azary 2011).

Common bile duct stones present in multiple ways. Central and right-sided upper abdominal pain is a common presentation (Anciaux 1986; Roston 1997). Jaundice, caused by an impacted stone in the common bile duct leading to obstruction of bile passage into the duodenum, is another presentation. It may subsequently resolve if the common bile duct stone passes spontaneously into the duodenum. This happens in 54% to 73% of people with common bile duct stones in whom cholecystectomy is performed for gallstones (Tranter 2003; Lefemine 2011). Another, more dangerous, complication of common bile duct stones is acute cholangitis. Cholangitis is clinically defined by Charcot's triad, which includes elevated body temperature, pain under the right ribcage, and jaundice (Raraty 1998; Salek 2009). Acute cholangitis is caused by an ascending bacterial infection of the common bile duct and the biliary tree along with biliary obstruction. This complication is present in 2% to 9% of people admitted for gallstone disease (Saik 1975; Tranter 2003), and a mortality of approximately 24% is recorded (Salek 2009). Common bile duct stones may also cause acute pancreatitis, accounting for 33% to 50% of all people with acute pancreatitis (Corfield 1985; Toh 2000). Acute pancreatitis is usually a self limiting disease and is usually sufficiently treated by conservative measures in its mild form (Neoptolemos 1988). However, a more severe pancreatitis may evolve in approximately 27% to 37% of people with common bile duct stone-induced pancreatitis, with mortality around 6% to 9% (Mann 1994; Toh 2000).

Suspicion of common bile duct stones can be confirmed by laboratory liver function tests (Barkun 1994), or imaging tests such as abdominal ultrasound (Ripolles 2009). Further testing may include endoscopic ultrasound (EUS) (Aljebreen 2008), magnetic resonance cholangiopancreatography (MRCP) (Stiris 2000), endoscopic retrograde cholangiopancreatography (ERCP) (Geron 1999), and intraoperative cholangiography (IOC) (Fiore

1997). Currently, these are the recommended tests for diagnosis of common bile duct stones and of these tests, IOC can only be done during an operation as the test requires surgical cannulation of the common bile duct during cholecystectomy. The other tests may be used preoperatively or postoperatively. Usually the first diagnostic tests that most people will undergo are liver function tests and abdominal ultrasound. While ultrasound findings about common bile duct stones are available as a part of the test that the person undergoes for the diagnosis of gallstones, liver function tests may be routinely used to screen people with gallstones to identify people who need further testing. Invasive diagnostic tests are usually reserved for people with suspected common bile duct stones based on non-invasive diagnostic tests, or when therapeutic measures are necessary (Freitas 2006).

There are other tests such as conventional computed tomogram, computed tomogram cholangiogram, laparoscopic ultrasound, and ERCP-guided intraductal ultrasound used for diagnosing common bile duct stones, but these are of limited use (Maple 2010).

Target condition being diagnosed

Common bile duct stones. We did not differentiate the target condition with respect to common bile duct stone size, degree of common bile duct obstruction, and the presence or absence of symptoms.

Index test(s)

Liver function tests measure different biochemical reactions in blood drawn from the person. Common bile duct stones are usually suspected in the presence of elevated gamma glutamyltransferase, alkaline phosphatase (ALP), and bilirubin levels. Liver function tests are usually the first laboratory tests that a person, whether asymptomatic or symptomatic, would undergo (along with abdominal ultrasound). Elevated bilirubin, gamma glutamyltransferase, or ALP raises suspicion of an obstruction of the common bile duct with a stone, tumour, or inflammatory changes (Barkun 1994; Giannini 2005). The normal values of these tests vary in different laboratories and are generally considered as the mean and 95% confidence interval (CI) in a normal reference population (Giannini 2005). A bilirubin value more than 1 mg/dL (greater than 17.1 µmol/L) is generally considered abnormal (FDA 2013). The mean ALP value in an adult population is between 50 and 170 IU/L (Eastman 1977). Gamma glutamyltransferase can be elevated because of alcohol abuse and drugs such as phenytoin and barbiturates (Rosalki 1971; Giannini 2005), and hence is of limited value in the diagnosis of obstruction to the flow of bile. Transaminases may also be elevated in the presence of obstructive jaundice due to common bile duct stones (Hayat 2005), but they may also be elevated in liver parenchymal damage. We fully acknowledge that even bilirubin and ALP can be elevated because of other reasons, but these are common tests used as triage tests for diagnosis of common bile stones particularly in people with gallstones. In this review, we assessed serum bilirubin and ALP as the index tests. We planned to investigate a test strategy that included abnormal values for either test or abnormal values for both tests.

Abdominal ultrasound uses sound waves of high frequency to visualise tissues and structures located in the abdomen, including the gallbladder and the common bile duct. The ultrasound probe is placed on the skin of the abdomen, using gel to enhance visibility.

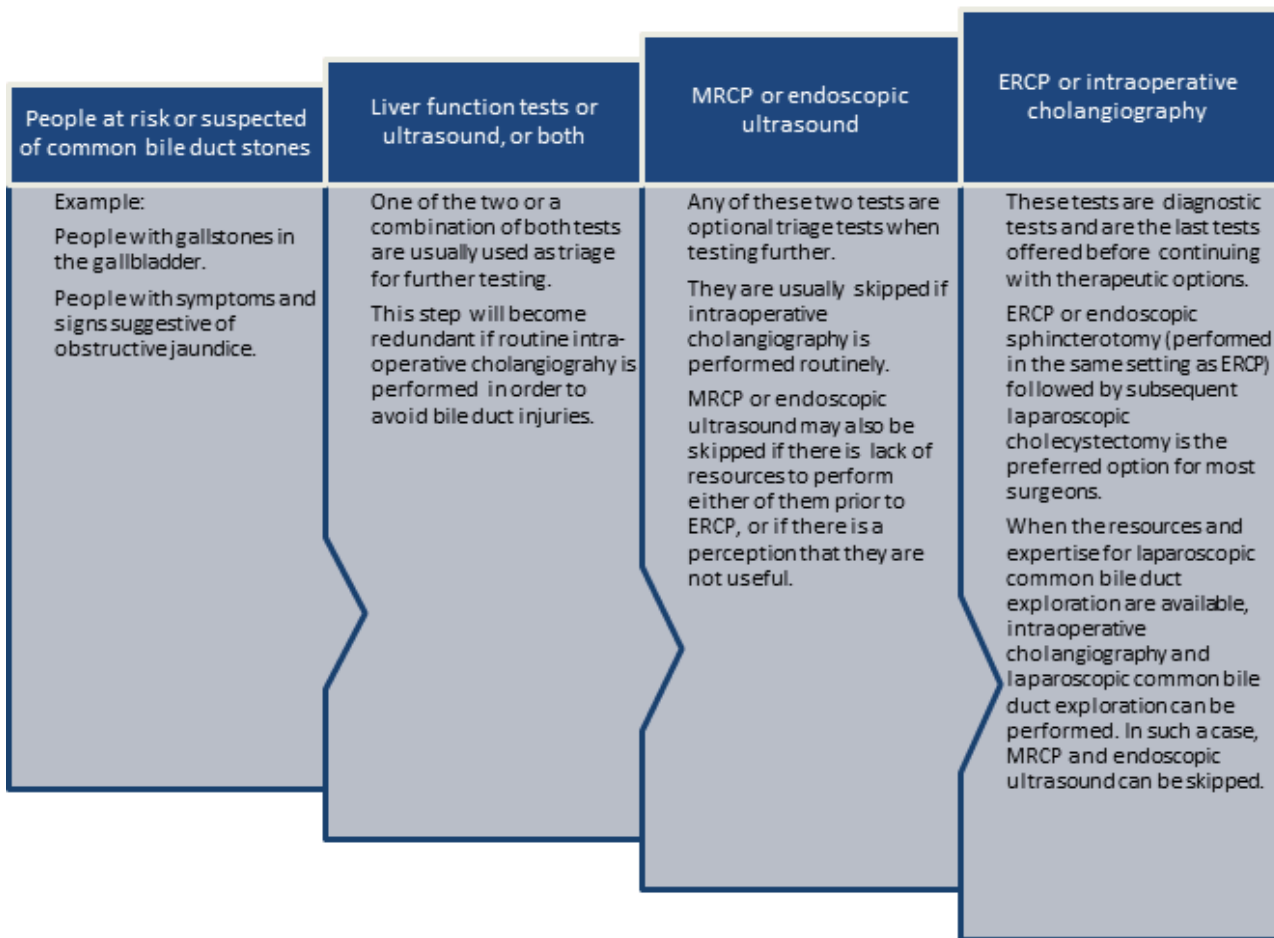
A test is considered positive when a hyperechoic round or oval structure is seen within the common bile duct (Ripolles 2009; RadiologyInfo 2011).

Clinical pathway

Figure 1 shows a diagnostic pathway. People that are at risk of having common bile duct stones or suspected of having common bile duct stones (such as people with gallbladder stones or people that show symptoms and signs of obstructive jaundice or pancreatitis) will undergo liver function tests and abdominal ultrasound as the first step. An abdominal ultrasound is usually available by the time the person is at risk or suspected of having common bile duct stones. Usually a combination of both tests is used as triage tests before further testing is done in the second step, but these can be used as the definitive diagnostic test to carry out a therapeutic option directly (e.g., endoscopic or surgical common bile duct exploration) (Williams 2008; ASGE Standards of Practice Committee 2010). MRCP or EUS are tests in the second step of the diagnostic pathway and are used as optional triage tests prior to the tests used in the third step of the diagnostic

pathway, but can also be used as definitive diagnostic tests to carry out a therapeutic option directly. MRCP and EUS are not usually combined, since the positive or negative results of one or the other is usually accepted for further clinical decision making, without taking into consideration the results of liver function tests or transabdominal ultrasound, as it is generally believed that MRCP and EUS have better diagnostic accuracy than liver function tests or transabdominal ultrasound. ERCP and IOC are used in the third step of the diagnostic pathway. Both tests are done just before the therapeutic intervention. Therapeutic interventions, such as endoscopic or surgical stone extraction, can then be undertaken during the same session. ERCP is done before endoscopic sphincterotomy and removal of common bile duct stones using Dormia basket or balloon during the same endoscopic session (Prat 1996; Maple 2010), and IOC is done before surgical common bile duct exploration and removal of common bile duct stones using surgical instruments during operation for cholecystectomy (Targarona 2004; Freitas 2006; Chen 2007; Williams 2008; ASGE Standards of Practice Committee 2010; Kelly 2010).

Figure 1. Diagnostic pathway for diagnosis of common bile duct stones. Note that ultrasound is generally performed in all people at risk or suspected of common bile duct stones. ERCP: endoscopic retrograde cholangiopancreatography; MRCP: magnetic resonance cholangiopancreatography.



Ultrasound and liver function tests are considered as triage tests before the person undergoes further investigation using MRCP, EUS, ERCP, or IOC and are generally performed in all people in whom common bile duct stones are suspected.

Implications of negative test results

In general, people with negative test results in one step do not undergo further testing. For example, a person with no suggestion of common bile duct stones on liver function tests and ultrasound will not undergo further testing for common bile duct stones. Similarly, people with no suggestion of common bile duct stones on MRCP or EUS will not undergo further testing for common bile duct stones and people with no suggestion of common bile duct stones on ERCP or IOC will not undergo common bile duct clearance. People with a false-negative test result can develop life-threatening complications of common bile duct stones, such as cholangitis and pancreatitis, but the natural history of such people in terms of the frequency with which these complications develop is unknown. However, it is generally recommended that common bile duct stones are removed when they are identified because of the serious complications that can occur (Williams 2008). Although this practice is not evidence-based, this shows the perception among hepato-pancreato biliary surgeons and gastroenterologists that it is important not to miss common bile duct stones.

Rationale

There are several benign (non-cancerous) and malignant (cancerous) conditions that may cause obstructive jaundice. Benign causes of obstructive jaundice include primary sclerosing cholangitis (Penz-Osterreicher 2011), primary biliary cirrhosis (Hirschfield 2011), chronic pancreatitis (Abdallah 2007), autoimmune pancreatitis (Lin 2008), inflammatory strictures of the common bile duct (Krishna 2008), and strictures of the common bile duct caused by prior instrumentation (Lillemoe 2000; Tang 2011). Malignant causes of obstructive jaundice include cholangiocarcinoma (Siddiqui 2011), cancer of the ampulla of Vater as well as other periampullary cancers (Hamade 2005; Choi 2011; Park 2011), and carcinoma of the pancreas (Singh 1990; Kalady 2004). It is important to differentiate between the causes of obstructive jaundice in order to initiate appropriate treatment. The correct diagnosis of common bile duct stones is an essential contribution to this differentiation.

Common bile duct stones are responsible for a range of complications. Common bile duct stones may lead to pancreatitis in about 33% to 50% of people (Corfield 1985; Toh 2000), and cause mortality in about 6% to 9% of these people (Mann 1994; Toh 2000). Acute cholangitis appears in 2% to 9% of people admitted for gallstone disease, with mortality around 24% (Salek 2009). Therefore, it is important to diagnose common bile duct stones in order to initiate treatment and prevent such complications.

The preferred option for the treatment of people with gallstones and common bile duct stones is currently endoscopic sphincterotomy with balloon trawling followed by laparoscopic cholecystectomy (Ludwig 2001; Spelsberg 2009). Other options include open cholecystectomy with open common bile duct exploration, laparoscopic cholecystectomy with laparoscopic common bile duct exploration, and laparoscopic cholecystectomy with endoscopic sphincterotomy (Hong 2006; Dasari 2013). Approximately half of people with jaundice, abnormal liver function

tests, and common bile duct dilation on ultrasound do not actually have common bile duct stones at ERCP (Hoyuela 1999), and these people have therefore undergone an unnecessary invasive procedure. Accurate diagnosis of common bile duct stones may avoid unnecessary procedures and complications associated with these procedures. Invasive tests can result in complications, for example, ERCP with endoscopic sphincterotomy can have life-threatening complications such as pancreatitis (Gurusamy 2011). Accurate diagnosis of common bile duct stones using non-invasive tests can avoid these complications.

Currently, there are no Cochrane reviews of studies assessing the accuracy of different tests for diagnosing common bile duct stones. This review is one of three reviews evaluating the diagnostic accuracy of different tests in the diagnosis of common bile duct stones and will help in the development of an evidence-based algorithm for diagnosis of common bile duct stones.

Ultrasound and liver function tests are non-invasive and almost used universally in the diagnosis of common bile duct stones, these tests are used for triage to determine whether people suspected of common bile duct stones undergo further tests. In this respect, it is important to know the false-negative rate as such people will be at risk of developing complications due to common bile duct stones. If the false-negative rate is high, then people may have to undergo other tests regardless of the results of ultrasound and liver function tests and so these tests fail in their role as triage tests. While false-positive results from liver function tests and ultrasound are not very significant in a clinical pathway where further tests such as MRCP or EUS are performed prior to ERCP followed immediately by endoscopic sphincterotomy or surgical exploration of common bile duct, the false-positive rate becomes very important in a clinical pathway where ERCP or surgical exploration is attempted directly after ultrasound or liver function tests as people are exposed to the risk of complications associated with these invasive procedures.

OBJECTIVES

To determine and compare the accuracy of ultrasound versus liver function tests for the diagnosis of common bile duct stones.

Secondary objectives

To investigate variation in the diagnostic accuracy of ultrasound and liver function tests according to the following potential sources of heterogeneity.

1. Studies at low risk of bias versus studies with unclear or high risk of bias (as assessed by the quality assessment of diagnostic accuracy studies assessment tool (QUADAS-2) tool (Table 1)).
2. Full-text publications versus abstracts (this may indicate publication bias if there is an association between the results of the study and the study reaching full publication) (Eloubeidi 2001).
3. Prospective versus retrospective studies.
4. Symptomatic versus asymptomatic common bile duct stones (the presence of symptoms may increase the pre-test probability). People with symptoms were defined as people showing upper right quadrant abdominal pain, jaundice, acute cholangitis, or acute pancreatitis (Anciaux 1986; Roston 1997; Raraty 1998; Toh 2000; Tranter 2003).

5. Prevalence of common bile duct stones in each included study. The prevalence of common bile duct stones in the population analysed by each included study may vary and cause heterogeneity. Prevalence may also change with people with comorbidities that would predispose them to common bile duct stones such as primary sclerosing cholangitis, Caroli's disease, hypercholesterolaemia, sickle cell anaemia, and sphincter of Oddi dysfunction.
6. Proportion of people with previous cholecystectomy. Cholecystectomy may cause dilation of the common bile duct ([Benjaminov 2013](#)), and subsequently change the accuracy of the index test particularly imaging modalities.

METHODS

Criteria for considering studies for this review

Types of studies

We included studies providing cross-sectional information comparing one or more of the index tests against a reference standard in the appropriate patient population (see [Participants](#)). We included studies irrespective of language or publication status, or whether data were collected prospectively or retrospectively. We included comparative studies in which ultrasound and liver function tests were performed in the same study population either by giving all participants the index tests or by randomly allocating participants to receive ultrasound or a liver function test. We excluded diagnostic case-control studies if there were at least four cross-sectional or comparative studies.

