Technoleg Iechyd Cymru
Health Technology Wales

Evidence Appraisal Report

Point-of-care ultrasound to diagnose gallstone disease

1. Purpose of the evidence appraisal report

This report aims to identify and summarise evidence that addresses the following question: Compared to other diagnostic techniques, what is the effectiveness of using hand-held or portable ultrasound to aid diagnosis in people with suspected gallstones?

Evidence Appraisal Reports are based on rapid systematic literature searches, with the aim of published evidence identifying the best clinical and economic evidence on health technologies. Researchers critically evaluate this evidence. The draft Evidence Appraisal Report is reviewed by experts and by Health Technology Wales (HTW) multidisciplinary advisory groups before publication.

2. Health problem

Gallstone disease, also referred to as cholelithiasis, occurs when hard fatty or mineral deposits (gallstones) form in the gallbladder (NICE 2014). Approximately 15% of the adult population are thought to have gallstone disease, and most of these experience no symptoms (NICE 2014). However, approximately 2% to 4% become symptomatic (NICE 2019), where stones irritate the gallbladder or block part of the biliary system and cause abdominal pain, inflammation and infection (NICE 2014). This can lead to more serious and potentially life-threatening conditions, such as cholecystitis, cholangitis, pancreatitis and jaundice. Based on the prevalence estimates above and Office of National Statistics population estimates, HTW estimates that approximately 245,035-367,553 adults aged over 20 in Wales have gallstones and 4,901-14,702 would become symptomatic annually (ONS 2020).

Diagnosing gallstones can be difficult using medical history and physical examination alone. Therefore, people presenting with suspected gallstones are usually referred for radiology-performed ultrasound (Ross 2011). This is usually performed in a radiology department by an accredited technician, and results are interpreted by a radiologist.

Effective point-of-care ultrasound could reduce the need for a radiology referral and therefore shorten time to diagnosis and treatment of gallstone disease.

3. Health technology

Point-of-care ultrasound (POCUS) covers any type of ultrasound that is performed at or near initial point of patient care. For the ultrasound diagnosis of gallstone disease, this can include the Gastroenterology Outpatient Clinic, A&E, GP practices and other outpatient or community settings. The ultrasound can be performed by different healthcare professionals.

Types of ultrasound devices used at point of care range from standard sized ultrasound machines (like those used in radiology departments), to smaller, more portable machines. The European Society of Radiology (ESR) states that “portable [ultrasound] devices can be subdivided into three groups: laptop-associated devices, hand-carried [ultrasound], and handheld [ultrasound] devices” (European Society of Radiology 2019). This review will focus on portable ultrasound devices as defined by the ESR.

4. Evidence search methods

We searched for evidence that could be used to answer the review question: Compared to other diagnostic techniques, what is the effectiveness of using hand-held/portable ultrasound to aid diagnosis in people with suspected gallstones?

The criteria used to select evidence for the appraisal are outlined in Appendix 1. These criteria were developed following comments from the Health Technology Wales (HTW) Assessment Group and UK experts.

HTW searched for evidence in Medline, Embase and the Cochrane Library, as well as relevant websites, including those for ongoing clinical trials. Date of last search was 5 January 2021. The search strategy is available upon request.

Appendix 2 summarises the selection of articles for inclusion in the review. Where available, we also included any evidence that addressed organisational or patient issues.

HTW undertook an additional literature search to target evidence on patient and/or carer issues relating to portable ultrasound for gallstone disease. However, we did not identify any additional evidence relating to portable point of care ultrasound.

5. Guidelines and Guidance

NICE clinical guideline 188 on “Gallstone disease: diagnosis and management” states to “offer liver function tests and ultrasound to people with suspected gallstone disease, and to people with abdominal or gastrointestinal symptoms that have been unresponsive to previous management” (NICE 2014). The guidelines do not specify on the type of ultrasound device.

6. Clinical effectiveness

Four systematic reviews were identified as potentially relevant (Carroll et al. 2013, Jain et al. 2017, Ross et al. 2011, Rykkje et al. 2019), but all four were only deemed partially relevant as they included studies that did not fit within the scope of this review. Three reviews included any ultrasound used at point-of-care, which included larger ultrasound machines (Carroll et al. 2013, Jain et al. 2017, Ross et al. 2011). The fourth from Rykkje et al. (2019) aimed to identify studies on the use of hand-held ultrasound devices for abdominal and pleural conditions. The citations of these reviews were checked and any relevant studies were included alongside the independently identified primary studies.

In total, we identified eight primary studies evaluating portable ultrasound for gallstones and/or cholecystitis. On further scrutiny, one study was excluded as the setting was not point of care, with the portable ultrasound and formal ultrasound performed by the same specialists (Stock et al. 2015). Two papers reported evidence for the same cohort; in this case the most recent publication included more detail on relevant outcomes, and the older paper was subsequently excluded (Gustafsson et al. 2016).

Characteristics for the remaining six studies are detailed in Table 1 (Gaszynski et al. 2019, Shekarchi et al. 2018, Gustafsson et al. 2018, Del Medico et al. 2018, Kell et al. 2002, Ismaeel et al. 2010, Colli et al. 2015).

The studies varied in design. Four studies evaluated larger ‘laptop-sized’ portable devices, whereas two studies evaluated a handheld device (Colli et al. 2015, Del Medico et al. 2018). Six studies reported diagnostic accuracy outcomes for gallstones or cholelithiasis specifically. Two studies reported diagnostic outcomes with POCUS compared to another test; this was either standard formal ultrasound (Gustafsson et al. 2018) or portable ultrasound performed by an expert operator (Del Medico et al. 2018). The reference standard was final diagnosis, formal ultrasound or radiology report. The professional performing the index text varied and included general practitioners, surgical trainees, surgeons, emergency physicians, and medical residents; all studies included a period of training. Due to the different designs of the studies, we judged that it would not be appropriate to pool outcomes.

Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) results are presented in Table 2. Additional outcomes are included in Appendix 3, Table 1.

6.1 Portable POCUS (all types)

Overall, for the diagnosis of gallstones, sensitivity of portable POCUS ranged from 75.0% to 100.0% and specificity from 84.4% to 100%. Two studies included comparative data in evaluating gallstones. In Gustafsson et al. (2018), both portable POCUS and formal ultrasound were evaluated against final diagnosis. Sensitivity was 87.1% with portable POCUS and 97.3% with formal ultrasound; specificity was 96.0% versus 98.9%. The second study evaluated a handheld device (Del Medico et al. 2018), and is discussed in more detail in Section 6.2.

Only one study reported on time to scan, defined as time from study entry to the scan being performed (Kell et al. 2002). The time to scan was significantly shorter for POCUS compared to radiology-performed ultrasound (3.1 hours versus 12 hours, $p < 0.05$); authors noted that patients with less severe pathology experienced the longest delay to radiology-performed ultrasound.

Other outcomes of interest (change in patient management, quality of life) were not reported in any of the studies identified.

6.2 Handheld POCUS

We identified two studies evaluating handheld POCUS (Colli et al. 2015, Del Medico et al. 2018). Del Medico et al. (2018) evaluated diagnosis of gallstones in two patient cohorts who received abdominal US: patients referred to the ambulatory service and patients on an internal medical ward. Patients received portable US by a non-expert operator (who received short training) and portable US performed by expert operators; standard abdominal ultrasound was used as the reference standard. Sensitivity was 75.0% by non-experts versus 93.8% by experts; specificity was 91.3% versus 100%.

In Colli et al. (2015), portable POCUS was performed in a range of settings: GP practices, hospital wards and gastroenterology outpatient clinics; the reference standard was additional testing, but the specifics of this testing were not clearly reported. Sensitivity of POCUS was 95.2% and specificity was 84.4%.

6.3 Impact on patient management

One study also evaluated the clinical impact of using hospital and non-hospital portable POCUS, and whether POCUS improved diagnostic accuracy of a physical examination for various indications (Colli et al. 2015). Participating practitioners included GPs, specialists in internal medicine and specialists in gastroenterology. Practitioners completed forms with the following:

- Demographic data
- A definition of the clinical problem (e.g. pain in the right upper quadrant)
- Results of the physical examination
- One of the ten clinical questions (i.e. asking for the diagnosis, e.g. Does the patient have gallstones?)
- The results of the portable POCUS examination (positive or negative)
- Any additional tests required
- Final diagnosis

This study looked at use of portable POCUS across multiple suspected conditions, and the results were mostly reported for all conditions; evidence for diagnosis of suspected gallstones alone was limited. For the 730 cases considered for gallstones, positive concordance between the clinical question ("Does the patient have gallstones?") and the portable POCUS occurred in 485 (66.4%) of cases; 245 (33.6%) had negative concordance. Physicians requested additional tests in 393 (57.5%) of the gallstone cohort; however, it is not clear what proportion of these were to confirm gallstones following positive physical examination and positive POCUS, and which were to determine diagnosis after discordance between physical examination and POCUS.

Table 1. Study characteristics

Reference	Study design	Population	Index test / reference standard	Relevant outcomes	Additional notes/Comments on applicability
Shekarchi et al. (2018)	Prospective observational study. Multicentre (n = 2), Iran 2015 to 2017	**Inclusion criteria:** Patients with right upper quadrant or epigastric pain suspected to be acute cholecystitis **Exclusion criteria:** history of biliary disease, jaundice, and cholecystectomy as well as intubated, pregnant and cases < 18 years **Setting:** emergency department n = 342 Mean age 53.92 ± 11.18 years (range 20 to 83 years). 63.2% female : 36.8% male	**Index test:** portable POCUS (HM-70 Samsung device or an Mturbo Sonosite) POCUS was performed by either trained emergency physicians who were expert in POCUS (179/342 [52.3%]) or trained 3rd year emergency medicine residents (163/342 [47.7%]). Emergency residents had an hour-long theoretical class of ultrasound principles and knobology and 1-hour theoretical class of right upper quadrant ultrasound Afterwards, they all had two hours of hands-on practice on a standard patient. They did their first individual examination after 12 cases of direct supervision. **Reference standard:** radiology ward reports (examinations were usually by a radiology resident under observation of a radiology attending physician).	• Sensitivity and specificity (for 4 different sonographic findings of acute cholecystitis, including gallstones) • PPV and NPV (for 4 different sonographic findings of acute cholecystitis, including gallstones)	Emergency physicians are described as experts but level of expertise/experience was not further described. Unknown number of POCUS trainees. The proportion of cases of cholecystitis with gallstones was not specified. It is not clear if radiology examinations were blinded to the index test. Patients and companions were blinded to the sonography findings.
Gustafsson et al. (2018)	Prospective observational study Single centre, Sweden October 2011 to November 2012	**Inclusion criteria**: Patients with suspected biliary disease and/or suspected appendicitis. Suspected biliary disease was defined as patients presenting with right upper quadrant pain and/or tenderness in the right upper quadrant during physical examination and/or with a referral to the radiology	**Index test:** portable US by study surgeon (LOGIQ e). Study surgeons (n = 6) included final year residents (n = 5) and a specialist in surgery (n = 1). Training consisted of a 1-week ultrasound course (including some hands-on training), followed by 3 weeks of training	• Sensitivity and specificity • PPV and NPV	Only data on suspected biliary disease was extracted for this study. Of the 183 patients referred for symptoms of gallstones/cholecystitis, diagnostic results for POCUS was given for 170 patients (POCUS not performed n = 11,

Reference	Study design	Population	Index test / reference standard	Relevant outcomes	Additional notes/Comments on applicability
		department regarding gallstones and/or cholecystitis. **Setting:** Radiology department n = 183 met the criteria for suspected biliary disease (symptoms of gallstones and/or cholecystitis). Mean age 54 years (range 19 to 92 years) 56.3% female : 43.7% male	in the radiology department. **Comparator:** standard ultrasound by radiologist (Philips iU22) **Reference standard:** final diagnosis by an independent external observer, a senior consultant surgeon, based on discharge diagnosis, operation logs and pathology reports from each patient's chart. Final diagnosis showed n = 74 had gallstones and n = 21 had acute cholecystitis.		reference data missing n = 2). Standard ultrasound results were given for 165 patients (standard ultrasound not performed n = 18). 74 patients were shown to have gallstones and 21 had acute cholecystitis. The proportion of patients who had gallstones as part of cholecystitis was not clear. Surgeons and radiologists were blinded to each other's results.
Del Medico et al. (2018)	Prospective observational study Single centre, Italy January 2016 to November 2016	**Inclusion criteria:** Patients referred/admitted to undergo abdominal ultrasound; the ability to give informed consent; a written request accounting for signs or symptoms of gallstone diseases (right hypochondrium/epigastric pain, cholestasis, dyspepsia, increased levels of transaminases, suspected gallstone disease). **Exclusion criterion:** previous cholecystectomy. **Setting:** ambulatory service (outpatients) and internal medicine ward (inpatients) n = 146	**Index test:** 'hand-held' portable US (VScan, GE Healthcare) by non-expert operators. The operators were four randomly selected internal medicine residents, attending first and second year with no prior ultrasound training. Training comprised a 2-h theoretical lesson, followed by performing 30 abdominal examinations focused on gallbladder under expert operator supervision. The internal medicine residents had no prior training in ultrasound before participating in this study. **Comparator:** 'hand-held' portable US (VScan, GE	• Sensitivity and specificity • Positive likelihood ratio & negative likelihood ratio	Patients underwent three different and sequential abdominal ultrasound examinations by three different operators, all blinded to the others and to medical history and clinical data. Two patients were excluded because the reference standard test (standard abdominal ultrasound) failed to determine the presence of cholelithiasis. Four further cases were considered impossible to determine when evaluated with portable POCUS —one by both the expert and non-expert operators, the other three cases only by the non-experts