Participants

People at risk of or suspected of having common bile duct stones, with or without prior diagnosis of cholelithiasis; with or without symptoms and complications of common bile duct stones, with or without prior treatment for common bile duct stones; and before or after cholecystectomy.

Index tests

Ultrasound and liver function tests (serum bilirubin and ALP).

Target conditions

Common bile duct stones.

Reference standards

We accepted the following reference standards.

- For test positives, we accepted confirmation of a common bile duct stone by extraction of the stone (irrespective of whether this is done by surgical or endoscopic methods).
- For test negatives, we acknowledged that there was no way of being absolutely sure that there were no common bile duct stones. However, we accepted negative results by surgical or endoscopic negative exploration of the common bile duct, or symptom-free follow-up for at least six months, as the reference standard. Surgical or endoscopic exploration is adequate, but it is not commonly used in people with negative index tests because of its invasive nature. Therefore, we accepted follow-up as a less adequate reference test. Negative exploration of common bile duct is likely to be a better reference standard than follow-up for at least six months since most stones already

present in the common bile duct are likely to be extracted using this method. Six months was an arbitrary choice, but we anticipated most common bile duct stones would manifest during this period.

Search methods for identification of studies

Electronic searches

We searched MEDLINE via PubMed (January 1946 to September 2012), EMBASE via OvidSP (January 1947 to September 2012), Science Citation Index Expanded via Web of Knowledge (January 1898 to September 2012), BIOSIS via Web of Knowledge (January 1969 to September 2012), and clinicaltrials.gov/ (September 2012). [Appendix 1](#) shows the search strategies. We used a common search strategy for the three reviews of which this review is one. The other two reviews assessed the diagnostic accuracy of EUS, MRCP, ERCP, and IOC ([Gurusamy 2015](#); [Giljaca 2015](#)). We also identified systematic reviews from the Database of Abstracts of Reviews of Effects (DARE), Health Technology Assessment (HTA), Medion, and ARIF (Aggressive Research Intelligence Facility) databases in order to search their reference lists (see [Searching other resources](#)).

Searching other resources

We searched the references of included studies and systematic reviews related to the topic for identifying further studies. We also searched for additional articles related to included studies by performing the 'related search' function in MEDLINE (PubMed) and EMBASE (OvidSP), and a 'citing reference' search (search the articles that cited the included articles) ([Sampson 2008](#)) in Science Citation Index Expanded and EMBASE (OvidSP).

Data collection and analysis

Selection of studies

Two authors (VG and DH or GP) searched the references independently for identification of relevant studies. We obtained full texts for the references that at least one of the authors considered relevant. Two authors (VG and DH or GP) independently assessed the full-text articles. One author (KG) arbitrated any differences in study selection. We selected studies that met the inclusion criteria for data extraction. We included abstracts that provided sufficient data to create a 2 x 2 table.

Data extraction and management

Two authors (KG and VG) independently extracted the following data from each included study.

1. First author of report.
2. Year of publication of report.
3. Study design (prospective or retrospective; cross-sectional studies or randomised clinical trials).
4. Inclusion and exclusion criteria for individual studies.
5. Total number of participants.
6. Number of males and females.
7. Mean age of the participants.
8. Tests carried out prior to index test.
9. Index test.
10. Reference standard.
11. Number of true positives, false positives, true negatives, and false negatives.

We sought further information on the diagnostic test accuracy data and assessment of methodological quality (see [Assessment of methodological quality](#)) from the authors of the studies, if necessary. We resolved any differences between the review authors by discussion until we reached a consensus. We extracted the data excluding the indeterminates but recorded the number of indeterminates and the reference standard results of the participants with indeterminate results.

Assessment of methodological quality

We adopted the quality assessment of diagnostic accuracy studies assessment tool (QUADAS-2) for assessment of the methodological quality of included studies as described in [Table 1](#) ([Whiting 2006](#); [Whiting 2011](#)). We considered studies classified at low risk of bias and low concern regarding applicability to the review question as studies at low risk of bias. We resolved any differences in the methodological quality assessment by discussion between the authors until we reached consensus. We sought further information from study authors in order to assess the methodological quality of included studies accurately.

Statistical analysis and data synthesis

We plotted study estimates of sensitivity and specificity on forest plots and in receiver operating characteristic (ROC) space to explore between-study variation in the performance of the tests. Because our focus of inference was summary points, we used the bivariate model to summarise jointly the sensitivity and specificity if the number of studies was adequate for a meta-analysis to be performed ([Reitsma 2005](#); [Chu 2006](#)). This model accounts for between-study variability in estimates of sensitivity and specificity through the inclusion of random effects for the logit sensitivity and logit specificity parameters of the bivariate model. We planned to calculate the summary sensitivity and specificity of serum bilirubin and serum ALP for each reported cut-off if there were sufficient data to enable meta-analyses. We performed meta-analysis using the `xmlogit` command in Stata version 13 (Stata-Corp, College Station, Texas, USA). Confidence regions on summary ROC plots generated using Review Manager 5 are excessively conservative when there are few studies and they may appear inconsistent with the estimated CIs ([RevMan 2012](#)). While estimation of the CIs relies on the standard errors, the confidence regions rely on the number of studies in addition to the standard errors and the covariance of the estimated mean logit sensitivity and specificity. Therefore, if there were fewer than 10 studies included in a meta-analysis, we used 10 as the number of studies for generating the regions. This number is arbitrary but seemed to provide a better approximation than using a small number of studies.

We planned to compare the diagnostic accuracy of ultrasound, serum bilirubin, and serum ALP by including covariate terms for test type in the bivariate model to estimate differences in the sensitivity and specificity of the tests. In addition, if studies that evaluated ultrasound and liver function tests in the same study population

had been available, we planned to perform a direct head-to-head comparison by limiting the test comparison to such studies.

We created a table of pre-test probabilities (using the observed median and range of prevalence from the included studies) against post-test probabilities. The post-test probabilities were calculated using the pre-test probabilities and the summary positive and negative likelihood ratios. We computed the summary likelihood ratios and their CIs by using the Stata `_diparm` command and functions of the parameter estimates from the bivariate model that we fitted to estimate the summary sensitivity and specificity of a test.

Investigations of heterogeneity

We visually inspected forest plots of sensitivity and specificity and summary ROC plots to identify heterogeneity. We investigated the sources of heterogeneity stated in the [Secondary objectives](#). Where possible, given the number of included studies, we planned to explore heterogeneity formally by adding each potential source of heterogeneity as a covariate in the bivariate model (meta-regression with one covariate at a time).

Sensitivity analyses

Exclusion of participants with uninterpretable results can result in overestimation of diagnostic test accuracy ([Schuetz 2012](#)). In practice, uninterpretable test results will generally be considered test negatives. Therefore, we planned to perform sensitivity analyses by including uninterpretable test results as test negatives if sufficient data were available.

Assessment of reporting bias

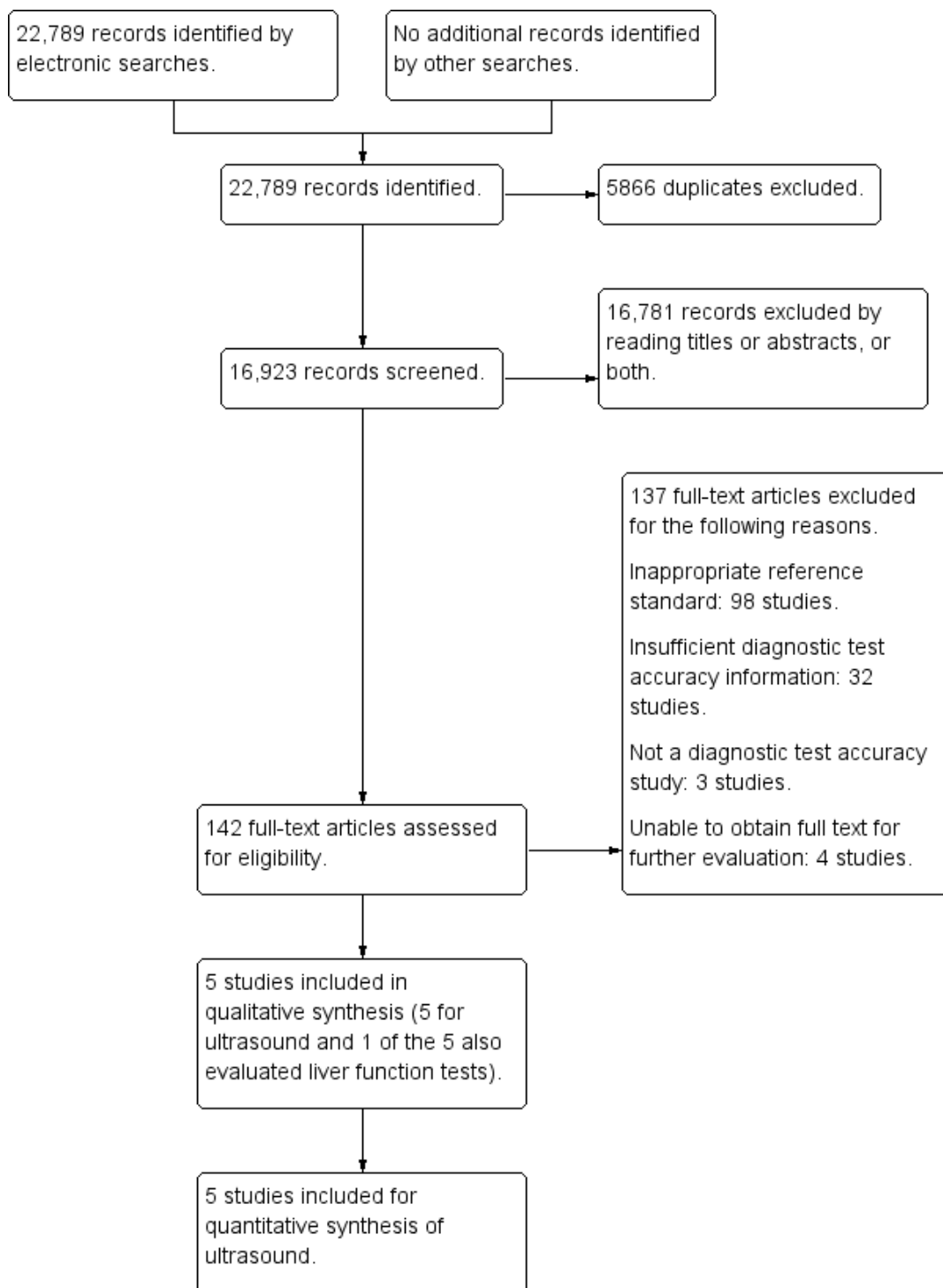
As described in the [Investigations of heterogeneity](#) section, we planned to investigate whether the summary sensitivities and specificities differed between studies that were published as full texts and studies that were available only as abstracts.

RESULTS

Results of the search

We identified 22,789 references through electronic searches of MEDLINE (8292 references), EMBASE (10,029 references), Science Citation Index Expanded and BIOSIS (4276 references), and DARE and HTA in *The Cochrane Library* (192 references). We identified no additional studies by searching the other sources. We excluded 5866 duplicates and 16,781 clearly irrelevant references through reading the abstracts. We retrieved 142 references for further assessment. We excluded 137 references for the reasons listed in the [Characteristics of excluded studies](#) table. Five studies fulfilled the inclusion criteria and provided data for the review. We were able to obtain additional information from the authors of one of the studies ([Kumar 1998](#)). [Figure 2](#) shows the flow of studies through the selection process.

Figure 2. Flow of studies through the screening process.



Characteristics of included studies

We included five studies (Busel 1989; Kumar 1998; Silverstein 1998; Admassie 2005; Rickes 2006; see [Characteristics of included studies](#) table). The studies included 523 participants. All five studies reported the diagnostic test accuracy of abdominal ultrasound. One study including 262 participants reported the diagnostic accuracy of serum bilirubin and serum ALP (Silverstein 1998). The

median pre-test probability of common bile duct stones in the five studies was 0.408 (range 0.095 to 0.658).

The five included studies were full-text publications. Three studies recruited participants prospectively (Kumar 1998; Silverstein 1998; Rickes 2006). Two studies included people with obstructive jaundice (Kumar 1998; Admassie 2005), two studies included people with jaundice or other symptoms such as pancreatitis (Silverstein 1998; Rickes 2006), and one study included people who

underwent common bile duct exploration for common bile duct stones (Busel 1989). Thus, it appears that all the studies included people with symptoms. One study included only participants who had not undergone previous cholecystectomy (Silverstein 1998), but the information was not available in the remaining studies. The proportion of people with common bile duct strictures was 6%

(Admassie 2005), and 12% (Kumar 1998), in the two studies that provided this information.

Methodological quality of included studies

Figure 3 and Figure 4 summarise the methodological quality of the included studies. All the studies were of poor methodological quality.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

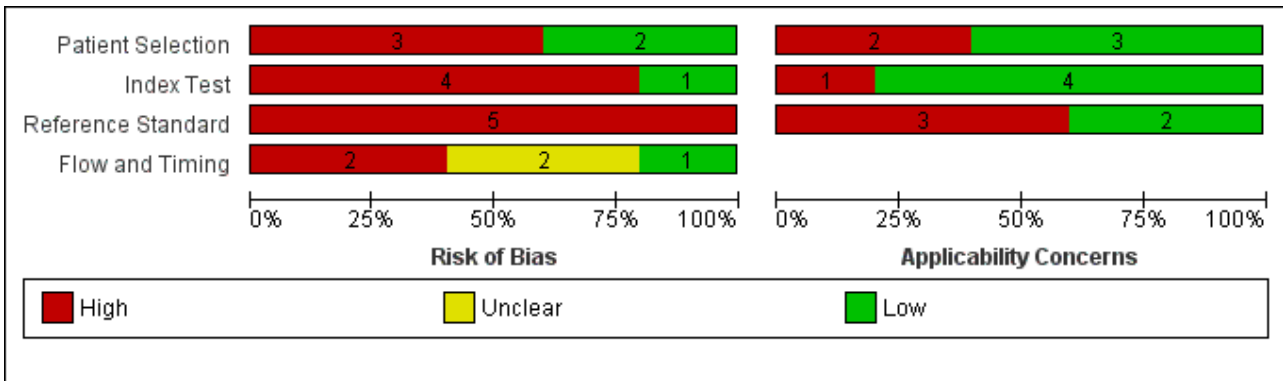





Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Admassie 2005	-	-	-	?	+	-	-
Busel 1989	-	-	-	?	-	+	-
Kumar 1998	+	+	-	+	+	+	+
Rickes 2006	-	-	-	-	-	+	+
Silverstein 1998	+	-	-	-	+	+	-

 High	 Unclear	 Low
---	--	---

Patient selection

Two studies were at low risk of bias and of low concern regarding applicability in the 'patient selection' domain (Kumar 1998; Silverstein 1998). The remaining studies were at high risk of bias with high concern about applicability because they did not mention whether a consecutive or random sample of participants was included.

Index test

Only one study was at low risk of bias in the 'index test' domain (Kumar 1998). The remaining studies were at high risk of bias because it was not clear whether the index test results were interpreted without knowledge of the reference standard results. Four studies were of low concern about applicability (Busel 1989; Kumar 1998; Silverstein 1998; Rickes 2006), while one study was of high concern because the criteria for a positive diagnosis were not reported (Admassie 2005).

Reference standard

None of the studies were at low risk of bias in the 'reference standard' domain. The studies were at high risk of bias because it was either not clear whether the reference standards were interpreted without knowledge of the index test results (Busel 1989; Silverstein 1998; Admassie 2005; Rickes 2006), or it was clear that

the reference standards were interpreted with knowledge of the index test results (Kumar 1998). Four studies were of low concern about applicability (Busel 1989; Kumar 1998; Admassie 2005; Rickes 2006), while one study was of high concern because endoscopic or surgical clearance of common bile duct was achieved in people with a positive test result and clinical follow-up was performed in people with a negative test result (Silverstein 1998).

Flow and timing

Only one study was at low risk of bias in the 'flow and timing' domain (Kumar 1998). The remaining studies were at high risk of bias because of the following reasons. Four studies did not report the time interval between the index test and reference standard (Busel 1989; Silverstein 1998; Admassie 2005; Rickes 2006), and in one study, the same reference standard was not used since endoscopic or surgical clearance of common bile duct was achieved in people with a positive test result and clinical follow-up was performed in people with a negative test result (Silverstein 1998). It was unclear whether all the participants were included in the analysis in two studies (Busel 1989; Admassie 2005), while two participants were excluded from the analysis in one study (Rickes 2006).

Findings

The forest plot of sensitivity and specificity (Figure 5) and the receiver operating characteristics (ROC) plot (Figure 6) summarises the individual study estimates of sensitivity and specificity (with

95% CIs) for ultrasound and the liver function tests. Results were available at two different cut-offs for bilirubin (greater than 22.23 µmol/L and greater than twice the normal limit) and ALP (greater than 125 IU/L and greater than twice the normal limit). Based on the five included studies, the minimum pre-test probability was 0.095, median was 0.408, and maximum was 0.658.