Reference	Study design	Population	Index test / reference standard	Relevant outcomes	Additional notes/Comments on applicability
		Median age 68.0 years (range 50.5 to 77 years) 62% female : 38% male	Healthcare) by an expert operator. Expert operators were certificated by Italian national ultrasound society. **Reference standard:** standard abdominal US		—and were included as mistakes in the intention to diagnose analysis. Cholelithiasis was found in 64 cases, with a prevalence of 44.4%
Colli et al. (2015)	Prospective cohort impact study (observational / descriptive) Multicentre, Italy	**Inclusion criteria:** Patients undergoing physical examination. **Setting:** Medical wards, gastroenterology outpatient clinic and local health authority agencies responsible for general practitioners. n = 1,962 records were analysed (n = 730 were for gallstones) Median age 71 years (IQR 56 to 80 years) 52% males; 48% females	**Index test:** portable POCUS (Vscan, GE Healthcare) used by physicians without any direct experience with ultrasound: GPs (n = 90), specialists in internal medicine (n = 30) and specialists in gastroenterology (n = 15). They attended a short training course conducted by an expert in diagnostic ultrasonography. The course included a preliminary frontal lesson explaining the general technical basis of US examinations (45 minutes), the collection of pertinent images and focused examinations of patients (120 minutes), and a subsequent one-week attendance at the referral hospital with training on patients. Physicians completed between 5 and 25 examinations. **Reference standard:** final diagnosis / further imaging testing or monitoring as appropriate	• Proportion of cases where additional testing was requested after the POCUS. • Diagnostic accuracy • Sensitivity and specificity • Positive and negative likelihood ratios • Concordance between clinical hypothesis (question) and POCUS examination.	Physicians completed a sheet including demographic data, a definition of the clinical problem (e.g. pain in the right upper quadrant, abdominal distension, dyspnoea, etc.), the results of the physical examination, one of the ten clinical questions, the results of the PUD examination (positive/negative), any additional tests required and their results, and the final diagnosis. Population inclusion/exclusion criteria not clearly defined. Of those questioned for gallstones (n = 730), 110 were from the hospital ward, 291 were from general practice and 329 were from gastroenterology outpatient clinic.

Reference	Study design	Population	Index test / reference standard	Relevant outcomes	Additional notes/Comments on applicability
Ismaeel et al. (2010)	Prospective observational study Single centre, UK Enrolment period not reported	**Inclusion criteria:** Patients where abdominal ultrasound was requested. **Setting:** acute medicine unit. n = 43, (n = 29 were evaluated for gallbladder) Mean age 58 years (range 20 to 93 years) 34 males ; 9 females Three patients were excluded as they did not undergo a subsequent ultrasound by the RAD.	**Index test:** portable POCUS (Sonosite MicroMaxx) by an acute medical trainee at the acute medicine unit. The acute medical trainee had enrolled as a part time student in a local Postgraduate Certificate in Medical Ultrasound (PGCMUS) university course. At the start of the study the acute medical trainee had observed 62 and performed 57 supervised ultrasound examinations. In total the acute medical trainee had attended 60 hours of practical ultrasound sessions and 92 hours of university lectures prior to commencing the study. **Reference standard:** standard abdominal ultrasound (Toshiba) at the radiology department. Nine consultant radiologists, two supervised trainee radiologists and five sonographers performed and reported the radiology department ultrasound scans. Their range of experience was varied from 2 to 30 years of ultrasound practice.	• Sensitivity & specificity • PPV and NPV • Agreement between tests.	Each participant underwent portable POCUS before the reference standard. Unclear if reference test was blinded to the results of portable POCUS. AU the images (100%) for each ultrasound examination by the acute medical trainee and the radiology department were classified as being of diagnostic value.
Kell et al. (2002)	Prospective observational study Single centre, Ireland Enrolment was over	**Inclusion criteria:** Patients referred for emergency surgical review with symptoms or signs of suspected hepatobiliary pathology. Symptoms were either suggestive of biliary colic or right upper quadrant pain	**Index test:** surgeon-performed ultrasonography (portable Sonosite, specific device not reported). Two surgical trainees were trained in ultrasonography.	• Sensitivity and specificity • PPV and NPV • Time to scan (time from study entry until	Diagnostic accuracy outcomes for the comparator were not clearly reported and could not be extracted. Patient flow for index / comparator not clear. Unclear

Reference	Study design	Population	Index test / reference standard	Relevant outcomes	Additional notes/Comments on applicability
	a 6 month period (date NR)	with clinical signs to support the diagnosis of hepatobiliary pathology. **Setting:** not clear. Patients referred for surgical review n = 53 Mean age 48 years (range 14 to 86 years) Male : female ratio of 1 : 2.6	Training involved a 5-day course comprising 25 h of American Medical Association Category I credit. **Comparator:** Radiology-performed ultrasound. **Reference standard:** final diagnosis (following operative treatment)	ultrasound is performed).	whether results were blinded. Diagnostic outcomes of interest for the comparator (sensitivity, specificity, PPV, NPV) was not reported.

NPV: negative predictive value; NR: not reported; PPV: positive predictive value; POCUS: point of care ultrasound

Table 2. Sensitivity and specificity, PPV NPV

Study	Index test/reference standard	Patients	Sensitivity	Specificity	PPV	NPV	Notes
Gustafsson et al. (2018)	Index test: POCUS Ref standard: final diagnosis at discharge	n = 170	87.1% (95% CI, 77.3–93.1%)	96.0% (95% CI, 90.1–98.4%)	93.8% (95% CI NR)	91.4% (95% CI NR)	
	Index test: standard US Ref standard: final diagnosis at discharge	n = 165	97.3% (95% CI, 90.6–99.3%)	98.9% (95% CI, 94.1–99.8%)	98.6% (95% CI NR)	97.8% (95% CI NR)	
Del Medico et al. (2018)	Index test: hand-held POCUS by a non-expert operator Ref standard: standard US	n = 144	75.0% (95% CI, 62.6–85.0%)	91.3% (95% CI, 82.8–96.4%)	87.3% (95% CI NR)	82.0% (95% CI NR)	
	Index test: hand-held POCUS by an expert operator Ref standard: standard US	n = 144	93.8% (95% CI, 84.8– 98.3%)	100% (95% CI, 95.5–100%)	100% (95% CI NR)	95.2% (95% CI NR)	
Shekarchi et al. (2018)	Index test: POCUS Ref standard: radiology report	n = 342	92.7 (95% CI 82.3-97.3%)	97.1 (95% CI 94.1-98.6%)	88.8 (95% CI 78.7-94.7%)	98.1 (95% CI 95.5-99.3%)	
Colli et al. (2015)	Index test: POCUS Reference standard: further imaging tests	n = 393	95.2% (95% CI NR)	84.4% (95% CI NR)	97.3% (95% CI NR)	75.4% (95% CI NR)	Reference standard not clearly defined.
Ismaeel et al. (2010)	Index test: POCUS Ref standard: standard US	n = 29	100% (95% CI NR)	94% (95% CI NR)	90% (95% CI NR)	100% (95% CI NR)	
Kell et al. (2002)	Index test: POCUS Ref standard: final diagnosis at discharge	n = 53	95.2% (95% CI NR)	100.0% (95% CI NR)	100.0% (95% CI NR)	84.6% (95% CI NR)	

CI: confidence interval; NPV: negative predictive value; NR: not reported; POCUS: point of care ultrasound; PPV: positive predictive value; US: ultrasound

6.4 Ongoing trials

HTW literature searches did not identify any ongoing trials evaluating portable POCUS for the diagnosis of gallstone disease.

7. Cost effectiveness

7.1 Health economic literature review

The titles and abstracts of 1,149 records identified in the search for this research question were screened and seven records were deemed potentially relevant. The full texts of these studies were reviewed against the inclusion/exclusion criteria. All seven studies were excluded from the review.

One study from the US perspective considered people with anterior gallbladder wall thickening or oedema, peri-cholecystic fluid, or gallbladder hydrops, all with or without the presence of gallstones confirmed by bedside ultrasound in the emergency department (Young et al. 2010). The study was excluded as it did not specify the type of device used, such as whether the device was hand-held or portable. One study was excluded as it is a review and 20 years old (Rodney & Pean 2000). One study was excluded as it is not an economic evaluation (Colucciello 2019).

Four of the studies were excluded as they are conference abstracts so could not be critically appraised (DSouza et al. 2010, Bansal et al. 2014, Wang 2017, Rusiecki et al. 2020). These are briefly summarised here. The most recently published conference abstract described a prospective study from the Canadian hospital perspective (Rusiecki et al. 2020). Rusiecki et al. (2020) compared the costs of patient investigation plans of emergency physicians using POCUS with the plans of emergency physicians who did not use POCUS as part of their assessment. The study considered laboratory and imaging costs only. It did not specify the type of device used for POCUS. The study considered a broad population, including for example people with abdominal pain, chest pain, and trauma. The study found overall similar health care costs in the POCUS and non-POCUS groups. When a subgroup analysis was conducted, the study found significantly lower costs in people presenting with 'flank pain' investigated with POCUS ($43.64 vs $248.82, p = 0.01).

The remaining three excluded conference abstracts were from the US perspective (Wang 2017, Bansal et al. 2014, DSouza et al. 2010). Wang (2017) retrospectively compared the use of POCUS with radiology department ultrasound, or both in people presenting to the emergency department with acute uncomplicated cholecystitis who received a cholecystectomy. The study did not specify the type of device used for POCUS. Billing charges were $31,171 in those diagnosed with radiology department ultrasound and $31,456 in people diagnosed using both methods. A health economic model by DSouza et al. (2010) compared testing with and without the use of a pocked-sized ultrasound device in people with symptomatic suspected gallstones. The study found that pocket-sized ultrasound could potentially save $141 per person aged under 65, and $165 per person in those aged over 65.

The remaining conference abstract (Bansal et al. 2014) did not specify the type of ultrasonography used.

7.2 Original health economic modelling

An original health economic model addressing the cost effectiveness of hand-held ultrasound in addition to physical examination, compared with the standard of care (physical examination with subsequent radiology-performed ultrasound) was constructed for this report. A decision tree analysis was developed which compared the two strategies.

The population considered was 'people with suspected symptomatic gallstones presenting to the emergency department'. The analysis took the perspective of the UK NHS and personal social services (PSS). A time horizon of one year was considered, reflecting the maximum period over which outcomes are likely to differ between the strategies. Discounting of future costs and benefits was not considered due to the short time horizon.

The results of the base case analysis are presented in Table 3, which shows the total and incremental costs and quality-adjusted life years (QALYs) over the time horizon (presented on a per patient basis) as well as the incremental cost effectiveness ration (ICER). Hand-held/portable US was found to be cheaper, though less effective in comparison with radiologist-performed ultrasound. The ICER of £58,631 per QALY indicates that POCUS is cost effective at the £20,000 per QALY threshold. (N.B. In the south west quadrant of the cost effectiveness plane, an ICER above £20,000 per QALY is cost effective, as it indicates that the money saved per QALY lost could be better spent elsewhere).

The result was found to be driven primarily by an assumption made in the base case of the model that there is a reduced waiting time for scanning in POCUS compared with radiology-performed ultrasound (RADUS), which was based on an outcome reported by a single study (Kell et al. (2002), Section 6.1). In a scenario analysis with no difference in waiting time for scanning, POCUS was dominated (more costly and less effective than RADUS). In a threshold analysis, it was found that the difference in waiting time for scanning needs to be 4.2 hours or greater for POCUS to be cost effective.

In two other scenario analyses, which applied the diagnostic accuracy data from the Del Medico et al. (2018) and Colli et al. (2015) studies, POCUS was found to be not cost effective (ICER: £15,441 per QALY) and dominated, respectively. POCUS was found to be cost effective in the remaining scenario analyses.

Probabilistic sensitivity analysis (PSA) was conducted to assess the combined parameter uncertainty in the model (Table 4). In this analysis, the mean values that were utilised in the base case were replaced with values drawn from distributions around the mean values and the model is run 10,000 times. At a threshold of £20,000 per QALY, POCUS was found to have a 49% probability of being cost effective, while there was a 51% probability that RADUS was cost effective.

See Appendix 4 for details.

Table 3. Base case results

Treatment strategy	Cost		QALYs		ICER (cost per QALY)
	Total	Incremental	Total	Incremental	
Radiologist-performed ultrasound	£1,617	-	0.905	-	-
Hand-held/portable ultrasound	£1,470	Saves £147	0.903	-0.0025	£58,631 per QALY

ICER: incremental cost effectiveness ratio; QALYs: quality-adjusted life years
N.B. In the South West quadrant of the cost-effectiveness plane, an ICER above £20,000 per QALY indicates that hand-held/portable ultrasound is more cost effective, as more money is saved per QALY lost.