Figure 5. Forest plot of ultrasound and liver function tests for detection of common bile duct stones. The plot shows study-specific estimates of sensitivity and specificity (with 95% confidence intervals). The studies are ordered according to whether recruitment was prospective or not, and sensitivity. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Ultrasound

Study	TP	FP	FN	TN	Prospective	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Admassie 2005	18	6	2	23	No	0.90 [0.68, 0.99]	0.79 [0.60, 0.92]		
Busel 1989	17	3	8	10	No	0.68 [0.46, 0.85]	0.77 [0.46, 0.95]		
Kumar 1998	11	1	0	38	Yes	1.00 [0.72, 1.00]	0.97 [0.87, 1.00]		
Rickes 2006	47	4	34	39	Yes	0.58 [0.47, 0.69]	0.91 [0.78, 0.97]		
Silverstein 1998	8	13	17	224	Yes	0.32 [0.15, 0.54]	0.95 [0.91, 0.97]		

Bilirubin (> 22.23 µmol/L)

Study	TP	FP	FN	TN	Prospective	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Silverstein 1998	21	22	4	215	Yes	0.84 [0.64, 0.95]	0.91 [0.86, 0.94]		

Bilirubin (> twice the normal limit)

Study	TP	FP	FN	TN	Prospective	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Silverstein 1998	10	6	14	231	Yes	0.42 [0.22, 0.63]	0.97 [0.95, 0.99]		

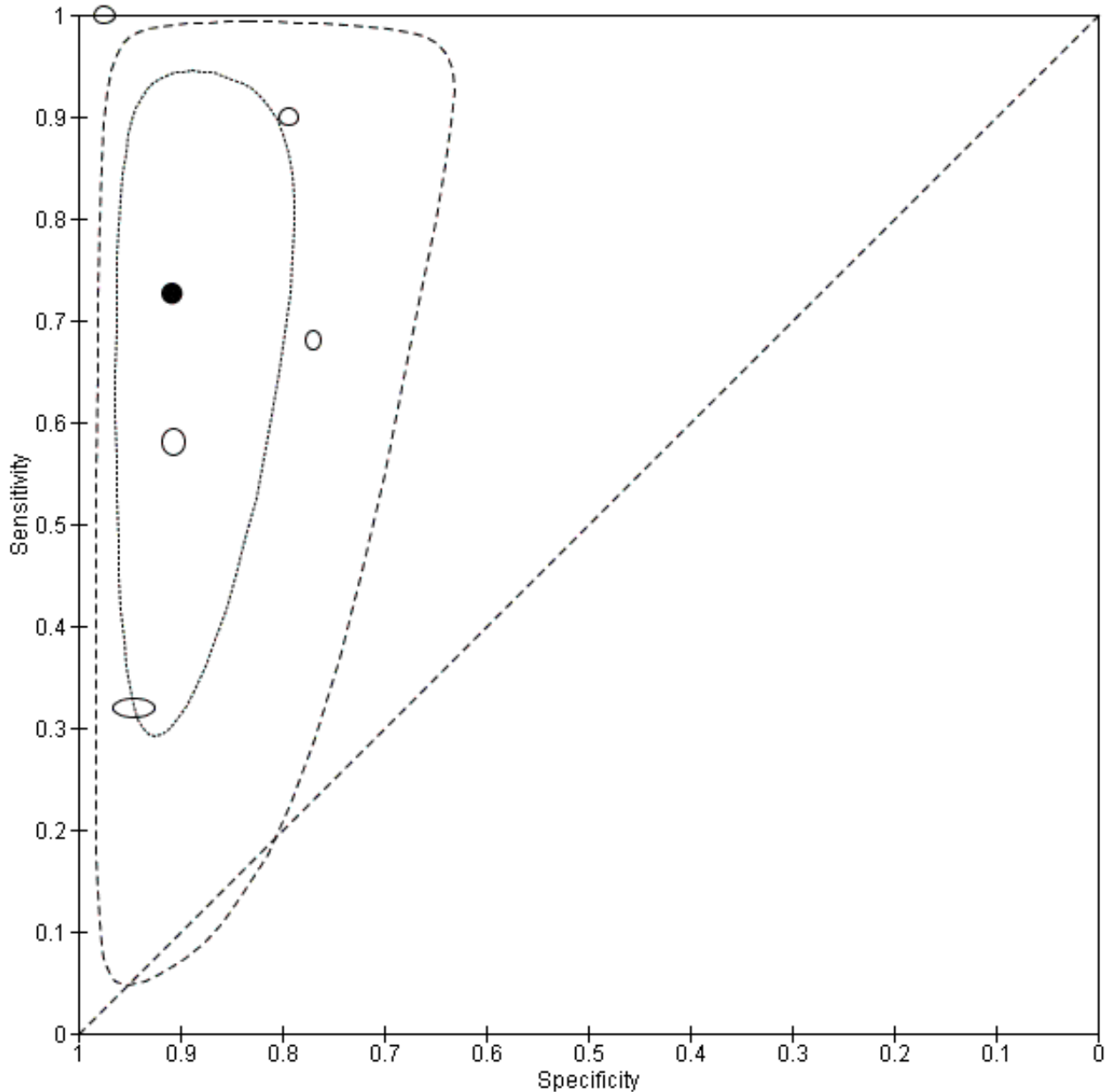
Alkaline phosphatase (> 125 IU/L)

Study	TP	FP	FN	TN	Prospective	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Silverstein 1998	23	49	2	188	Yes	0.92 [0.74, 0.99]	0.79 [0.74, 0.84]		

Alkaline phosphatase (> twice the normal limit)

Study	TP	FP	FN	TN	Prospective	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Silverstein 1998	9	6	15	231	Yes	0.38 [0.19, 0.59]	0.97 [0.95, 0.99]		

Figure 6. Summary receiver operating characteristic (ROC) plot of ultrasound for detection of common bile duct stones. Each ellipse on the plot represents the pair of sensitivity and specificity from a study and the size of the ellipse is scaled according to the sample size of the study. The solid black circle represents the summary sensitivity and specificity, and this summary point is surrounded by a 95% confidence region (dotted line) and 95% prediction region (dashed line).



Ultrasound

Five studies including 162 cases out 523 participants reported the diagnostic test accuracy of ultrasound. Sensitivities ranged from 0.32 to 1.00 and specificities ranged from 0.77 to 0.97. The summary sensitivity was 0.73 (95% CI 0.44 to 0.90) and specificity was 0.91 (95% CI 0.84 to 0.95). The positive likelihood ratio was 7.89 (95% CI 4.33 to 14.4) and the negative likelihood ratio was 0.30 (95% CI 0.13 to 0.72). At the median pre-test probability of common bile duct stones of 0.408, post-test probability associated with positive test

results was 0.85 (95% CI 0.75 to 0.91) and negative test results was 0.17 (95% CI 0.08 to 0.33). At the minimum pre-test probability of 0.095, the post-test probability associated with positive test results was 0.45 (95% CI 0.31 to 0.60) and negative test results was 0.03 (95% CI 0.01 to 0.07). At the maximum pre-test probability of 0.658, the post-test probability associated with positive test results was 0.94 (95% CI 0.89 to 0.97) and negative test results was 0.37 (95% CI 0.20 to 0.58).

Liver function tests

Although we planned to estimate the summary sensitivity and specificity at each cut-off point for serum bilirubin and serum ALP, only one study evaluated the tests and so we did not perform meta-analyses. The study included 262 participants of whom 25 had common bile duct stones (Silverstein 1998). The diagnostic accuracy of serum bilirubin and serum ALP were reported at two cut-offs for each test. For bilirubin at a cut-off of greater than 22.23 $\mu\text{mol/L}$, sensitivity was 0.84 (95% CI 0.65 to 0.94) and specificity was 0.91 (0.86 to 0.94). For bilirubin at a cut-off of greater than twice the normal limit, sensitivity was 0.42 (95% CI 0.22 to 0.63) and specificity was 0.97 (95% CI 0.95 to 0.99).

For ALP at a cut-off of greater than 125 IU/L, sensitivity was 0.92 (95% CI 0.74 to 0.99) and specificity was 0.79 (95% CI 0.74 to 0.84). For ALP at a cut-off of greater than twice the normal limit, sensitivity was 0.38 (95% CI 0.19 to 0.59) and specificity was 0.97 (95% CI 0.95 to 0.99).

Test strategies involving combinations of bilirubin and ALP were not reported.

Combinations of ultrasound and liver function tests

None of the studies assessed the diagnostic accuracy of a combination of ultrasound and liver function tests.

Comparison of ultrasound and liver function tests

We were unable to compare tests formally because only one study evaluated liver function tests (Silverstein 1998). This study also assessed ultrasound and reported higher sensitivities for bilirubin and ALP than ultrasound at both cut-offs but the specificities of the markers were higher at only the greater than twice the normal limit cut-off (Figure 5). For ultrasound, sensitivity was 0.32 (95% CI 0.15 to 0.54) and specificity was 0.95 (95% CI 0.91 to 0.97). For bilirubin at a cut-off greater than 22.23 $\mu\text{mol/L}$, sensitivity was 0.84 (95% CI 0.64 to 0.95) and specificity was 0.91 (0.86 to 0.94). For ALP at a cut-off greater than 125 IU/L, sensitivity was 0.92 (95% CI 0.74 to 0.99) and specificity was 0.79 (95% CI 0.74 to 0.84).

Exploration of heterogeneity

We performed none of the planned investigations of heterogeneity because few studies were included in the review.

Sensitivity analyses

We performed no sensitivity analyses. One study reported the exclusion of two participants, but their reference standard results were not available (Rickes 2006).

DISCUSSION

Summary of main results

Summary of findings 1 shows the results for ultrasound and liver function tests. There was considerable variation in the estimates of sensitivity and specificity between the five studies that evaluated ultrasound. For ultrasound, the summary sensitivity was 0.73 (95% CI 0.44 to 0.90) and specificity was 0.91 (95% CI 0.84 to 0.95). At the median pre-test probability of 40.8%, the post-test probability for ultrasound was 0.85 (95% CI 0.75 to 0.91) for a positive test result and 0.17 (95% CI 0.08 to 0.33) for a negative test result. Bilirubin and ALP were the two liver function tests assessed in this

review. Only one study evaluated both markers and the study also evaluated ultrasound. The study reported two cut-offs for defining test positivity for bilirubin and ALP. For bilirubin, at cut-offs of greater than 22.23 $\mu\text{mol/L}$, the sensitivity was 0.84 (95% CI 0.65 to 0.94) and specificity was 0.91 (95% CI 0.86 to 0.94) and at cut-offs of greater than twice the normal limit, sensitivity was 0.42 (95% CI 0.22 to 0.63) and specificity was 0.97 (95% CI 0.95 to 0.99). For ALP, at cut-offs of greater than 125 IU/L, sensitivity was 0.92 (95% CI 0.74 to 0.99) and specificity was 0.79 (95% CI 0.74 to 0.84) and at cut-offs of greater than twice the normal limit, sensitivity was 0.38 (95% CI 0.19 to 0.59) and specificity was 0.97 (95% CI 0.95 to 0.99). None of the studies assessed the diagnostic accuracy of a combination of ultrasound and liver function tests.

An ideal test should have few false-positive test results (so that people without common bile duct stones are not diagnosed to have common bile duct stones by the test performed) and few false-negative test results (so that people with common bile duct stones are not missed by the test). Since relatively non-invasive tests such as MRCP and EUS follow abnormal transabdominal ultrasound or liver function tests before people undergo invasive tests or treatment, it is acceptable for ultrasound and liver function tests to have false-positive test results. However, it is not acceptable for ultrasound and liver function tests to have false-negative results. This is because people with negative ultrasound and liver function tests may not be investigated further and so common bile duct stones may be missed exposing the person to the risk of developing pancreatitis and cholangitis, both of which are life-threatening conditions. Although some researchers have suggested that there is no need to treat asymptomatic common bile duct stones (Caddy 2005), it is generally recommended that common bile duct stones are removed when they are identified because of the serious complications associated with their presence (Williams 2008). At the minimum pre-test probability of 9.5%, which is close to the generally accepted 10% to 20% probability of common bile duct stones in people undergoing cholecystectomy (Williams 2008), ALP at a cut-off of greater than 125 IU/L appeared to have the lowest post-test probability for a negative test result at 1.0%. This means that approximately 1 out of 100 people with common bile duct stones would be missed by the test but this depends on the pre-test probability of common bile duct stones. While people with abnormal liver function tests or ultrasound undergo further investigations such as MRCP to determine whether they have stones, it is unlikely that people with normal liver function tests and ultrasound would be subjected to further investigations. Liver function tests identify only whether there is obstruction in the common bile duct while ultrasound has the potential to identify the cause of obstruction by visualisation of a common bile duct stone. Therefore, strategies such as 'either test positive' may improve the sensitivity of the test as some people with normal ultrasound may have an abnormal liver function test and people with normal liver function tests may have an abnormal transabdominal ultrasound. In contrast, a strategy of 'both tests positive' may decrease the sensitivity but improve the specificity. However, these strategies were not evaluated in any of the studies.

Strengths and weaknesses of the review

We searched the literature thoroughly including full-text publications and abstracts without any language restrictions. The use of diagnostic test accuracy filters may lead to the loss of some studies (Doust 2005), and so we did not use filters. Two authors

independently identified and extracted data from the studies potentially decreasing errors related to single data extraction (Buscemi 2006). We used appropriate reference standards; relying on ERCP or IOC as reference standards can result in incorrect information about diagnostic test accuracy (Gurusamy 2015), and should be avoided.

The major limitation in the review process was our inability to explore heterogeneity formally because few studies were included in the review. For ultrasound, visual inspection of the forest plot of sensitivity and specificity and the prediction ellipse on the summary ROC plot indicated considerable differences, especially in sensitivity, between studies and it was not possible to determine the reason for the differences. All studies included people with symptoms (i.e., people with obstructive jaundice or pancreatitis). Four studies used surgical or endoscopic clearance as the reference standard in all participants (Busel 1989; Kumar 1998; Admassie 2005; Rickes 2006). In one study, participants with positive index test underwent surgical or endoscopic extraction of stones while the remaining participants had a clinical follow-up of a minimum of two years (Silverstein 1998). It is quite possible that the participants included in the first four studies were at greater risk of having common bile duct stones than the last study. This was evident from pre-test probabilities of common bile duct stones, which ranged from 22% to 66% in studies in which surgical or endoscopic clearance was performed in all participants (Busel 1989; Kumar 1998; Admassie 2005; Rickes 2006) compared to 10% in the study in which different reference standards were used depending upon the results of the index test (Silverstein 1998). Another potential source of heterogeneity was the criterion used to define a positive ultrasound. Two studies used hyperechoic shadowing within the common bile duct as the criterion for a positive ultrasound (Busel 1989; Rickes 2006); one study used dilated common bile duct (Silverstein 1998); one study used a combination of hyperechoic shadowing and dilated common bile duct (Kumar 1998); and the criterion was not stated in one study (Admassie 2005). It was also not possible to perform a comparison of the accuracy of the tests because we found too few studies for inclusion in the review.

The major limitation of the included studies was that none of the studies was of good methodological quality. There was a high proportion of studies at high risk of bias and of high concern regarding applicability in all the four domains. This makes the results potentially unreliable. We considered endoscopic or surgical extraction of common bile duct stones in all participants as a better reference standard than a combination of extraction of common bile duct stones in participants with positive index test and clinical follow-up in participants with negative index test. However, we acknowledge that even this ideal reference standard can result in misclassification and hence alteration in diagnostic test accuracy if one or more stones reach the small bowel without the knowledge of the person who performed the common bile duct stone extraction.

Despite all these shortcomings, these studies are the best available evidence on the topic.

Applicability of findings to the review question

Most of the participants included in this review were people who had signs or symptoms related to common bile duct stones. The diagnostic accuracy of the tests when applied in people without symptoms with common bile duct stones may be lower. Therefore,

unless further information becomes available, relying on these tests to make a diagnosis of common duct stones in people without symptoms cannot be recommended. It should also be noted that this review assessed the diagnostic accuracy of these tests only for the diagnosis of common bile duct stones and not for the diagnosis of other conditions such as benign or malignant biliary stricture and periampullary tumours.

Previous research

This is the first systematic review on this topic using appropriate reference standards.

AUTHORS' CONCLUSIONS

Implications for practice

Many people may have common bile duct stones in spite of having a negative ultrasound or liver function test. Such people may have to be tested with other modalities if the clinical suspicion of common bile duct stones is very high because of their symptoms. It is recommended that further non-invasive tests be used to confirm common bile duct stones to avoid the risks of invasive testing.

It should be noted that our results are based on studies of poor methodological quality and the results should be interpreted with caution.

Implications for research

Further studies of high methodological quality are necessary to determine the diagnostic test accuracy of ultrasound and liver function tests. In particular, it would be useful to compare the diagnostic test accuracy of either test positive versus both tests positive approach to inform the role of these triage tests and whether it is safe not to perform other tests if liver function tests and ultrasound do not suggest common bile duct stones. We acknowledge that differential verification cannot always be avoided if endoscopic sphincterotomy and extraction of stones was used as the reference standard because of the complications associated with this procedure (Gurusamy 2011). Surgical exploration of common bile ducts is a major surgical procedure and cannot be undertaken lightly. People with positive test results are likely to undergo endoscopic sphincterotomy and extraction of stones or surgical exploration of common bile ducts while people with negative test results are likely to be followed up. It is recommended that such people be followed up for at least six months to ensure that they do not develop the symptoms of common bile duct stones. Future studies that avoid inappropriate exclusions to ensure that true diagnostic accuracy can be calculated would be informative. Long-term follow-up of people with negative test results would help in understanding the implications of false-negative results and aid clinical decision making.