Table 4. Probabilistic sensitivity analysis results

Treatment strategy	Cost		QALYs		ICER (cost per QALY)
	Total	Incremental	Total	Incremental	
Radiologist-performed ultrasound	£1,643	-	0.91	-	-
Hand-held/portable ultrasound	£1,584	Saves £59	0.90	-0.0033	£17,769 per QALY

ICER: incremental cost effectiveness ratio; QALYs: quality-adjusted life years
N.B. In the South West quadrant of the cost-effectiveness plane, an ICER above £20,000 per QALY indicates that hand-held/portable ultrasound is more cost effective, as more money is saved per QALY lost.

8. Organisational Issues

Expert comments during HTW consultation noted that training and experience of the operators is a key consideration of POCUS effectiveness. Of the six studies identified in this review, operator experience and expertise varied. Five studies evaluated POCUS used by inexperienced operators who underwent a period of training, and the sixth study had a mix of POCUS experts or inexperienced operators (Shekarchi et al. 2018). All studies included a period of training, but again varied in length and content (including theory versus practical elements). Details of operators and training are described in Table 1.

One study compared handheld POCUS performed by non-experts with handheld POCUS performed by certified expert operators (Del Medico et al. 2018). Both sensitivity and specificity were lower in POCUS performed by 'non-experts' than expert-performed POCUS (Table 2).

9. Patient issues

HTW looked for evidence on patient issues in the main searches as well as through a more targeted search for patient and/or carer issues. We did not identify any additional patient evidence on the use of portable POCUS for the diagnosis of gallstone disease.

HTW did identify one study (Lindelius et al. 2009) that reported the impact of POCUS on patient satisfaction, although this was not using a portable device (Hawk 2102, B-K Medical). People presenting at the emergency department with abdominal pain received standard examination and were then randomised to receive either surgeon-performed POCUS (n = 392), or no POCUS (n = 391). Patient-reported satisfaction used a visual analogue scale from zero to ten, with zero representing lowest level of satisfaction and ten representing the highest. When leaving the emergency department, initial patient satisfaction was higher in the group that received POCUS (median 9.5 versus 9.2, p = 0.005). However, at six-week follow-up there was no difference between the two groups (median 8.0 versus 8.0, p = 0.958).

10. Conclusions

The aim of this review was to address the following question: Compared to other diagnostic techniques, what is the effectiveness of using hand-held or portable ultrasound to aid diagnosis in people with suspected gallstones?

Six unique studies were identified that evaluated the diagnostic accuracy of portable POCUS for the diagnosis of gallstones. Sensitivity ranged from 75% to 100% and specificity ranged from 84.4% to 100%. These studies varied in design, from the device used, the operators performing the POCUS, to the reference standard applied.

One study reported comparative data on portable 'laptop-sized' POCUS against formal ultrasound, with final diagnosis as the reference standard. The study showed similar specificities for both types of test, but higher sensitivities were reported with formal ultrasound.

Another study compared test accuracy for handheld POCUS carried out by non-expert or expert operators, with formal ultrasound as the reference standard. Both sensitivity and specificity were higher for tests carried out by expert operators.

We identified one study reporting on patient management; although this study reported concordance between examination and POCUS, the study did not clearly report actual impact or change in patient management. For example, it was not clear what proportion of the gallstones cohort would have been referred for unnecessary additional testing without the addition of POCUS.

No health economic studies were included in the literature review. One original health economic analysis found in the deterministic analysis that hand-held/portable US was cheaper, though less effective than radiologist-performed ultrasound. The ICER of £58,631 per QALY indicates that POCUS is cost effective at the £20,000 per QALY threshold. (N.B. In the south west quadrant of the cost effectiveness plane, an ICER above £20,000 per QALY is cost effective, as it indicates that the money saved per QALY lost could be better spent elsewhere). However, the result was not robust across all scenario analyses. For example, when a key assumption made in the base case of the model was changed so that no difference was assumed between POCUS and RADUS in waiting time for scanning, POCUS was dominated (more costly and less effective than RADUS). In a threshold analysis, it was found that the difference in waiting time for scanning needs to be 4.2 hours or greater for POCUS to be cost effective. In probabilistic sensitivity analysis, POCUS was found to have a 49% probability of being cost effective at a threshold of £20,000 per QALY gained. The average ICER in probabilistic sensitivity analysis was £17,769 per QALY gained, which is below the NICE threshold of £20,000 per QALY and indicates that POCUS is not cost effective.

11. Contributors

The HTW staff and contract researchers involved in writing this report were:

- B Coles, J Washington – literature searches
- L Elston – clinical author
- S Hughes – health economics author
- A Evans – PPI lead
- D Jarrom – clinical quality assurance
- M Prettyjohns – health economics quality assurance
- K McDermott – project management

The HTW Assessment Group advised on methodology throughout the scoping and development of the report.

A range of clinical experts from the UK provided material and commented on a draft of this report. Their views were documented and have been actioned accordingly. All contributions from reviewers were considered by HTW's Assessment Group. However, reviewers had no role in authorship or editorial control, and the views expressed are those of Health Technology Wales.

Experts who contributed to this appraisal:

- T George – Consultant Gastroenterologist, Wrexham Maelor Hospital
- K Gurusamy - Professor of Evidence-based Medicine and Surgery, University College London
- J Green - Consultant Gastroenterologist, Cardiff & Vale UHB
- M Smith - Senior Lecturer, Cardiff University
- M Khan - Consultant Gastroenterologist and Clinical Lead Wales Neuroendocrine Cancer Service, Cardiff & Vale UHB

12. References

Bansal R, Singhvi G, Shah R, et al. (2014). Tu1950 Assessment of potential cost savings in gall stone pancreatitis. Gastroenterology. 146(5): S-880. doi: http://dx.doi.org/10.1016/S0016-5085%2814%2963205-4

Carroll PJ, Gibson D, El-Faedy O, et al. (2013). Surgeon-performed ultrasound at the bedside for the detection of appendicitis and gallstones: systematic review and meta-analysis. American Journal of Surgery. 205(1): 102-8. doi: https://doi.org/10.1016/j.amjsurg.2012.02.017

CEVR. (2020). CEA Registry. Center for the Evaluation of Value and Risk in Health. Available at: http://healtheconomicsdev.tuftsmedicalcenter.org/cear2/search/search.aspx [Accessed 26 Oct 2020].

Colli A, Prati D, Fraquelli M, et al. (2015). The use of a pocket-sized ultrasound device improves physical examination: results of an in- and outpatient cohort study. PloS One. 10(3): e0122181. doi: https://doi.org/10.1371/journal.pone.0122181

Colucciello S. (2019). Assessing abdominal pain in adults: a rational, cost-effective, and evidence-based strategy. Emergency Medicine Practice. 21(6): 1-32.

Cook J, Richardson J, Street A. (1994). A cost utility analysis of treatment options for gallstone disease: methodological issues and results. Health Economics. 3(3): 157-68. doi: https://doi.org/10.1002/hec.4730030305

DCHW. (2020). Annual PEDW Data Tables 2018/2019. Patient Episode Database for Wales. Digital Health and Care Wales, NHS Wales. Available at: https://nwis.nhs.wales/information-services/welsh-data-hub/pedw-data-online/ [Accessed 6 May 2021].

Del Medico M, Altieri A, Carnevale-Maffe G, et al. (2018). Pocket-size ultrasound device in cholelithiasis: diagnostic accuracy and efficacy of short-term training. Internal and Emergency Medicine. 13(7): 1121-6. doi: https://doi.org/10.1007/s11739-018-1901-3

DSouza A, Patel P, Shah M. (2010). Economic benefit of using a pocket-sized ultrasound device for the initial diagnostic test in symptomatic patients with suspected gallstone disease: 1077. American Journal of Gastroenterology. 105(SUPPL. 1): S389.

European Society of Radiology. (2019). ESR statement on portable ultrasound devices. Insights into Imaging. 10(1): 89. doi: https://doi.org/10.1186/s13244-019-0775-x

Gaszynski R, Lim C, Chan DL, et al. (2019). Surgical ultrasonography at the bedside: a comparison of surgical trainees with trained sonographers for symptomatic cholelithiasis - a first Australian experience. ANZ Journal of Surgery. 89(5): 492-6. doi: https://doi.org/10.1111/ans.14928

Gurusamy K, Wilson E, Burroughs AK, et al. (2012). Intra-operative vs pre-operative endoscopic sphincterotomy in patients with gallbladder and common bile duct stones: cost-utility and value-of-information analysis. Applied Health Economics and Health Policy. 10(1): 15-29. doi: https://doi.org/10.2165/11594950-000000000-00000

Gustafsson C, Lindelius A, Torngren S, et al. (2018). Surgeon-performed ultrasound in diagnosing acute cholecystitis and appendicitis. World Journal of Surgery. 42(11): 3551-9. doi: https://doi.org/10.1007/s00268-018-4673-z

Gustafsson C, McNicholas A, Sonden A, et al. (2016). Accuracy of surgeon-performed ultrasound in detecting gallstones: a validation study. World Journal of Surgery. 40(7): 1688-94. doi: https://doi.org/10.1007/s00268-016-3468-3

Howard K, Lord SJ, Speer A, et al. (2006). Value of magnetic resonance cholangiopancreatography in the diagnosis of biliary abnormalities in postcholecystectomy patients: a probabilistic cost-effectiveness analysis of diagnostic strategies. International Journal of Technology Assessment in Health Care. 22(1): 109-18. doi: https://doi.org/10.1017/s0266462306050902

Ismaeel SM, Day NJ, Earnshaw D, et al. (2010). Training requirements for point of care ultrasound in acute medicine. Acute Medicine. 9(2): 87-90.

Jain A, Mehta N, Secko M, et al. (2017). History, physical examination, laboratory testing, and emergency department ultrasonography for the diagnosis of acute cholecystitis. Academic Emergency Medicine. 24(3): 281-97. doi: https://doi.org/10.1111/acem.13132

Kell MR, Aherne NJ, Coffey C, et al. (2002). Emergency surgeon-performed hepatobiliary ultrasonography. British Journal of Surgery. 89(11): 1402-4. doi: https://doi.org/10.1046/j.1365-2168.2002.02297.x

Lindelius A, Torngren S, Nilsson L, et al. (2009). Randomized clinical trial of bedside ultrasound among patients with abdominal pain in the emergency department: impact on patient satisfaction and health care consumption. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 17: 60. doi: https://doi.org/10.1186/1757-7241-17-60

Morris S, Gurusamy KS, Sheringham J, et al. (2015). Cost-effectiveness analysis of endoscopic ultrasound versus magnetic resonance cholangiopancreatography in patients with suspected common bile duct stones. PloS One. 10(3): e0121699. doi: https://doi.org/10.1371/journal.pone.0121699

NHS England. (2020). National schedule of NHS costs. 2018/19 National Cost Collection data. Available at: https://www.england.nhs.uk/national-cost-collection/#ncc1819 [Accessed 26 Oct 2020].

NHS Supply Chain. Vscan from GE Healthcare. Available at: https://media.supplychain.nhs.uk/media/documents/JAH091/Marketing/23903_JAH091%20MM.pdf [Accessed 6 Oct 2020].

NHS Supply Chain. (2020). Online catalogue. Supply Chain Coordination Limited. Available at: https://my.supplychain.nhs.uk/catalogue [Accessed 26 Oct 2020].

NICE. (2014). Gallstone disease: diagnosis and management. Clinical guideline CG188. National Institute for Health and Social Care Excellence. Available at: https://www.nice.org.uk/guidance/cg188 [Accessed 31 Oct 2020].

NICE. (2019). Gallstones. Clinical Knowledge Summary. National Institute for Health and Care Excellence. Available at: https://cks.nice.org.uk/topics/gallstones/ [Accessed 4 Nov 2020].

ONS. (2020). Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland. Office for National Statistics. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland [Accessed 4 Nov 2020].

PSSRU. (2019). Unit costs of health and social care 2019. Personal Social Services Research Unit. Available at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/ [Accessed 26 Oct 2020].

Rodney W, Pean C. (2000). Acute abdominal pain in the elderly: guide to a cost-effective work-up. Consultant. 40(1): 25-35.

Ross M, Brown M, McLaughlin K, et al. (2011). Emergency physician-performed ultrasound to diagnose cholelithiasis: a systematic review. Academic Emergency Medicine. 18(3): 227-35. doi: https://doi.org/10.1111/j.1553-2712.2011.01012.x

Rusiecki D, Douglas S, Bell C. (2020). P041: Point-of-care ultrasound utilization and monetary outcomes (POCUMON) study. Canadian Journal of Emergency Medicine. 22(S1): S79. doi: https://doi.org/10.1017/cem.2020.248

Rykkje A, Carlsen JF, Nielsen MB. (2019). Hand-held ultrasound devices compared with high-end ultrasound systems: a systematic review. Diagnostics. 9(2): 61. doi: http://dx.doi.org/10.3390/diagnostics9020061

Shekarchi B, Rafsanjani SZH, Fomani NSR, et al. (2018). Emergency department bedside ultrasonography for diagnosis of acute cholecystitis: a diagnostic accuracy study. Emergency. 6(1): e11.