ACKNOWLEDGEMENTS

We thank the Cochrane Hepato-Biliary Group staff, contact editors Mirella Fraquelli and Agostino Colli, and the UK Cochrane Diagnostic Test Accuracy Review Support Unit for their advice in the preparation of this review.

This project was funded by the National Institute for Health Research.

Disclaimer of the Department of Health: "The views and opinions expressed in the review are those of the authors and do not necessarily reflect those of the National Institute for

Health Research (NIHR), National Health Services (NHS), or the Department of Health".

REFERENCES

References to studies included in this review

Admassie 2005 {published data only}

Admassie D, Hyesus A, Denke A. Validity of ultrasonography in diagnosing obstructive jaundice. *East African Medical Journal* 2005;**82**(7):379-81.

Busel 1989 {published data only}

Busel D, Espinoza Ugarte A, Osorio M, Baeza A, Silva G, Niedmann J, et al. Ultrasonics in the diagnosis of choledocholithiasis. *Revista Medica de Chile* 1989;**117**(1):40-1.

Kumar 1998 {published data only}

Kumar M, Prashad R, Kumar A, Sharma R, Acharya SK, Chattopadhyay TK. Relative merits of ultrasonography, computed tomography and cholangiography in patients of surgical obstructive jaundice. *Hepato-Gastroenterology* 1998;**45**(24):2027-32.

Rickes 2006 {published data only}

Rickes S, Treiber G, Monkemuller K, Peitz U, Csepregi A, Kahl S, et al. Impact of the operator's experience on value of high-resolution transabdominal ultrasound in the diagnosis of choledocholithiasis: a prospective comparison using endoscopic retrograde cholangiography as the gold standard. *Scandinavian Journal of Gastroenterology* 2006;**41**(7):838-43.

Silverstein 1998 {published data only}

Silverstein JC, Wavak E, Millikan KW. A prospective experience with selective cholangiography. *American Surgeon* 1998;**64**(7):654-9.

References to studies excluded from this review

Abdul Ghani 1989 {published data only}

Abdul Ghani AK. Selective per-operative cholangiography and scoring method for selection. *Bangladesh Medical Research Council Bulletin* 1989;**15**(2):81-9.

Almersjo 1966 {published data only}

Almersjo O, Hultborn KA, Jensen C. Diagnosis of choledocholithiasis. Relations between historical, laboratory, operative, and peroperative roentgen data. *Acta Chirurgica Scandinavica* 1966;**131**(1):112-26.

Al Samman 2000 {published data only}

Al Samman M, Nelson BK, Kwong KL, Mohandas A. The accuracy of liver biochemistry and abdominal ultrasound in detecting choledocholithiasis. *Gastroenterology* 2000;**118**(4 Suppl 2):A448.

Ang 2007 {published data only}

Ang TL, Ming FK, Kiong TE. A prospective study of the impact of clinical parameters and EUS in the management of suspected choledocholithiasis. *Journal of Gastroenterology and Hepatology* 2007;**22**(Suppl 2):A220.

Arroyo 1989 {published data only}

Arroyo A, Guerra M, Strauszer T. Value of ultrasonography in the diagnosis of distal choledocolithiasis. *Revista Medica de Chile* 1989;**117**(11):1251-4.

Bardach 1984 {published data only}

Bardach G, Iskandrani G, Grabenwoger F. Reliability of sonography and i.v. cholangiography in preoperative investigation of the common bile duct. *Rontgenpraxis* 1984;**37**(10):381-3.

Barkun 1994 {published data only}

Barkun AN, Barkun JS, Fried GM, Ghitulescu G, Steinmetz O, Pham C, et al. Useful predictors of bile duct stones in patients undergoing laparoscopic cholecystectomy. *Annals of Surgery* 1994;**220**(1):32-9.

Barr 1999 {published data only}

Barr LL, Frame BC, Coulanjon A. Proposed criteria for preoperative endoscopic retrograde cholangiography in candidates for laparoscopic cholecystectomy. *Surgical Endoscopy* 1999;**13**(8):778-81.

Bhatt 2005 {published data only}

Bhatt C, Shah PS, Prajapati HJ, Modi J. Comparison of diagnostic accuracy between USG and MRCP in biliary and pancreatic pathology. *Indian Journal of Radiology and Imaging* 2005;**15**(2):177-81.

Birkigt 1989 {published data only}

Birkigt HG, Lossner C, Buhtz C, Winckelmann B. Preoperative ultrasonic of common bile duct stones and its importance to surgical management - result of a prospective controlled trial. *Acta Chirurgica Hungarica* 1989;**30**(3):171-7.

Blackbourne 1994 {published data only}

Blackbourne LH, Earnhardt RC, Siström CL, Abbitt P, Jones RS. The sensitivity and role of ultrasound in the evaluation of biliary obstruction. *American Surgeon* 1994;**60**(9):683-90.

Boboev 2012 {published data only}

Boboev B. Ultrasonography in the diagnosis of cholelithiasis and its complications. *Vestnik Khirurgii Imeni I. I. Grekova* 2012;**171**(2):21-4.

Bocchetti 1989 {published data only}

Bocchetti T, Caramanico L, Benedettini MM, Izzo L, Costi U. Ultrasonography and biliary calculi: a retrospective study on the diagnostic reliability. *Giornale di Chirurgia* 1989;**10**(5):269-71.

Bokobza 1988 {published data only}

Bokobza B, Leblanc I, Michot F, Bouvier P, Teniere P. Peroperative investigations during surgery for biliary lithiasis. Endoscopy and ultrasonography. *Revue Francaise de Gastro-Enterologie* 1988;**24**(236):729-31.

Borsch 1985 {published data only}

Borsch G, Wedmann B, Brand J, Zumtobel V. Evaluation of pre- and postoperative sonography in biliary tract surgery. *Deutsche Medizinische Wochenschrift* 1985;**110**(36):1359-64.

Bose 2001 {published data only}

Bose SM, Mazumdar A, Prakash S, Kocher R, Katariya S, Pathak CM. Evaluation of the predictors of choledocholithiasis: comparative analysis of clinical, biochemical, radiological, radionuclear, and intraoperative parameters. *Surgery Today* 2001;**31**(2):117-22.

Calero Ayala 1983 {published data only}

Calero Ayala B, Mino Fugarolas G, De Dios Vega JF, Jimenez Sanchez JR. Ultrasound and ERCP in the diagnosis of common bile duct stones. *Gastroenterologia y Hepatologia* 1983;**6**(7):341-4.

Cao 1986 {published data only}

Cao HG. Ultrasonography in the diagnosis of common bile duct stones (an analysis of 112 cases). *Chinese Journal of Radiology* 1986;**20**(5):279-82.

Carrascosa 1983 {published data only}

Carrascosa J, Martinez A, Lloves A, Suarez A, Raffeta J, Magnanini F. Critical evaluation of echography in the diagnosis of cholestatic jaundice. *Acta Gastroenterologica Latinoamericana* 1983;**13**(4):689-98.

Chan 2008 {published data only}

Chan T, Yaghoubian A, Rosing D, Lee E, Lewis RJ, Stabile BE, et al. Total bilirubin is a useful predictor of persisting common bile duct stone in gallstone pancreatitis. *American Surgeon* 2008;**74**(10):977-80.

Chang 1998 {published data only}

Chang L, Lo SK, Stabile BE, Lewis RJ, De Virgilio C. Gallstone pancreatitis: a prospective study on the incidence of cholangitis and clinical predictors of retained common bile duct stones. *American Journal of Gastroenterology* 1998;**93**(4):527-31.

Chavez-Valencia 2009 {published data only}

Chavez-Valencia V, Espinosa-Ortega H, Espinoza-Peralta D, Arce-Salinas C. Diagnostic performance of biliary ultrasound vs. magnetic resonance cholangiogram in patients with recurrent biliary obstruction. *Revista de Gastroenterologia de Mexico* 2009;**74**(4):329-33.

Cockbain 2010 {published data only}

Cockbain A, Toogood G. Predicting common bile duct stones prior to laparoscopic cholecystectomy: a 10 year experience. *HPB: the Official Journal of the International Hepato Pancreato Biliary Association* 2010;**12**(Suppl S1):442-3.

Contini 1995 {published data only}

Contini S, Dallavalle R, Campanella G. Ultrasound (US) as a diagnostic-tool for common bile-duct (CBD) stones in preoperative screening - is it worthwhile. *British Journal of Surgery* 1995;**82**(Suppl 1):10.

Contini 1997 {published data only}

Contini S, Dallavalle R, Campanella G. Reliability of ultrasounds in detecting common bile duct stones before biliary surgery - a retrospective analysis. *Panminerva Medica* 1997;**39**(2):106-10.

Contractor 1997 {published data only}

Contractor QQ, Boujemla M, Contractor TQ, El Essawy OM. Abnormal common bile duct sonography - the best predictor of choledocholithiasis before laparoscopic cholecystectomy. *Journal of Clinical Gastroenterology* 1997;**25**(2):429-32.

Cranley 1980 {published data only}

Cranley B, Logan H. Exploration of the common bile duct: the relevance of the clinical picture and the importance of preoperative cholangiography. *British Journal of Surgery* 1980;**67**(12):869-72.

Cwik 2003 {published data only}

Cwik G, Wallner G, Ciechanski A, Zinkiewicz K, Zgodzinski W, Polkowski W. Endoscopic sphincterotomy in 100 patients scheduled for laparoscopic cholecystectomy: ultrasound evaluation. *Hepato-Gastroenterology* 2003;**50**(53):1225-8.

Danaci 2002 {published data only}

Danaci M, Polat V, Kamali A, Belet U, Incesu L, Ozen N, et al. Magnetic resonance cholangiography and ultrasonography in the diagnosis of choledocholithiasis. *Ondokuz Mayis Universitesi Tip Dergisi* 2002;**19**(3):195-201.

de Dios Vega 1982 {published data only}

de Dios Vega JF, Mino Fugarolas G, Rodriguez Vives A, Guerrero Gimenez P, Gomez Camacho F, Robles Olid J. Ultrasonics and endoscopic retrograde cholangiopancreatography. Correlation in biliopancreatic pathology. *Revista Espanola de las Enfermedades del Aparato Digestivo* 1982;**62**(1):45-9.

Delacruzcaro 1963 {published data only}

Delacruzcaro F. Value of the liver function tests in surgery of the biliary tract. *Medicina* 1963;**31**:505-6.

Derodra 1986 {published data only}

Derodra J. Predictive ability of choledocholithiasis indicators. *Annals of Surgery* 1986;**203**(3):335.

Dwivedi 1989 {published data only}

Dwivedi M, Acharya SK, Nundy S, Tandon BN. Accuracy of abdominal ultrasonography and the role of a second investigation in surgical obstructive jaundice. *Gastroenterologia Japonica* 1989;**24**(5):573-9.

Ellul 1991 {published data only}

Ellul JPM, Wilkinson ML, McColl I, Dowling RH. Do clinical and biochemical features of cholestasis predict choledocholithiasis in patients with gall-bladder stones - a prospective ERCP study including the effects of sphincterotomy on natural-history. *Gut* 1991;**32**(10):A1233.

Espinoza 1984 {published data only}

Espinoza P, Kunstlinger F, Liguory C. Accuracy of ultrasonography for the diagnosis of choledocholithiasis. *Gastroenterologie Clinique et Biologique* 1984;**8**(1):42-6.

Famos 1990 {published data only}

Famos M, Stadler P, Schneekloth G. Selective intraoperative cholangiography. *Helvetica Chirurgica Acta* 1990;**56**(6):897-901.

Fujino 1974 {published data only}

Fujino T, Takagi T, Nakajima S. Serum glutamyltranspeptidase activity in biliary diseases. *Japanese Journal of National Medical Services* 1974;**28**(5):361-366, 347.

Garcia-Caballero 1994 {published data only}

Garcia-Caballero M, Martin-Palanca A, Vara-Thorbeck C. Common bile duct stones after laparoscopic cholecystectomy and its treatment. The role of ultrasound and intravenous and intraoperative cholangiography. *Surgical Endoscopy* 1994;**8**(10):1182-5.

Gierblinski 2003 {published data only}

Gierblinski I, Boroszko D, Dziuba E, Basaj J. The value of ultrasonography in diagnosis of biliary tract disorders. *Gastroenterologia Polska* 2003;**10**(3):227-31.

Gologan 2006 {published data only}

Gologan E, Balan G, Stanciu C. Conventional ultrasound in the diagnosis of bile duct calculi. *Revista Medico* 2006;**110**(4):824-9.

Gregg 1979 {published data only}

Gregg JA, McDonald DG. Endoscopic retrograde cholangiopancreatography and grey-scale abdominal ultrasound in the diagnosis of jaundice. *American Journal of Surgery* 1979;**137**(5):611-5.

Gross 1983 {published data only}

Gross BH, Harter LP, Gore RM. Ultrasonic evaluation of common bile duct stones: prospective comparison with endoscopic retrograde cholangiopancreatography. *Radiology* 1983;**146**(2):471-4.

Gui 1995 {published data only}

Gui GP, Borgstein RL, Beaugie JM. Assessment of the common bile duct before cholecystectomy: a direct comparison between ultrasonography and intravenous cholangiography. *Journal of the Royal College of Surgeons of Edinburgh* 1995;**40**(1):16-9.

Hammarstrom 1998 {published data only}

Hammarstrom LE, Ranstam J. Factors predictive of bile duct stones in patients with acute calculous cholecystitis. *Digestive Surgery* 1998;**15**(4):323-7.

Holzinger 1999 {published data only}

Holzinger F, Baer HU, Wildi S, Vock P, Buchler MW. The role of intravenous cholangiography in the era of laparoscopic cholecystectomy: is there a renaissance?. *Deutsche Medizinische Wochenschrift* 1999;**124**(46):1373-8.

Hunt 1990 {published data only}

Hunt DR, Reiter L, Scott AJ. Pre-operative ultrasound measurement of bile duct diameter: basis for selective cholangiography. *Australian and New Zealand Journal of Surgery* 1990;**60**(3):189-92.

Jarvinen 1978 {published data only}

Jarvinen H. Abnormal liver function tests in acute cholecystitis; the predicting of common duct stones. *Annals of Clinical Research* 1978;**10**(6):323-7.

Jolobe 2012 {published data only}

Jolobe OMP. Limitations of gammaglutamyl transaminase as an indicator of biliary obstruction. *European Journal of Internal Medicine* 2012;**23**(3):e75.

Jovanovic 2012 {published data only}

Jovanovic P, Salkic N, Zerem E, Ljuca F. Gammaglutamyl transaminase in biliary obstruction. *European Journal of Internal Medicine* 2012;**23**(3):e76.

Katz 2004 {published data only}

Katz D, Nikfarjam M, Sfakiotaki A, Christophi C. Selective endoscopic cholangiography for the detection of common bile duct stones in patients with cholelithiasis. *Endoscopy* 2004;**36**(12):1045-9.

Kelly 2011 {published data only}

Kelly NM, Tham T. Abnormal liver function tests in acute cholecystitis with and without choledocholithiasis. *Gastroenterology* 2011;**140**(5 Suppl 1):S450.

Khan 2011 {published data only}

Khan RM, Sapkota B, Batool A, Khan MO, Sharma P, Gonzalez MO, et al. Obstructive jaundice & spectrum of liver function tests (LFT). *Gastrointestinal Endoscopy* 2011;**73**(4 Suppl 1):AB369-70.

Kruis 1997 {published data only}

Kruis W, Roehrig H, Hardt M, Pohl C, Schlosser D. A prospective evaluation of the diagnostic work-up before laparoscopic cholecystectomy. *Endoscopy* 1997;**29**(7):602-8.

Lacaine 1980 {published data only}

Lacaine F, Corlette MB, Bismuth H. Preoperative evaluation of the risk of common bile duct stones. *Archives of Surgery* 1980;**115**(9):1114-6.

Lakatos 2004 {published data only}

Lakatos L, Mester G, Reti G, Nagy A, Lakatos PL. Selection criteria for preoperative endoscopic retrograde cholangiopancreatography before laparoscopic cholecystectomy and endoscopic treatment of bile duct stones: results of a retrospective, single center study between 1996-2002. *World Journal of Gastroenterology* 2004;**10**(23):3495-9.

Lakoma 1996 {published data only}

Lakoma S, Malczak J, Dziekan R, Kryskow A. Choledocholithiasis - the evaluation of diagnostic effectiveness of USG and ERCP in own material. *Acta Endoscopica Polona* 1996;**6**(4):165-7.

Leander 1994 {published data only}

Leander P, Ekberg O, Almqvist P. Radiology in laparoscopic cholecystectomy - a retrospective study. *Acta Radiologica* 1994;**35**(5):437-41.

Lee 2008 {published data only}

Lee ES, Chan T, Rasing D, Lewis R, Yaghoobian A, de Virgilio C. Gallstone pancreatitis: predictive value of serum total bilirubin for common bile duct stones. *Journal of Investigative Medicine* 2008;**56**(1):211-2.