StatsWales. (2020). Performance against 4 hour waiting times target by hospital. Available at: https://statswales.gov.wales/Catalogue/Health-and-Social-Care/NHS-Hospital-Waiting-Times/Accident-and-Emergency/performanceagainst4hourwaitingtimestarget-by-hospital [Accessed 26 Oct 2020].

Stock KF, Klein B, Steubl D, et al. (2015). Comparison of a pocket-size ultrasound device with a premium ultrasound machine: diagnostic value and time required in bedside ultrasound examination. Abdominal Imaging. 40(7): 2861-6. doi: http://dx.doi.org/10.1007/s00261-015-0406-z

Sutherland JM, Mok J, Liu G, et al. (2020). A cost-utility study of laparoscopic cholecystectomy for the treatment of symptomatic gallstones. Journal of Gastrointestinal Surgery. 24(6): 1314-9. doi: https://doi.org/10.1007/s11605-019-04268-z

Wang A. (2017). 667 Financial impact of cholecystitis diagnosed by emergency department point of care ultrasound (POCUS). Academic Emergency Medicine. 24(Supplement 1): S233-4. doi: http://dx.doi.org/10.1111/acem.13203

Williams E, Beckingham I, El Sayed G, et al. (2017). Updated guideline on the management of common bile duct stones (CBDS). Gut. 66(5): 765-82. doi: https://doi.org/10.1136/gutjnl-2016-312317

Wilson E, Gurusamy K, Gluud C, et al. (2010). Cost-utility and value-of-information analysis of early versus delayed laparoscopic cholecystectomy for acute cholecystitis. British Journal of Surgery. 97(2): 210-9. doi: https://doi.org/10.1002/bjs.6872

Young N, Kinsella S, Raio CC, et al. (2010). Economic impact of additional radiographic studies after registered diagnostic medical sonographer (RDMS)-certified emergency physician-performed identification of cholecystitis by ultrasound. Journal of Emergency Medicine. 38(5): 645-51. doi: https://doi.org/10.1016/j.jemermed.2008.10.016

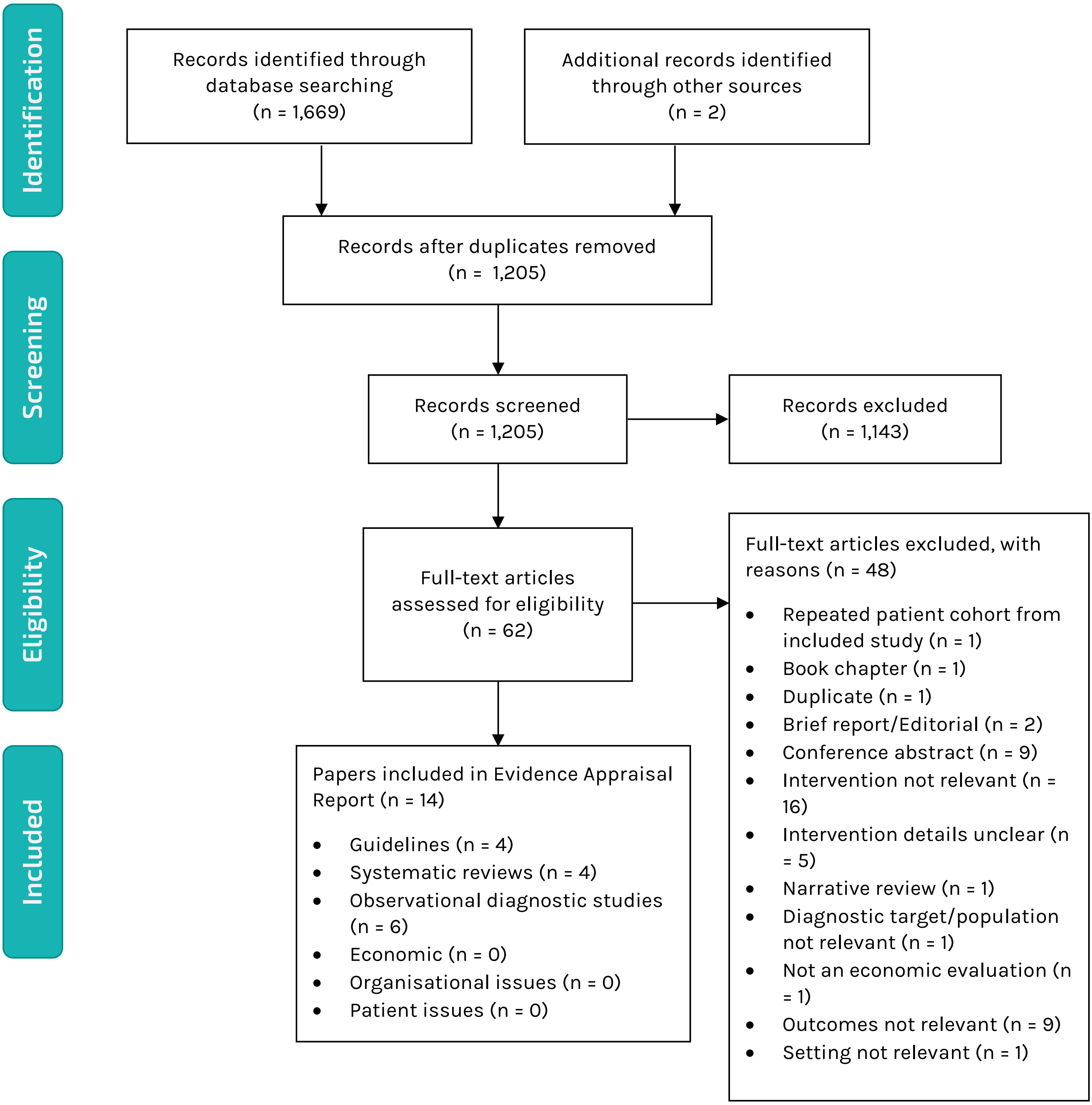
Appendix 1. PICO framework

Research Question	Compared to other diagnostic techniques, what is the effectiveness of using hand-held or portable ultrasound to aid diagnosis in people with suspected gallstones?