Leitman 1993 {published data only}

Leitman IM, Fisher ML, McKinley MJ, Rothman R, Ward RJ, Reiner DS, et al. The evaluation and management of known or suspected stones of the common bile duct in the era of minimal access surgery. *Surgery Gynecology and Obstetrics* 1993;**176**(6):527-33.

Lichtenbaum 2000 {published data only}

Lichtenbaum RA, McMullen HF, Newman RM. Preoperative abdominal ultrasound may be misleading in risk stratification for presence of common bile duct abnormalities. *Surgical Endoscopy* 2000;**14**(3):254-7.

Lindsell 1990 {published data only}

Lindsell DR. Ultrasound imaging of pancreas and biliary tract. *Lancet* 1990;**335**(8686):390-3.

Link 1995 {published data only}

Link J, Marienhoff N, Benecke P, Zwaan M, Wenk H. The value of sonography prior to laparoscopic cholecystectomy. *Fortschritte Auf Dem Gebiete Der Rontgenstrahlen Und Der Neuen Bildgebenden Verfahren* 1995;**162**(1):20-2.

Liu 1984 {published data only}

Liu GL. The value of ultrasonography in the diagnosis of choledocholithiasis. *Chinese Journal of Surgery* 1984;**22**(11):669-70, 701.

Lomanto 1999 {published data only}

Lomanto D, Carlei F, Paganini AM, Nardovino M, De Simone P, De Angelis L, et al. Validation of predictive factors for choledocholithiasis. *Gastroenterology* 1999;**116**(4):A1330-1.

Lui 1986 {published data only}

Lui P, Ng HS, Teh LB, Kwok KC, Ong YY, Tan L, et al. Ultrasonography in the diagnosis of cholestatic jaundice. *Annals of the Academy of Medicine* 1986;**15**(2):182-5.

Magnuson 1997 {published data only}

Magnuson TH, Regan F, Ahrendt SA, Schaefer DC, Bender JS. Magnetic resonance cholangiography compared to ultrasound and other clinical criteria in the preoperative prediction of choledocholithiasis. *Gastroenterology* 1997;**112**(4 Suppl S):A1456.

Maurea 2009 {published data only}

Maurea S, Caleo O, Mollica C, Imbriaco M, Mainenti PP, Palumbo C, et al. Comparative diagnostic evaluation with MR cholangiopancreatography, ultrasonography and CT in patients with pancreatobiliary disease. *Radiologia Medica* 2009;**114**(3):390-402.

Miao 2008 {published data only}

Miao L, Fan ZN, Ji GZ, Wen W, Wang X, Xiong GY, et al. Comparative study of ultrasonography, magnetic resonance

cholangiopancreatography and endoscopic retrograde cholangiopancreatography in common duct stones. *Chinese Journal of Surgery* 2008;**46**(19):1465-7.

Moon 2005 {published data only}

Moon JH, Cho YD, Cha SW, Cheon YK, Ahn HC, Kim YS, et al. The detection of bile duct stones in suspected biliary pancreatitis: comparison of MRCP, ERCP, and intraductal US. *American Journal of Gastroenterology* 2005;**100**(5):1051-7.

Neoptolemos 1986a {published data only}

Neoptolemos JP, London N, Bailey I. The role of clinical and biochemical criteria and endoscopic retrograde cholangiopancreatography in the urgent diagnosis of common bile duct stones in acute pancreatitis. *Surgery* 1986;**100**(4):732-42.

Neoptolemos 1986b {published data only}

Neoptolemos JP, London N, Bailey I, Shaw D, Carrlocke DL, Fossard DP. Detection of common bile-duct stones during acute gallstone pancreatitis - the role of biochemical criteria and ERCP. *British Journal of Surgery* 1986;**73**(12):1030-1.

Notash 2008 {published data only}

Notash AY, Salimi J, Golfam F, Habibi G, Alizadeh K. Preoperative clinical and paraclinical predictors of choledocholithiasis. *Hepatobiliary and Pancreatic Diseases International* 2008;**7**(3):304-7.

Oconnor 1986 {published data only}

Oconnor HJ, Hamilton I, Ellis WR, Watters J, Lintott DJ, Axon ATR. Ultrasound detection of choledocholithiasis - prospective comparison with ERCP in the postcholecystectomy patient. *Gastrointestinal Radiology* 1986;**11**(2):161-4.

Oehy 1987 {published data only}

Oehy K, Vogt B, Aubert de la Rue R, Hofer B. Comparison of imaging techniques in preoperative diagnosis of choledocholithiasis. *Schweizerische Medizinische Wochenschrift* 1987;**117**(29):1109-12.

Onken 1996 {published data only}

Onken JE, Brazer SR, Eisen GM, Williams DM, Bouras EP, DeLong ER, et al. Predicting the presence of choledocholithiasis in patients with symptomatic cholelithiasis. *American Journal of Gastroenterology* 1996;**91**(4):762-7.

Padda 2009 {published data only}

Padda MS, Singh S, Tang SJ, Rockey DC. Liver test patterns in patients with acute calculous cholecystitis and/or choledocholithiasis. *Alimentary Pharmacology & Therapeutics* 2009;**29**(9):1011-8.

Pamos 2003 {published data only}

Pamos S, Benages A, Medina E, Martinez Sanjuan V. Prospective evaluation of magnetic resonance cholangiopancreatography in patients with biliary disease: comparative study with conventional ultrasonography and endoscopic retrograde cholangiopancreatography diagnostic algorithm. *Digestive and Liver Disease* 2003;**35**(3):186-92.

Pancione 1994 {published data only}

Pancione L, Lupo F, Ballotto L, Ghezzi L, Ferro C. Ultrasonography and retrograde transpapillary cholangiopancreatography combined with sphincterotomy in the diagnosis of choledochal lithiasis. Our experience with 87 patients. *La Radiologia Medica* 1994;**88**(4):437-44.

Parra Perez 2007 {published data only}

Parra Perez V, Vargas Cardenas G, Astete Benavides M, Valdivia Roldan M, Moran Tisoc L, Nunez Calixto N, et al. Choledocolithiasis predictors in high-risk population subjected to endoscopic retrograde pancreatocholangiography at "Hospital Nacional Arzobispo Loayza". *Revista de Gastroenterologia del Peru* 2007;**27**(2):161-71.

Pasanen 1993 {published data only}

Pasanen PA, Partanen KP, Pikkarainen PH, Alhava EM, Janatuinen EK, Pirinen AE. A comparison of ultrasound, computed tomography and endoscopic retrograde cholangiopancreatography in the differential diagnosis of benign and malignant jaundice and cholestasis. *European Journal of Surgery, Acta Chirurgica* 1993;**159**(1):23-9.

Pasanen 1994 {published data only}

Pasanen PA, Partanen K, Pikkarainen P, Alhava E, Pirinen A, Janatuinen E. A prospective study on the value of ultrasound, computed tomography and endoscopic retrograde cholangiopancreatography in the diagnosis of unjaundiced cholestasis. *In Vivo* 1994;**8**(2):227-30.

Patwardhan 1987 {published data only}

Patwardhan RV, Smith OJ, Farmelant MH. Serum transaminase levels and cholescintigraphic abnormalities in acute biliary tract obstruction. *Archives of Internal Medicine* 1987;**147**(7):1249-53.

Pedersen 1987 {published data only}

Pedersen OM, Nordgard K, Kvinnsland S. Value of sonography in obstructive jaundice. Limitations of bile duct caliber as an index of obstruction. *Scandinavian Journal of Gastroenterology* 1987;**22**(8):975-81.

Peng 2005 {published data only}

Peng WK, Sheikh Z, Paterson-Brown S, Nixon SJ. Role of liver function tests in predicting common bile duct stones in acute calculous cholecystitis. *British Journal of Surgery* 2005;**92**(10):1241-7.

Pereira-Lima 2000 {published data only}

Pereira-Lima JC, Jakobs R, Busnelo JV, Benz C, Blaya C, Riemann JF. The role of serum liver enzymes in the diagnosis of choledocholithiasis. *Hepato-Gastroenterology* 2000;**47**(36):1522-5.

Pickuth 2000 {published data only}

Pickuth D, Spielmann RP. Detection of choledocholithiasis: comparison of unenhanced spiral CT, US, and ERCP. *Hepato-Gastroenterology* 2000;**47**(36):1514-7.

Pizzuto 1997 {published data only}

Pizzuto G, Surgo D, Nasorri L, Tretola V, Rallo G, Grasso E, et al. Prospective evaluation of a preoperative work-up in

patients candidated to laparoscopic cholecystectomy. *Chirurgia* 1997;**10**(1):5-9.

Pourseidi 2007 {published data only}

Pourseidi B, Khorram-Manesh A. Triple non-invasive diagnostic test for exclusion of common bile duct stones before laparoscopic cholecystectomy. *World Journal of Gastroenterology* 2007;**13**(43):5745-9.

Prian 1977 {published data only}

Prian GW, Norton LW, Eule J Jr, Eiseman B. Clinical indications and accuracy of grey scale ultrasonography in the patient with suspected biliary tract disease. *American Journal of Surgery* 1977;**134**(6):705-11.

Rakoczy 1985 {published data only}

Rakoczy A, Noszczyk W, Serafin-Krol M, Zalewski J. Value of ultrasonic examination for detection of biliary calculi. *Polski Przegląd Radiologii* 1985;**49**(1):27-9.

Rieger 1995 {published data only}

Rieger R, Wayand W. Yield of prospective, noninvasive evaluation of the common bile-duct combined with selective ERCP/sphincterotomy in 1390 consecutive laparoscopic cholecystectomy patients. *Gastrointestinal Endoscopy* 1995;**42**(1):6-12.

Rigauts 1992 {published data only}

Rigauts H, Marchal G, Vansteenbergen W, Ponette E. Comparison of ultrasound and ERCP in the detection of the cause of obstructive biliary disease. *Fortschritte Auf Dem Gebiete Der Rontgenstrahlen Und Der Neuen Bildgebenden Verfahren* 1992;**156**(3):252-7.

Robertson 1996 {published data only}

Robertson GS, Jagger C, Johnson PR, Rathbone BJ, Wicks AC, Lloyd DM, et al. Selection criteria for preoperative endoscopic retrograde cholangiopancreatography in the laparoscopic era. *Archives of Surgery* 1996;**131**(1):89-94.

Robinson 2003 {published data only}

Robinson TN, Biffl WL, Moore EE, Heimbach JK, Calkins CM, Burch J. Routine preoperative laboratory analyses are unnecessary before elective laparoscopic cholecystectomy. *Surgical Endoscopy and Other Interventional Techniques* 2003;**17**(3):438-41.

Roig 1995 {published data only}

Roig MVP, Espinosa RG, Rodero DR. Risk of common bile-duct stones - selective criteria for preoperative ERCP or intraoperative cholangiography in laparoscopic cholecystectomy. *British Journal of Surgery* 1995;**82**(Suppl 1):15.

Roston 1997 {published data only}

Roston AD, Jacobson IM. Evaluation of the pattern of liver tests and yield of cholangiography in symptomatic choledocholithiasis: a prospective study. *Gastrointestinal Endoscopy* 1997;**45**(5):394-9.

Salazar 1993 {published data only}

Salazar S, Useche E, Vetencourt R. Ultrasonography in the differential diagnosis of obstructive jaundice: comparison with endoscopic retrograde cholangiopancreatography. *GEN* 1993;**47**(2):73-7.

Saltzstein 1982 {published data only}

Saltzstein EC, Peacock JB, Thomas MD. Preoperative bilirubin, alkaline phosphatase and amylase levels as predictors of common duct stones. *Surgery Gynecology and Obstetrics* 1982;**154**(3):381-4.

Santucci 1996 {published data only}

Santucci L, Natalini G, Sarpi L, Fiorucci S, Solinas A, Morelli A. Selective endoscopic retrograde cholangiography and preoperative bile duct stone removal in patients scheduled for laparoscopic cholecystectomy: a prospective study. *American Journal of Gastroenterology* 1996;**91**(7):1326-30.

Sasukevich 2005 {published data only}

Sasukevich VN, Nazarenko VA. Ultrasonic diagnosis of choledocholithiasis. *Khirurgiia* 2005;**6**:58-60.

Sauerbruch 1979 {published data only}

Sauerbruch T, Wotzka R, Rehkamp D, Rummel T. Upper-abdominal ultrasonography and ERCP?. *Deutsche Medizinische Wochenschrift* 1979;**104**(5):165-6, 169-71.

Sgourakis 2005 {published data only}

Sgourakis G, Dedemadi G, Stamatelopoulos A, Leandros E, Voros D, Karaliotas K. Predictors of common bile duct lithiasis in laparoscopic era. *World Journal of Gastroenterology* 2005;**11**(21):3267-72.

Sheen 2008 {published data only}

Sheen AJ, Asthana S, Al-Mukhtar A, Attia M, Toogood GJ. Preoperative determinants of common bile duct stones during laparoscopic cholecystectomy. *International Journal of Clinical Practice* 2008;**62**(11):1715-9.

Sheen-Chen 1990 {published data only}

Sheen-Chen SM, Chou FF. Value of five predictive parameters for choledocholithiasis. *Journal of Surgical Association Republic of China* 1990;**23**(2):128-31.

Shiozawa 2005 {published data only}

Shiozawa S, Tsuchiya A, Kim DH, Usui T, Masuda T, Kubota K, et al. Useful predictive factors of common bile duct stones prior to laparoscopic cholecystectomy for gallstones. *Hepato-Gastroenterology* 2005;**52**(66):1662-5.

Shiozawa 2011 {published data only}

Shiozawa S, Kim DH, Usui T, Tsuchiya A, Masuda T, Inose S, et al. Indication of endoscopic retrograde cholangiography by noninvasive predictive factors of common bile duct stones before laparoscopic cholecystectomy: a prospective clinical study. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques* 2011;**21**(1):28-32.

Sigel 1983 {published data only}

Sigel B, Machi J, Beitler JC, Donahue PE, Bombeck CT, Baker RJ, et al. Comparative accuracy of operative ultrasonography and cholangiography in detecting common duct calculi. *Surgery* 1983;**94**(4):715-20.

Singh 2000 {published data only}

Singh G, Gupta PC, Sridar G, Katariya RN. Role of selective intra-operative cholangiography during cholecystectomy. *Australian and New Zealand Journal of Surgery* 2000;**70**(2):106-9.

Skorka 1982 {published data only}

Skorka B, Preuss HJ, Baldauf K. Informative value and validity of pre-, per- and postoperative radiological bile-duct diagnostics. *Zeitschrift fur Arztliche Fortbildung* 1982;**76**(1-2):24-9.

Soto 2000 {published data only}

Soto JA, Munera F, Alvarez OA, Alvarez LG. Diagnosis of choledocholithiasis: comparison of MR cholangiography and real-time ultrasonography. *Radiology* 2000;**217**(Suppl S):462.

Stain 1994 {published data only}

Stain SC, Masri LS, Froes ET, Sharma V, Parekh D. Laparoscopic cholecystectomy - laboratory predictors of choledocholithiasis. *American Surgeon* 1994;**60**(10):767-71.

Sugiyama 1997 {published data only}

Sugiyama M, Atomi Y. Endoscopic ultrasonography for diagnosing choledocholithiasis: a prospective comparative study with ultrasonography and computed tomography. *Gastrointestinal Endoscopy* 1997;**45**(2):143-6.

Swobodnik 1986 {published data only}

Swobodnik W, Seitz K, Kluppelberg U, Wechsler JG, Rettenmaier G, Ditschuneit H. Results of sonographic diagnosis of choledochal calculi. *Ultraschall in der Medizin* 1986;**7**(3):108-13.

Tham 1994 {published data only}

Tham TCK, Collins JSA, Watson RGP. Comparison of ultrasound scan and liver-function tests with endoscopic retrograde cholangiography in patients with suspected common bile-duct stones. *Gastroenterology* 1994;**106**(4 Suppl S):A361.

Tham 1996 {published data only}

Tham TC, Collins JS, Watson RG, Ellis PK, McIlrath EM. Diagnosis of common bile duct stones by intravenous cholangiography: prediction by ultrasound and liver function tests compared with endoscopic retrograde cholangiography. *Gastrointestinal Endoscopy* 1996;**44**(2):158-63.

Thornton 1992 {published data only}

Thornton JR, Lobo AJ, Lintott DJ, Axon AT. Value of ultrasound and liver function tests in determining the need for endoscopic retrograde cholangiopancreatography in unexplained abdominal pain. *Gut* 1992;**33**(11):1559-61.

Tobin 1984 {published data only}

Tobin MV, Gilmore IT, Lamb GHR. Choledocholithiasis - discrepancies between ultrasound and retrograde cholangiogram findings. *Gut* 1984;**25**(5):A560.

Tobin 1986 {published data only}

Tobin MV, Mendelson RM, Lamb GH, Gilmore IT. Ultrasound diagnosis of bile duct calculi. *British Medical Journal (Clinical Research Ed.)* 1986;**293**(6538):16-7.

Trondsen 1998 {published data only}

Trondsen E, Edwin B, Reiertsen O, Faerden AE, Fagertun H, Rosseland AR. Prediction of common bile duct stones prior to cholecystectomy - a prospective validation of a discriminant analysis function. *Archives of Surgery* 1998;**133**(2):162-6.

Tsyb 1991 {published data only}

Tsyb AF, Dergachev AI. Ultrasonic diagnosis of diseases of the gallbladder and bile ducts. *Vestnik Rentgenologii i Radiologii* 1991;**3**:75-81.

Uomo 1992 {published data only}

Uomo G, Rabitti PG, Laccetti M, Marcopido B, Picciotto FP, Visconti M, et al. The role of clinical, biochemical and echographic data in identifying the biliary pathogenesis of acute pancreatitis. *Recenti Progressi in Medicina* 1992;**83**(4):206-9.

van Santvoort 2009 {published data only}

van Santvoort HC, Bakker OJ, Besselink MG, Bollen TL, Fischer K, Gooszen H, et al. Predictors of common bile duct stones during early endoscopic retrograde cholangiopancreatography in acute biliary pancreatitis. *Pancreatology* 2009;**9**(4):434.

van Santvoort 2011 {published data only}

van Santvoort HC, Bakker OJ, Besselink MG, Bollen TL, Fischer K, Nieuwenhuijs VB, et al. Prediction of common bile duct stones in the earliest stages of acute biliary pancreatitis. *Endoscopy* 2011;**43**(1):8-13.

Varghese 2000 {published data only}

Varghese JC, Liddell RP, Farrell MA, Murray FE, Osborne DH, Lee MJ. Diagnostic accuracy of magnetic resonance cholangiopancreatography and ultrasound compared with direct cholangiography in the detection of choledocholithiasis. *Clinical Radiology* 2000;**55**(1):25-35.

Videhult 2011 {published data only}

Videhult P, Sandblom G, Rudberg C, Rasmussen IC. Are liver function tests, pancreatitis and cholecystitis predictors of common bile duct stones? Results of a prospective, population-based, cohort study of 1171 patients undergoing cholecystectomy. *HPB: the Official Journal of the International Hepato Pancreato Biliary Association* 2011;**13**(8):519-27.

Voyles 1994 {published data only}

Voyles CR, Sanders DL, Hogan R. Common bile-duct evaluation in the era of laparoscopic cholecystectomy - 1050 cases later. *Annals of Surgery* 1994;**219**(6):744-52.

Wang 1995 {published data only}

Wang CH, Mo LR, Yau MP, Lin RC, Kuo JY, Huang MH. Preoperative assessment of choledocholithiasis in laparoscopic cholecystectomy. *Journal of the Formosan Medical Association* 1995;**94**(5):228-31.

Wang 2001 {published data only}

Wang CH, Mo LR, Lin RC, Kuo JY, Chang KK. Rapid diagnosis of choledocholithiasis using biochemical tests in patients undergoing laparoscopic cholecystectomy. *Hepato-Gastroenterology* 2001;**48**(39):619-21.

Watkin 1994 {published data only}

Watkin DS, Haworth JM, Leaper DJ, Thompson MH. Assessment of the common bile duct before cholecystectomy using ultrasound and biochemical measurements: validation based on follow-up. *Annals of the Royal College of Surgeons of England* 1994;**76**(5):317-9.

Welbourn 1995 {published data only}

Welbourn CR, Haworth JM, Leaper DJ, Thompson MH. Prospective evaluation of ultrasonography and liver function tests for preoperative assessment of the bile duct. *British Journal of Surgery* 1995;**82**(10):1371-3.

Wermke 1987 {published data only}

Wermke W, Schulz HJ. Sonographic diagnosis of bile duct calculi. Results of a prospective study of 222 cases of choledocholithiasis. *Ultraschall in der Medizin* 1987;**8**(3):116-20.

Wermke 1992 {published data only}

Wermke W. Sonographic diagnosis of concretions of the bile ducts. A prospective study in respect of the effects of objective and subjective factors on the on-target efficacy in choledocholithiasis. *Ultraschall in der Medizin* 1992;**13**(6):246-54.

Yadav 2000 {published data only}

Yadav RK, Kaushal V, Griwan MS, Garg P, Popli P, Maken S, et al. Ultrasonography and computed tomography evaluation in patients of obstructive jaundice. *Asian Oceanian Journal of Radiology* 2000;**5**(3):138-43.

Yang 1990 {published data only}

Yang HL, Li ZZ, Sun YG. Reliability of ultrasonography in diagnosis of biliary lithiasis. *Chinese Medical Journal* 1990;**103**(8):638-41.

Yang 2008 {published data only}

Yang MH, Chen TH, Wang SE, Tsai YF, Su CH, Wu CW, et al. Biochemical predictors for absence of common bile duct stones in patients undergoing laparoscopic cholecystectomy. *Surgical Endoscopy* 2008;**22**(7):1620-4.

Yriberri 2007 {published data only}

Yriberri Urena S, Monge Zapa V. Alterations in laboratory tests as predictors of choledocholithiasis in patients who underwent ERCP: experience in a national private center. *Revista de Gastroenterologia del Peru* 2007;**27**(3):253-8.

Zusmer 1978 {published data only}

Zusmer NR, Harwood SJ, Pevsner NH. Grey scale ultrasonography: evaluating the jaundiced patient. *Southern Medical Journal* 1978;**71**(5):498-501.

Additional references

Abdallah 2007

Abdallah AA, Krige JEJ, Bornman PC. Biliary tract obstruction in chronic pancreatitis. *HPB* 2007;**9**(6):421-8.

Aljebreen 2008

Aljebreen A, Azzam N, Eloubeidi MA. Prospective study of endoscopic ultrasound performance in suspected choledocholithiasis. *Journal of Gastroenterology and Hepatology* 2008;**23**(5):241-5.

Anciaux 1986

Anciaux ML, Pelletier G, Attali P, Meduri B, Liguory C, Etienne JP. Prospective study of clinical and biochemical features of symptomatic choledocholithiasis. *Digestive Diseases and Sciences* 1986;**31**(5):449-53.

Arnold 1970

Arnold DJ, Zollinger RW, Bartlett RM, Telesz W, King RG, Strait J, et al. 28621 cholecystectomies in Ohio. Results of a survey in Ohio hospitals by the gallbladder survey committee, Ohio chapter, American College of Surgeons. *American Journal of Surgery* 1979;**119**(6):714-7.

ASGE Standards of Practice Committee 2010

Maple JT, Ben-Menachem T, Anderson MA, Appalaneni V, Banerjee S, ASGE Standards of Practice Committee 2010. The role of endoscopy in the evaluation of suspected choledocholithiasis. *Gastrointestinal Endoscopy* 2010;**71**(1):1-9.

Attili 1995

Attili AF, De Santis A, Capri R, Repice AM, Maselli S. The natural history of gallstones: the GREPCO experience. The GREPCO group. *Hepatology* 1995;**21**(3):655-60.

Barbara 1987

Barbara L, Sama C, Morselli Labate AM, Taroni F, Rusticali AG, Festi D, et al. A population based study on the prevalence of gallstone disease: the Sirmione study. *Hepatology* 1987;**7**(5):913-7.

Benjaminov 2013

Benjaminov F, Leichtman G, Naftali T, Half EE, Konikoff FM. Effects of age and cholecystectomy on common bile duct diameter as measured by endoscopic ultrasonography. *Surgical Endoscopy* 2013;**27**(1):303-7.

Buscemi 2006

Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. *Journal of Clinical Epidemiology* 2006;**59**(7):697-703.

Caddy 2005

Caddy GR, Kirby J, Kirk SJ, Allen MJ, Moorehead RJ, Tham TC. Natural history of asymptomatic bile duct stones at time of cholecystectomy. *Ulster Medical Journal* 2005;**74**(2):108-12.

Caddy 2006

Caddy GR, Tham DC. Gallstone disease: symptoms, diagnosis and endoscopic management of common bile duct stones. *Best Practice & Research. Clinical Gastroenterology* 2006;**20**(6):1085-101.

Chen 2007

Chen YK, Pleskow DK. SpyGlass single-operator peroral cholangiopancreatography system for the diagnosis and therapy of bile duct disorders: a clinical feasibility study (with video). *Gastrointestinal Endoscopy* 2007;**65**(6):832-41.

Choi 2011

Choi SB, Kim Wb, Song TJ, Suh SO, Kim YC, Choi SY. Surgical outcomes and prognostic factors for ampulla of Vater cancer. *Scandinavian Journal of Surgery* 2011;**100**(2):92-8.

Chu 2006

Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology* 2006;**59**(12):1331-2.

Corfield 1985

Corfield AP, Cooper MJ, Williamson RCN. Acute pancreatitis: a lethal disease of increasing incidence. *Gut* 1985;**26**(7):724-9.

Dasari 2013

Dasari BV, Tan CJ, Gurusamy KS, Martin DJ, Kirk G, McKie L, et al. Surgical versus endoscopic treatment of bile duct stones. *Cochrane Database of Systematic Reviews* 2013, Issue 9. [DOI: [10.1002/14651858.CD003327.pub3](https://doi.org/10.1002/14651858.CD003327.pub3)]

Doust 2005

Doust JA, Pietrzak E, Sanders S, Glasziou PP. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *Journal of Clinical Epidemiology* 2005;**58**(5):444-9.

Eastman 1977

Eastman JR, Bixler D. Serum alkaline phosphatase: normal values by sex and age. *Clinical Chemistry* 1977;**23**(9):1769-70.

Eloubeidi 2001

Eloubeidi MA, Wade SB, Provenzale D. Factors associated with acceptance and full publication of GI endoscopic research originally published in abstract form. *Gastrointestinal Endoscopy* 2001;**53**(3):275-82.

FDA 2013

US Food, Drug Administration. Investigations operations manual, 2013. www.fda.gov/iceci/inspections/iom/default.htm (accessed 23 January 2014).

Fiore 1997

Fiore NF, Ledniczky G, Wiebke EA, Broadie TA, Pruitt AL, Goulet RJ, et al. An analysis of perioperative cholangiography in one thousand laparoscopic cholecystectomies. *Surgery* 1997;**122**(4):817-23.

Freitas 2006

Freitas ML, Bell RL, Duffy AJ. Choledocholithiasis: evolving standard for diagnosis and management. *World Journal of Gastroenterology* 2006;**12**(20):3162-7.

Geron 1999

Geron N, Reshef R, Shiller M, Kniaz D, Eitan A. The role of endoscopic retrograde cholangiopancreatography in the laparoscopic era. *Surgical Endoscopy, Ultrasound and Interventional Techniques* 1999;**13**(5):452-6.

Giannini 2005

Giannini EG, Testa R, Savarino V. Liver enzyme alteration: a guide for clinicians. *Canadian Medical Association Journal* 2005;**172**(3):367-79.

Giljaca 2015

Giljaca V, Gurusamy KS, Takwoingi Y, Higgie D, Poropat G, Štimac D, et al. Endoscopic ultrasound versus magnetic resonance cholangiopancreatography for common bile duct stones. *Cochrane Database of Systematic Reviews* 2015, Issue 2. [DOI: [10.1002/14651858.CD011549](https://doi.org/10.1002/14651858.CD011549)]

Gurusamy 2010

Gurusamy KS, Davidson BR. Surgical treatment of gallstones. *Gastroenterology Clinics of North America* 2010;**39**(2):229-44, viii.

Gurusamy 2011

Gurusamy K, Sahay SJ, Burroughs AK, Davidson BR. Systematic review and meta-analysis of intraoperative versus preoperative endoscopic sphincterotomy in patients with gallbladder and suspected common bile duct stones. *British Journal of Surgery* 2011;**98**(7):908-16.

Gurusamy 2015

Gurusamy KS, Giljaca V, Higgie D, Poropat G, Štimac D, Davidson BR. Endoscopic retrograde cholangiopancreatography versus intraoperative cholangiography for diagnosis of common bile duct stones. *Cochrane Database of Systematic Reviews* 2015, Issue 2. [DOI: [10.1002/14651858.CD010339](https://doi.org/10.1002/14651858.CD010339)]

Halldestam 2004

Halldestam I, Enell EL, Kullman E, Borch K. Development of symptoms and complications in individuals with asymptomatic gallstones. *British Journal of Surgery* 2004;**91**(6):734-8.

Hamade 2005

Hamade AM, Al-Bahrani AZ, Oweru AMA, Hamoodi AA, Abid GH, Bani Hani OI, et al. Therapeutic, prophylactic, and pre-resection applications of laparoscopic gastric and biliary bypass for patients with periampullary malignancy. *Surgical Endoscopy* 2005;**19**(10):1333-40.

Hayat 2005

Hayat JO, Loew CJ, Asress KN, McIntyre AS, Gorard DA. Contrasting liver function test patterns in obstructive jaundice due to biliary strictures [corrected] and stones. *QJM* 2005;**98**(1):35-40.

Hirschfield 2011

Hirschfield GM, Gershwin ME. Primary biliary cirrhosis: one disease with many faces. *Israel Medical Association Journal* 2011;**13**(1):55-9.

Hong 2006

Hong DF, Xin Y, Chen DW. Comparison of laparoscopic cholecystectomy combined with intraoperative endoscopic sphincterotomy and laparoscopic exploration of the common bile duct for cholecystocholedocholithiasis. *Surgical Endoscopy and Other Interventional Techniques* 2006;**20**(3):424-7.

Hoyuela 1999

Hoyuela C, Cougat E, Pretcha P, Collera P, Espinos J, Marco C. Must ERCP be routinely performed if choledocholithiasis is suspected?. *Digestive Surgery* 1999;**16**(5):411-4.

Kalady 2004

Kalady MF, Peterson B, Baillie J, Onaitis MW, Abdul-Wahab OI, Howden JK, et al. Pancreatic duct strictures: identifying risk of malignancy. *Annals of Surgical Oncology* 2004;**11**(6):581-8.

Kelly 2010

Kelly MD. Results of laparoscopic bile duct exploration via choledochotomy. *ANZ Journal of Surgery* 2010;**80**(10):694-8.

Krishna 2008

Krishna RP, Kumar A, Singh RK, Sikora S, Saxena R, Kapoor VK. Xanthogranulomatous inflammatory strictures of extrahepatic biliary tract: presentation and surgical management. *Journal of Gastrointestinal Surgery* 2008;**12**(5):836-41.

Lefemine 2011

Lefemine V, Morgan RJ. Spontaneous passage of common bile duct stones in jaundiced patients. *Hepatobiliary and Pancreatic Diseases International* 2011;**10**(2):209-13.

Lill 2010

Lill S, Rantala A, Pekkala E, Sarparanta H, Huhtinen H, Rautava P, et al. Elective laparoscopic cholecystectomy without routine intraoperative cholangiography: a retrospective analysis of 1101 consecutive cases. *Scandinavian Journal of Surgery* 2010;**99**(4):197-200.

Lillemoen 2000

Lillemoen KD, Melton GB, Cameron JL, Pitt HA, Campbell KA, Talamini MA, et al. Postoperative bile duct strictures: management and outcome in the 1990s. *Annals of Surgery* 2000;**232**(3):430-41.

Lin 2008

Lin LF, Huang PT, Ho KS, Tung JN. Autoimmune chronic pancreatitis. *Journal of the Chinese Medical Association* 2008;**71**(1):14-22.

Loria 1994

Loria P, Dilengite MA, Bozzoli M, Carubbi F, Messori R, Sassatelli R, et al. Prevalence rates of gallstone disease in Italy. The Chianciano population study. *European Journal of Epidemiology* 1994;**10**(2):143-50.

Ludwig 2001

Ludwig K, Kockerling F, Hohenberger W, Lorenz D. Surgical therapy for cholecysto-/choledocholithiasis. Results of a Germany-wide questionnaire sent to 859 clinics including 123,090 cholecystectomies [Die chirurgische Therapie der cholecysto-/choledocholithiasis. Ergebnisse einer deutschlandweiten Umfrage an 859 Kliniken mit 123.090 cholecystektomien]. *Der Chirurg* 2001;**72**(10):1170-8.

Mann 1994

Mann DV, Hershman MJ, Hittinger R, Glazer G. Multicentre audit of death from acute pancreatitis. *British Journal of Surgery* 1994;**81**(6):890-3.

Maple 2010

Maple JT, Ben-Menachem T, Anderson MA, Appalaneeni V, Banerjee S, ASGE Standards of Practice Committee 2010. The role of endoscopy in the evaluation of suspected choledocholithiasis. *Gastrointestinal Endoscopy* 2010;**71**(1):1-9.

Neoptolemos 1988

Neoptolemos JP, London NJ, James D, Carr-Locke DL, Bailey IA, Fossard DP. Controlled trial of urgent endoscopic retrograde cholangiopancreatography and endoscopic sphincterotomy versus conservative treatment for acute pancreatitis due to gallstones. *Lancet* 1988;**2**(8618):979-83.

Park 2011

Park SY, Park CH, Cho SB, Lee WS, Kim JC, Cho CK, et al. What is appropriate procedure for preoperative biliary drainage in patients with obstructive jaundice awaiting pancreaticoduodenectomy. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques* 2011;**21**(5):344-8.

Penz-Osterreicher 2011

Penz-Osterreicher M, Osterreicher CH, Trauner M. Fibrosis in autoimmune and cholestatic liver disease. *Best Practices & Research. Clinical Gastroenterology* 2011;**25**(2-4):245-58.

Prat 1996

Prat F, Amouyal G, Amouyal P, Pelletier G, Fritsch J, Choury AD, et al. Prospective controlled study of endoscopic ultrasonography and endoscopic retrograde cholangiography in patients with suspected common-bileduct lithiasis. *Lancet* 1996;**347**(8994):75-9.

RadiologyInfo 2011

RadiologyInfo. What is ultrasound imaging of the abdomen?. www.radiologyinfo.org/en/info.cfm?pg=abdominus (accessed 25 October 2012).

Raraty 1998

Raraty MG, Finch M, Neoptolemos JP. Acute cholangitis and pancreatitis secondary to common duct stones: management update. *World Journal of Surgery* 1998;**22**(11):1155-61.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982-90.

RevMan 2012 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

Ripolles 2009

Ripolles T, Ramirez-Fuentes C, Martinez-Perez MJ, Delgado F, Blanc E, Lopez A. Tissue harmonic sonography in the diagnosis of common bile duct stones: a comparison with endoscopic retrograde cholangiography. *Journal of Clinical Ultrasound* 2009;**37**(9):501-6.

Rosalki 1971

Rosalki SB, Tarlow D, Rau D. Plasma gamma-glutamyl transpeptidase elevation in patients receiving enzyme-inducing drugs. *Lancet* 1971;**2**(7720):376-7.

Saik 1975

Saik RP, Greenburg AG, Farris JM, Peksin GW. Spectrum of cholangitis. *American Journal of Surgery* 1975;**130**(2):143-50.

Salek 2009

Salek J, Livote E, Sideridis K, Bank S. Analysis of risk factors predictive of early mortality and urgent ERCP in acute cholangitis. *Journal of Clinical Gastroenterology* 2009;**43**(2):171-5.

Sampson 2008

Sampson M, Shojania KG, McGowan J, Daniel R, Rader T, Iansavichene AE, et al. Surveillance search techniques identified the need to update systematic reviews. *Journal of Clinical Epidemiology* 2008;**61**(8):755-62.

Sarli 2000

Sarli L, Costi R, Gobbi S, Sansebastiano G, Roncoroni L. Asymptomatic bile duct stones: selection criteria for intravenous cholangiography and/or endoscopic retrograde cholangiography prior to laparoscopic cholecystectomy. *European Journal of Gastroenterology & Hepatology* 2000;**12**(11):1175-80.

Schuetz 2012

Schuetz GM, Schlattmann P, Dewey M. Use of 3x2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studies. *BMJ (Clinical Research Edition)* 2012;**345**:e6717.

Siddiqui 2011

Siddiqui AA, Sreenarasimhaiah J, Lara LF, Harford W, Lee C, Eloubeidi MA. Endoscopic ultrasound-guided transduodenal placement of a fully covered metal stent for palliative biliary drainage in patients with malignant biliary obstruction. *Surgical Endoscopy* 2011;**25**(2):549-55.

Singh 1990

Singh SM, Londmire WP, Reber HA. Surgical palliation for pancreatic cancer: the UCLA experience. *Annals of Surgery* 1990;**212**(2):132-9.

Spelsberg 2009

Spelsberg FW, Nusser F, Huttli TK, Obeidat FW, Lang RA, Jauch KW, et al. Management of cholecysto- and choledocholithiasis - survey and analysis of 16 615 cholecystectomies and common bile duct explorations in Bavaria. *Zentralblatt fur Chirurgie* 2009;**134**(2):120-6.

Stiris 2000

Stiris MG, Tennoe B, Aadland E, Lunde OC. MR cholangiopancreatography and endoscopic retrograde cholangiopancreatography in patients with suspected common bile duct stones. *Acta Radiologica* 2000;**41**(3):269-72.

Tang 2011

Tang S, Singh S, Singh S. Sphincterotome stricturoplasty for long ampullary stenoses and benign biliary strictures (with video). *Surgical Endoscopy* 2011;**25**(4):1313-8.

Targarona 2004

Targarona EM, Bendahan GE. Management of common bile duct stones: controversies and future perspectives. *HPB* 2004;**6**(3):140-3.

Toh 2000

Toh SKC, Phillips S, Johnson CD. A prospective audit against national standards of the presentation and management of acute pancreatitis in the south of England. *Gut* 2000;**46**(2):239-43.

Tranter 2003

Tranter SE, Thompson MH. Spontaneous passage of bile duct stones: frequency of occurrence and relation to clinical

presentation. *Annals of the Royal College of Surgeons of England* 2003;**85**(3):174-7.

Whiting 2006

Whiting PF, Weswood ME, Rutjes AWS, Reitsma JB, Bossuyt PNM, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Medical Research Methodology* 2006;**6**:9.

Whiting 2011

Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, the QUADAS-2 group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

Williams 2008

Williams EJ, Green J, Beckingham I, Parks R, Martin D, Lombard M. Guidelines on the management of common bile duct stones (CBDS). *Gut* 2008;**57**(7):1004-21.

Yousefpour Azary 2011

Yousefpour Azary S, Kalbasi H, Setayesh A, Mousavi M, Hashemi A, Khodadoostan M, et al. Predictive value and main determinants of abnormal features of intraoperative cholangiography during cholecystectomy. *Hepatobiliary & Pancreatic Diseases International* 2011;**10**(3):308-12.

References to other published versions of this review
Giljaca 2013

Giljaca V, Gurusamy KS, Vaughan J, Stimac D, Davidson BR. Tests for diagnosis of common bile duct stones. *Cochrane Database of Systematic Reviews* 2013, Issue 1. [DOI: [10.1002/14651858.CD010339](https://doi.org/10.1002/14651858.CD010339)]

CHARACTERISTICS OF STUDIES
Characteristics of included studies [ordered by study ID]
Admassie 2005
Study characteristics

Patient sampling	Type of study: retrospective. Consecutive or random sample: unclear.
Patient characteristics and setting	Sample size: 49. Females: 34 (69.4%). Age: 50 years. Presentation: Inclusion criteria: 1. People undergoing surgery for obstructive jaundice. Setting: Surgery Department, Ethiopia.
Index tests	Index test: ultrasound. Technical specifications: not stated.

Ultrasound versus liver function tests for diagnosis of common bile duct stones (Review)

Admassie 2005 (Continued)

 Performed by: not stated.
 Criteria for positive diagnosis: not stated.

Target condition and reference standard(s)	Target condition: common bile duct stones. Reference standard: surgical confirmation of cause for obstructive jaundice. Technical specifications: not applicable. Performed by: surgeons. Criteria for positive diagnosis: surgical confirmation of cause for obstructive jaundice.
Flow and timing	Number of indeterminates for whom the results of reference standard was available: not stated. Number of participants who were excluded from the analysis: not stated.
Comparative	
Notes	Attempted to contact the authors in June 2013. Received no replies.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Unclear		
		High	High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	High
DOMAIN 4: Flow and Timing			

Admassie 2005 (Continued)

Was there an appropriate interval between index test and reference standard?	Unclear
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Unclear
Unclear	

Busel 1989
Study characteristics

Patient sampling	Type of study: retrospective. Consecutive or random sample: unclear.
Patient characteristics and setting	Sample size: 38. Females: not stated. Age: not stated. Presentation: Inclusion criteria: 1. People undergoing common bile duct exploration for common bile duct stones. Setting: Surgery and Gastroenterology Departments, Chile.
Index tests	Index test: ultrasound. Technical specifications: sectorial transducer 3.5 MHz (manufacturer not stated). Performed by: not stated. Criteria for positive diagnosis: echogenic image with or without acoustic shadow in the common bile duct.
Target condition and reference standard(s)	Target condition: common bile duct stones. Reference standard: surgical exploration of common bile duct. Technical specifications: not applicable. Performed by: surgeons. Criteria for positive diagnosis: surgical exploration of common bile duct.
Flow and timing	Number of indeterminates for whom the results of reference standard was available: not stated. Number of participants who were excluded from the analysis: not stated.
Comparative	
Notes	Attempted to contact the authors in June 2013. Received no replies.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
------	--------------------	--------------	------------------------

Busel 1989 *(Continued)*
DOMAIN 1: Patient Selection

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
		High	High

DOMAIN 2: Index Test All tests

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
		High	Low

DOMAIN 3: Reference Standard

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	High

DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
		Unclear	

Kumar 1998
Study characteristics

Patient sampling	Type of study: prospective. Consecutive or random sample: consecutive.
Patient characteristics and setting	Sample size: 50. Females: 29 (58.0%). Age: 55 years. Presentation:

Kumar 1998 (Continued)

Inclusion criteria:

1. People with surgical obstructive jaundice.

Setting: Department of Surgery, India.

Index tests	Index test: ultrasound. Technical specifications: ATL MK 600 (or UM 9) 3.5 MHz scanner. Performed by: radiologist. Criteria for positive diagnosis: echogenic shadow in the dilated common bile duct suggestive of common bile duct stone.
Target condition and reference standard(s)	Target condition: common bile duct stones. Reference standard: surgical confirmation of cause for obstructive jaundice. Technical specifications: not applicable. Performed by: surgeons. Criteria for positive diagnosis: surgical confirmation of cause for obstructive jaundice.
Flow and timing	Number of indeterminates for whom the results of reference standard was available: 0. Number of participants who were excluded from the analysis: 0.
Comparative	
Notes	Attempted to contact the authors in June 2013. Received reply in June 2013.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Kumar 1998 (Continued)

Were the reference standard results interpreted without knowledge of the results of the index tests? No

High
Low
DOMAIN 4: Flow and Timing

Was there an appropriate interval between index test and reference standard? Yes

Did all patients receive the same reference standard? Yes

Were all patients included in the analysis? Yes

Low
Rickes 2006
Study characteristics

Patient sampling	Type of study: prospective. Consecutive or random sample: unclear.
Patient characteristics and setting	<p>Sample size: 126.</p> <p>Females: 86 (68.3%).</p> <p>Age: 63 years.</p> <p>Presentation:</p> <p>Inclusion criteria</p> <ol style="list-style-type: none"> 1. People with suspected choledocholithiasis on the basis of a combination of epigastric or right upper quadrant pain with fever or jaundice, acute pancreatitis, or abnormal liver function tests. <p>Exclusion criteria:</p> <ol style="list-style-type: none"> 1. Daily alcohol intake exceeded 80 g. 2. Taking hepatotoxic drugs. 3. Serum hepatitis B or C antibodies were present. 4. Refused to undergo ultrasound or endoscopic retrograde cholangiopancreatography. <p>Setting: Department of Gastroenterology, Hepatology and Infectious Diseases, Germany.</p>
Index tests	<p>Index test: ultrasound.</p> <p>Technical specifications: Siemen Elegra (2 to 5 MHz).</p> <p>Performed by: experienced and inexperienced operators.</p> <p>Criteria for positive diagnosis: hyperechoic structure within the common bile duct.</p>
Target condition and reference standard(s)	<p>Target condition: common bile duct stones.</p> <p>Reference standard: endoscopic extraction of stones in all participants.</p> <p>Technical specifications: not stated.</p> <p>Performed by: experienced endoscopists.</p> <p>Criteria for positive diagnosis: endoscopic extraction of stones in all participants.</p>

Rickes 2006 (Continued)

Flow and timing
 Number of indeterminates for whom the results of reference standard was available: not stated.
 Number of participants who were excluded from the analysis: 2 (1.6%).

Comparative

Notes
 Attempted to contact the authors in June 2013. Received no replies.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Silverstein 1998

Study characteristics

Patient sampling	Type of study: prospective. Consecutive or random sample: consecutive.
Patient characteristics and setting	Sample size: 262. Females: 206 (78.6%). Age: 47 years. Presentation: Inclusion criteria: 1. People undergoing cholecystectomy with ≥ 1 of the following features 2. Abnormal liver function tests. 3. History of obstructive jaundice or pancreatitis. 4. Common bile duct diameter > 6 mm in calibre on ultrasound. Setting: Department of Surgery, USA.
Index tests	Index test: ultrasound. Technical specifications: not stated. Performed by: not stated. Criteria for positive diagnosis: common bile duct diameter > 6 mm. Index test: bilirubin. Technical specifications: not applicable. Performed by: not applicable. Criteria for positive diagnosis: cut-offs used to define test positivity were > 22.23 $\mu\text{mol/L}$ and > twice the normal limit. Index test: alkaline phosphatase. Technical specifications: not applicable. Performed by: not applicable. Criteria for positive diagnosis: cut-offs used to define test positivity were > 125 IU/L and > twice the normal limit.
Target condition and reference standard(s)	Target condition: common bile duct stones. Reference standard: endoscopic or surgical extraction of stones in people with positive intraoperative cholangiogram and clinical follow-up of minimum 2 years in other people (participants underwent selective cholangiogram). Technical specifications: not applicable. Performed by: endoscopists, surgeons, and clinicians. Criteria for positive diagnosis: endoscopic or surgical extraction of stones in people with positive intraoperative cholangiogram and clinical follow-up of minimum 2 years in other people (participants underwent selective cholangiogram).
Flow and timing	Number of indeterminates for whom the results of reference standard was available: not stated. Number of participants who were excluded from the analysis: not stated.
Comparative	
Notes	Attempted to contact the authors in June 2013. Received no replies.

Methodological quality

Silverstein 1998 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
		High	Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
		High	High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
		High	

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abdul Ghani 1989	Inappropriate reference standard.
Al Samman 2000	Inappropriate reference standard.
Almersjo 1966	Inappropriate reference standard.
Ang 2007	Inappropriate reference standard.
Arroyo 1989	Inappropriate reference standard.
Bardach 1984	Inappropriate reference standard.
Barkun 1994	Inappropriate reference standard.
Barr 1999	Inappropriate reference standard.
Bhatt 2005	Insufficient diagnostic test accuracy information.
Birkigt 1989	Inappropriate reference standard.
Blackbourne 1994	Inappropriate reference standard.
Boboev 2012	Inappropriate reference standard.
Bocchetti 1989	Insufficient diagnostic test accuracy information.
Bokobza 1988	Not a diagnostic test accuracy study.
Borsch 1985	Inappropriate reference standard.
Bose 2001	Inappropriate reference standard.
Calero Ayala 1983	Insufficient diagnostic test accuracy information.
Cao 1986	Insufficient diagnostic test accuracy information.
Carrascosa 1983	Inappropriate reference standard.
Chan 2008	Inappropriate reference standard.
Chang 1998	Inappropriate reference standard.
Chavez-Valencia 2009	Inappropriate reference standard.
Cockbain 2010	Inappropriate reference standard.
Contini 1995	Inappropriate reference standard.
Contini 1997	Inappropriate reference standard.
Contractor 1997	Inappropriate reference standard.
Cranley 1980	Inappropriate reference standard.
Cwik 2003	Inappropriate reference standard.

Study	Reason for exclusion
Danaci 2002	Inappropriate reference standard.
de Dios Vega 1982	Insufficient diagnostic test accuracy information.
Delacruzcaro 1963	Insufficient diagnostic test accuracy information.
Derodra 1986	Not a diagnostic test accuracy study.
Dwivedi 1989	Insufficient diagnostic test accuracy information.
Ellul 1991	Inappropriate reference standard.
Espinoza 1984	Inappropriate reference standard.
Famos 1990	Inappropriate reference standard.
Fujino 1974	Inappropriate reference standard.
Garcia-Caballero 1994	Insufficient diagnostic test accuracy information.
Gierblinski 2003	Inappropriate reference standard.
Gologan 2006	Inappropriate reference standard.
Gregg 1979	Inappropriate reference standard.
Gross 1983	Inappropriate reference standard.
Gui 1995	Inappropriate reference standard.
Hammarstrom 1998	Inappropriate reference standard.
Holzinger 1999	Inappropriate reference standard.
Hunt 1990	Inappropriate reference standard.
Jarvinen 1978	Inappropriate reference standard.
Jolobe 2012	Insufficient diagnostic test accuracy information.
Jovanovic 2012	Not a diagnostic test accuracy study.
Katz 2004	Insufficient diagnostic test accuracy information.
Kelly 2011	Inappropriate reference standard.
Khan 2011	Insufficient diagnostic test accuracy information.
Kruis 1997	Inappropriate reference standard.
Lacaine 1980	Insufficient diagnostic test accuracy information.
Lakatos 2004	Inappropriate reference standard.
Lakoma 1996	Inappropriate reference standard.

Study	Reason for exclusion
Leander 1994	Inappropriate reference standard.
Lee 2008	Inappropriate reference standard.
Leitman 1993	Inappropriate reference standard.
Lichtenbaum 2000	Inappropriate reference standard.
Lindsell 1990	Inappropriate reference standard.
Link 1995	Inappropriate reference standard.
Liu 1984	Inappropriate reference standard.
Lomanto 1999	We were unable to obtain this reference.
Lui 1986	Insufficient diagnostic test accuracy information.
Magnuson 1997	Inappropriate reference standard.
Maurea 2009	Insufficient diagnostic test accuracy information.
Miao 2008	Insufficient diagnostic test accuracy information.
Moon 2005	Inappropriate reference standard.
Neoptolemos 1986a	Inappropriate reference standard.
Neoptolemos 1986b	Insufficient diagnostic test accuracy information.
Notash 2008	Insufficient diagnostic test accuracy information.
Oconnor 1986	Inappropriate reference standard.
Oehy 1987	Insufficient diagnostic test accuracy information.
Onken 1996	Inappropriate reference standard.
Padda 2009	Inappropriate reference standard.
Pamos 2003	Inappropriate reference standard.
Pancione 1994	Inappropriate reference standard.
Parra Perez 2007	Inappropriate reference standard.
Pasanen 1993	Insufficient diagnostic test accuracy information.
Pasanen 1994	Inappropriate reference standard.
Patwardhan 1987	Inappropriate reference standard.
Pedersen 1987	Insufficient diagnostic test accuracy information.
Peng 2005	Insufficient diagnostic test accuracy information.

Study	Reason for exclusion
Pereira-Lima 2000	Insufficient diagnostic test accuracy information.
Pickuth 2000	Inappropriate reference standard.
Pizzuto 1997	Insufficient diagnostic test accuracy information.
Pourseidi 2007	Inappropriate reference standard.
Priani 1977	Inappropriate reference standard.
Rakoczy 1985	Inappropriate reference standard.
Rieger 1995	Insufficient diagnostic test accuracy information.
Rigauts 1992	Inappropriate reference standard.
Robertson 1996	Inappropriate reference standard.
Robinson 2003	Inappropriate reference standard.
Roig 1995	Insufficient diagnostic test accuracy information.
Roston 1997	Insufficient diagnostic test accuracy information.
Salazar 1993	Inappropriate reference standard.
Saltzstein 1982	Inappropriate reference standard.
Santucci 1996	Inappropriate reference standard.
Sasukevich 2005	Inappropriate reference standard.
Sauerbruch 1979	Inappropriate reference standard.
Sgourakis 2005	Insufficient diagnostic test accuracy information.
Sheen 2008	Inappropriate reference standard.
Sheen-Chen 1990	We were unable to obtain this reference.
Shiozawa 2005	Insufficient diagnostic test accuracy information.
Shiozawa 2011	Insufficient diagnostic test accuracy information.
Sigel 1983	Inappropriate reference standard.
Singh 2000	Inappropriate reference standard.
Skorka 1982	Inappropriate reference standard.
Soto 2000	We were unable to obtain this reference.
Stain 1994	Inappropriate reference standard.
Sugiyama 1997	Inappropriate reference standard.

Study	Reason for exclusion
Swobodnik 1986	Inappropriate reference standard.
Tham 1994	Inappropriate reference standard.
Tham 1996	Inappropriate reference standard.
Thornton 1992	Inappropriate reference standard.
Tobin 1984	Inappropriate reference standard.
Tobin 1986	Inappropriate reference standard.
Trondsen 1998	Insufficient diagnostic test accuracy information.
Tsyb 1991	Inappropriate reference standard.
Uomo 1992	Inappropriate reference standard.
van Santvoort 2009	Insufficient diagnostic test accuracy information.
van Santvoort 2011	Inappropriate reference standard.
Varghese 2000	Inappropriate reference standard.
Videhult 2011	Inappropriate reference standard.
Voyles 1994	Insufficient diagnostic test accuracy information.
Wang 1995	Inappropriate reference standard.
Wang 2001	Inappropriate reference standard.
Watkin 1994	Inappropriate reference standard.
Welbourn 1995	Inappropriate reference standard.
Wermke 1987	Inappropriate reference standard.
Wermke 1992	Inappropriate reference standard.
Yadav 2000	We were unable to obtain this reference.
Yang 1990	Inappropriate reference standard.
Yang 2008	Inappropriate reference standard.
Yriberry 2007	Inappropriate reference standard.
Zusmer 1978	Inappropriate reference standard.

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Ultrasound	5	523
2 Bilirubin (> 22.23 µmol/L)	1	262
3 Bilirubin (> twice the normal limit)	1	261
4 Alkaline phosphatase (> 125 IU/L)	1	262
5 Alkaline phosphatase (> twice the normal limit)	1	261

Test 1. Ultrasound.

Test 2. Bilirubin (> 22.23 µmol/L).

Test 3. Bilirubin (> twice the normal limit).

Test 4. Alkaline phosphatase (> 125 IU/L).

Test 5. Alkaline phosphatase (> twice the normal limit).

ADDITIONAL TABLES

Table 1. Application of the QUADAS-2 tool for assessing methodological quality of included studies

Domain	Signalling question	Signalling question	Signalling question	Risk of bias	Concerns for applicability
Domain 1: Participant sampling					
Participant sampling	Was a consecutive or random sample of participants enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?	Could the selection of participants have introduced bias?	Are there concerns that the included participants and setting did not match the review question?

Table 1. Application of the QUADAS-2 tool for assessing methodological quality of included studies (Continued)

	<p>Yes: all consecutive participants or random sample of participants with suspected common bile duct stones were enrolled.</p> <p>No: selected participants were enrolled.</p> <p>Unclear: this was not clear from the report.</p>	<p>Yes: case-control design was avoided.</p> <p>No: case-control design was not avoided.</p> <p>Unclear: this was not clear from the report.</p>		<p>Yes: the study avoided inappropriate exclusions (i.e., people who were difficult to diagnose).</p> <p>No: the study excluded participants inappropriately.</p> <p>Unclear: this was not clear from the report.</p>	<p>Low risk: 'yes' for all signalling questions.</p> <p>High risk: 'no' or 'unclear' for at least 1 signalling question.</p>	<p>Low concern: the selected participants represented the people in whom the tests would be used in clinical practice (see diagnostic pathway (Figure 1)).</p> <p>High concern: there was high concern that participant selection was performed in a such a way that the included participants did not represent the people in whom the tests would be used in clinical practice.</p>
Domain 2: Index test						
Index test(s)	Were the index test results interpreted without knowledge of the results of the reference standard?	If a threshold was used, was it pre-specified?	-	Could the conduct or interpretation of the index test have introduced bias?	Were there concerns that the index test, its conduct, or interpretation differed from the review question?	
	<p>Yes: index test results were interpreted without knowledge of the results of the reference standard.</p> <p>No: index test results were interpreted with knowledge of the results of the reference standard.</p> <p>Unclear: this was not clear from the report.</p>	<p>Yes: if the criteria for a positive test result were pre-specified.</p> <p>No: if the criteria for a positive test result were not pre-specified.</p> <p>Unclear: this was not clear from the report.</p>	-	<p>Low risk: 'yes' for all signalling questions.</p> <p>High risk: 'no' or 'unclear' for at least 1 of the 2 signalling questions.</p>	<p>High concern: there was high concern that the conduct or interpretation of the index test differed from the way it was likely to be used in clinical practice.</p> <p>Low concern: there was low concern that the conduct or interpretation of the index test differed from the way it was likely to be used in clinical practice.</p>	
Domain 3: Reference standard						
Target condition and reference standard(s)	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index tests?	-	Could the reference standard, its conduct, or its interpretation have introduced bias?	Are there concerns that the target condition as defined by the reference standard did not match the review question?	
	<p>Yes: all participants underwent the acceptable reference standard.</p>	<p>Yes: reference standard results were interpreted without knowledge of the results of the index test.</p> <p>No: reference standard results were interpreted with</p>	-	<p>Low risk: 'yes' for all signalling questions.</p> <p>High risk: 'no' or 'unclear' for at least 1 of</p>	<p>Low concern: participants underwent endoscopic or surgical exploration for common bile duct stone.</p> <p>High concern: no participants underwent endoscop-</p>	

Table 1. Application of the QUADAS-2 tool for assessing methodological quality of included studies (Continued)

	<p>No: if all participants did not undergo an acceptable reference standard. Such studies were excluded from the review.</p> <p>Unclear: if the reference standard that the participants underwent was not stated. Such studies were excluded from the review.</p>	<p>the knowledge of the results of the index test.</p> <p>Unclear: this was not clear from the report.</p>	<p>the 2 signalling questions.</p>	<p>ic or surgical exploration for common bile duct stone.</p>	
Domain 4: Flow and timing					
Flow and timing	<p>Was there an appropriate interval between index test and reference standard?</p>	<p>Did all participants receive the same reference standard?</p>	<p>Were all participants included in the analysis?</p>	<p>Could the participant flow have introduced bias?</p>	-
	<p>Yes: the interval between index test and reference standard was shorter \leq 4 weeks (arbitrary choice).</p> <p>No: the interval between index test and reference standard was $>$ 4 weeks.</p> <p>Unclear: this was not clear from the report.</p>	<p>Yes: all participants underwent endoscopic or surgical exploration for common bile duct stone irrespective of the index test results.</p> <p>No: participants underwent endoscopic or surgical exploration if the index test results were positive and underwent clinical follow-up for at least 6 months if the index test results were negative.</p> <p>Unclear: this was not clear from the report. Such studies were excluded.</p>	<p>Yes: all participants meeting the selection criteria (selected participants) were included in the analysis, or data on all the selected participants were available so that a 2 x 2 table including all selected participants could be constructed.</p> <p>No: not all participants meeting the selection criteria were included in the analysis or the 2 x 2 table could not be constructed using data on all selected participants.</p> <p>Unclear: this was not clear from the report.</p>	<p>Low risk: 'yes' for all signalling questions.</p> <p>High risk: 'no' or 'unclear' for at least 1 signalling question.</p>	-

APPENDICES

Appendix 1. Search strategies

Database	Period of Search	Search Strategy
MEDLINE (PubMed)	1946 to September 2012.	(((bile duct[tiab] or biliary[tiab] OR CBD[tiab]) AND (stone[tiab] OR stones[tiab] OR calculus[tiab] OR calculi[tiab])) OR choledocholithiasis[tiab] OR cholelithiasis[tiab] OR "Choledocholithiasis"[Mesh] OR "Common Bile Duct Calculi "[MESH] OR "Cholelithiasis "[MESH]) AND (CT[tiab] OR tomodensitometry[tiab] OR MRI[tiab] OR NMRI[tiab] OR zeugmatogra*[tiab] OR ((computed[tiab] OR computerised[tiab] OR computerized[tiab] OR magneti*[tiab] OR MR[tiab] OR NMR[tiab] OR proton[tiab]) AND (tomogra*[tiab] OR scan[tiab] OR scans[tiab] OR imaging[tiab] OR cholangiogra*[tiab])) OR "Tomography, X-Ray Computed"[Mesh] OR "Magnetic Resonance Imaging"[Mesh] OR echogra*[tiab] OR ultrason*[tiab] OR ultrasound[tiab] OR EUS[tiab] OR "Ultrasonography"[Mesh] OR "Endosonography"[Mesh] OR cholangiogra*[tiab] OR cholangio?pancreatogra*[tiab] OR cholangiosco*[tiab] OR choledochosco*[tiab] OR ERCP[tiab] OR MRCP[tiab] OR "Cholangiography"[Mesh] OR "Cholangiopancreatography, Magnetic Resonance"[Mesh] OR liver function test[tiab] OR liver function tests[tiab] OR "Liver Function Tests"[Mesh])
EMBASE (OvidSP)	1947 to September 2012.	<ol style="list-style-type: none"> 1. (((bile duct or biliary or CBD) adj5 (stone or stones or calculus or calculi)) or choledocholithiasis or cholelithiasis).tw. 2. exp common bile duct stone/ or exp bile duct stone/ or exp cholelithiasis/ 3. 1 or 2 4. (CT or tomodensitometry or MRI or NMRI or zeugmatogra* or ((computed or computerised or computerized or magneti* or MR or NMR or proton) adj5 (tomogra* or scan or scans or imaging or cholangiogra*))).tw. 5. exp computer assisted tomography/ 6. exp nuclear magnetic resonance imaging/ 7. (echogra* or ultrason* or ultrasound or EUS).tw. 8. exp ultrasound/ 9. (cholangiogra* or cholangio?pancreatogra* or cholangiosco* or choledochosco* or ERCP or MRCP).tw. 10. exp cholangiography/ 11. (liver function test or liver function tests).tw. 12. exp liver function test/ 13. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 14. 3 and 13
Science Citation Index Expanded (ISI Web of Knowledge)	1898 to September 2012.	<p>#1 TS=(((bile duct or biliary OR CBD) AND (stone OR stones OR calculus OR calculi)) OR choledocholithiasis OR cholelithiasis)</p> <p>#2 TS=(CT OR tomodensitometry OR MRI OR NMRI OR zeugmatogra* OR ((computed OR computerised OR computerized OR magneti* OR MR OR NMR OR proton) AND (tomogra* OR scan OR scans OR imaging OR cholangiogra*)))</p> <p>#3 TS=(echogra* OR ultrason* OR ultrasound OR EUS)</p> <p>#4 TS=(cholangiogra* OR cholangio?pancreatogra* OR cholangiosco* OR choledochosco* OR ERCP OR MRCP)</p> <p>#5 TS=(liver function test OR liver function tests)</p> <p>#6 #5 OR #4 OR #3 OR #2</p> <p>#7 #1 AND #6</p>

(Continued)

BIOSIS (ISI Web of Knowledge)	1969 to September 2012.	<p>#1 TS=((bile duct or biliary OR CBD) AND (stone OR stones OR calculus OR calculi)) OR choledocholithiasis OR cholelithiasis</p> <p>#2 TS=(CT OR tomodensitometry OR MRI OR NMRI OR zeugmatogra* OR ((computed OR computerised OR computerized OR magneti* OR MR OR NMR OR proton) AND (tomogra* OR scan OR scans OR imaging OR cholangiogra*)))</p> <p>#3 TS=(echogra* OR ultrason* OR ultrasound OR EUS)</p> <p>#4 TS=(cholangiogra* OR cholangio?pancreatogra* OR cholangiosco* OR choledochosco* OR ERCP OR MRCP)</p> <p>#5 TS=(liver function test OR liver function tests)</p> <p>#6 #5 OR #4 OR #3 OR #2</p> <p>#7 #1 AND #6</p>
clinicaltrials.gov/	September 2012.	(bile duct) OR CBD OR choledocholithiasis OR cholelithiasis
Database of Abstracts of Reviews of Effects (DARE) and Health Technology Assessment (HTA) in <i>The Cochrane Library</i> (Wiley)	September 2012.	<p>#1 (((bile duct or biliary or CBD) NEAR/5 (stone OR stones OR calculus OR calculi)) OR choledocholithiasis OR cholelithiasis):ti,ab,kw</p> <p>#2 MeSH descriptor Choledocholithiasis explode all trees</p> <p>#3 (#1 OR #2)</p> <p>#4 (CT OR tomodensitometry OR MRI OR NMRI OR zeugmatogra* OR ((computed OR computerised OR computerized OR magneti* OR MR OR NMR OR proton) NEAR/5 (tomogra* OR scan OR scans OR imaging OR cholangiogra*))) :ti,ab,kw</p> <p>#5 MeSH descriptor Tomography, X-Ray Computed explode all trees</p> <p>#6 MeSH descriptor Magnetic Resonance Imaging explode all trees</p> <p>#7 (echogra* OR ultrason* OR ultrasound OR EUS):ti,ab,kw</p> <p>#8 MeSH descriptor Ultrasonography explode all trees</p> <p>#9 MeSH descriptor Endosonography explode all trees</p> <p>#10 (cholangiogra* OR cholangio?pancreatogra* OR cholangiosco* OR choledochosco* OR ERCP OR MRCP):ti,ab,kw</p> <p>#11 MeSH descriptor Cholangiography explode all trees</p> <p>#12 MeSH descriptor Cholangiopancreatography, Magnetic Resonance explode all trees</p> <p>#13 (liver function test OR liver function tests):ti,ab,kw</p> <p>#14 MeSH descriptor Liver Function Tests explode all trees</p> <p>#15 (#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14)</p> <p>#16 (#3 AND #15)</p>
Medion (www.medion-database.nl/)	September 2012.	We conducted four separate searches of the abstract using the terms: bile duct CBD choledocholithiasis cholelithiasis

(Continued)

ARIF (www.birmingham.ac.uk/research/activities/mds/projects/HaPS/PHEB/ARIF/databases/index.aspx) September 2012. (bile duct) OR CBD OR choledocholithiasis OR cholelithiasis

CONTRIBUTIONS OF AUTHORS

KG extracted the data and wrote the review.
VG, GP, and DH evaluated studies for inclusion.
VG extracted data from included studies.
YT and KG analysed the data.
YT contributed to the draft of the review.
DS and BRD critically commented on the review.

DECLARATIONS OF INTEREST

None.

SOURCES OF SUPPORT

Internal sources

- University College London, UK.

External sources

- University of Rijeka - Medical Faculty, Croatia.
TransMedRi project, EU FP7 REGPOT-2010-5 programme (grant agreement No. 256686)
- National Institute of Health Research, UK.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. We used the statistical package Stata instead of SAS to fit the bivariate model.
2. We performed one main analysis and performed no sensitivity analyses. One study reported the exclusion of two participants but their reference standard results were not available.
3. Author order changed: Kurinchi Selvan Gurusamy, Vanja Giljaca, David Higgle, Goran Poropat, Davor Stimac, Brian R Davidson.

NOTES

This review is based on a common protocol which needed to be split in to three reviews([Giljaca 2013](#)).

INDEX TERMS

Medical Subject Headings (MeSH)

*Liver Function Tests; Alkaline Phosphatase [blood]; Bilirubin [blood]; Biomarkers [blood]; Choledocholithiasis [*diagnosis] [*diagnostic imaging]; Ultrasonography

MeSH check words

Humans