	Inclusion criteria	Exclusion criteria
Population	People with gastroenterological or abdominal symptoms leading to clinical suspicion of gallstones, or any other case where gallstones are suspected and diagnostic investigation is required	People with already diagnosed gallstones
Index test	Point-of-care 'portable' or hand-held ultrasound. Any setting, including, but not limited to: Gastroenterology OP clinic, A&E, Medical Admissions unit, GP Surgery	Point-of-care devices that are not 'portable', e.g. mobile ultrasound machines. The European Society of Radiology states that "Portable US devices can be subdivided into three groups: laptop-associated devices, hand-carried US, and handheld US devices"
Comparison/ Comparators	Any comparison. Standard care is physical examination initially, followed up with high-resolution ultrasound in radiology department	
Reference standard	Any suitable reference standard	
Outcome measures	• Sensitivity and specificity • Positive predictive value (PPV) and negative predictive value (NPV) • Time to diagnosis • Changes in patient management • Quality of life	
Study design	We will prioritise the following study types, in the order listed: • Systematic reviews of interventions or diagnostic accuracy • Randomised trials or studies of diagnostic accuracy that employ a suitable reference standard • Non-randomised comparative trials • Single-arm trials that report any relevant outcome We will only include evidence for "lower priority" evidence where outcomes for each condition/symptom of interest are not reported by a "higher priority" source.	

	Inclusion criteria	Exclusion criteria
	We will also search for economic evaluations or original research that can form the basis of an assessment of costs/cost comparison.	
Search limits	No limits.	
Other factors	No other factors.	

Appendix 2. PRISMA flow diagram outlining selection of papers for clinical and cost effectiveness



Appendix 3. Other clinical outcomes

Table 1. Other outcomes

Study	Comparator / Reference standard	Patients	Outcome	Notes
Time to scan				
Kell et al. (2002)	Index test: POCUS Comparator: radiology US Ref standard: final diagnosis at discharge	n = 53	POCUS: 3.1 hours; radiology US: 12.0 hours ($p < 0.05$)	Patients with less severe pathology had the longest delay to radiology-performed US. The greatest delays for radiology-performed US were in patients who either did not have pathology at scan or who were found to have biliary colic.
Diagnostic accuracy				
Gustafsson et al. (2018)	Index test: POCUS Ref standard: final diagnosis at discharge	n = 170	92.3% (95% CI 87.4–95.5%)	
	Index test: standard US Ref standard: final diagnosis at discharge	n = 165	98.2% (95% CI 94.8–99.4%)	
Inter-rater variability/agreement				
Gustafsson et al. (2018)	Index test: POCUS Comparator: standard US	n = 160	κ = 0.79 (good agreement)	
Gaszynski et al. (2019)	Index test: POCUS Comparator: accredited formal US or CT	n = 100	κ = 0.62, 95% CI 0.46–0.77)	
Shekarchi et al. (2018)	Index test: POCUS Ref standard/Comparator: radiology report	n = 342	κ = 0.884	
Ismaeel et al. (2010)	Index test: POCUS Ref standard: standard US	n = 29	κ = 0.925 (95% CI 0.782 to 1.0)	
Positive likelihood ratio and negative likelihood ratio				
Gustafsson et al. (2018)	Index test: POCUS Ref standard: final diagnosis at discharge	n = 170	PLR: 21.8 (95% CI 8.30–57.2) NLR: 0.13 (95% CI 0.07–0.25)	
	Index test: standard US Ref standard: final diagnosis at discharge	n = 165	PLR: 89.5 (95% CI 12.7–629) NLR: 0.03 (95% CI 0.01–0.11)	

Study	Comparator / Reference standard	Patients	Outcome	Notes
Shekarchi et al. (2018)	Index test: POCUS Ref standard: radiology report	n = 342	PLR: 4.30 (95% CI: 2.42 – 7.62) NLR: 0.017 (95% CI: 0.007 – 0.041)	

CI: confidence interval; NLR: negative likelihood ratio; PLR: positive likelihood ratio; POCUS: point-of-care ultrasound; US: ultrasound.

Appendix 4. Original health economic modelling

1. Background and objective

An original health economic model was constructed to assess the cost effectiveness of hand-held ultrasound in addition to physical examination, compared with the standard of care (physical examination with subsequent radiology-performed ultrasound).

The emergency department setting was selected for the health economic model. This decision was based on the settings of the studies included in this report. Gustafsson et al. (2018) considered people referred to radiology and hand-held/portable ultrasound was carried out by surgeons. Kell et al. (2002) considered people referred for surgical review, with surgical trainees carrying out hand-held/portable ultrasound. Shekarchi et al. (2018) considered the emergency department setting. The setting for Ismaeel et al. (2010) was an acute medicine unit. Colli et al. (2015) and Del Medico et al. (2018) included a mix of settings. Devices were assigned to medical wards, gastroenterology outpatient clinics and general practitioners in Colli et al. (2015). Del Medico et al. (2018) included the outpatient and inpatient settings. As neither study reported diagnostic accuracy data separately for different settings, the emergency department setting was selected for the health economic model.

2. Methods

2.1 Model structure

A decision tree model was developed using Microsoft Excel to compare the cost effectiveness of hand-held/portable ultrasound and radiology-performed ultrasound for the diagnosis of gallstones. The analysis took the perspective of the UK NHS and personal social services (PSS). A time horizon of one year was considered, reflecting the maximum period over which outcomes are likely to differ between the strategies. Discounting of future costs and benefits was not considered due to the short time horizon.

People in the model receive either hand-held POCUS or RADUS and then enter the decision tree shown in figures 1a and b. The two strategies considered in the analysis are:

1. **Hand-held POCUS**
   People presenting in the emergency department undergo standard physical examination, hand-held POCUS and liver function tests.
2. **Standard of care (radiology-performed ultrasound)**
   People presenting in the emergency department undergo standard physical examination, radiology-performed ultrasound and liver function tests.

People with positive ultrasound results (true positives and false positives) and stones in the gallbladder have either day case or elective inpatient laparoscopic cholecystectomy, or elective inpatient open cholecystectomy. Those with positive ultrasound results who also have stones in the common bile duct receive either ERCP only, ERCP and elective inpatient laparoscopic cholecystectomy or ERCP and elective inpatient open cholecystectomy. False negatives with abnormal liver function test results are assumed to re-present and undergo Magnetic Resonance Cholangiopancreatography (MRCP). After MRCP, people either have ERCP only or ERCP with laparoscopic or open cholecystectomy. If they have normal liver function tests, they are assumed to either require emergency laparoscopic cholecystectomy (and RADUS), or to re-present at a later date for RADUS and undergo either day case, elective inpatient laparoscopic cholecystectomy, or open cholecystectomy.

People with true negative ultrasound results and normal liver function test results are discharged. People with true negative ultrasound results and abnormal liver function test results receive MRCP and are then discharged.

Figure 1a

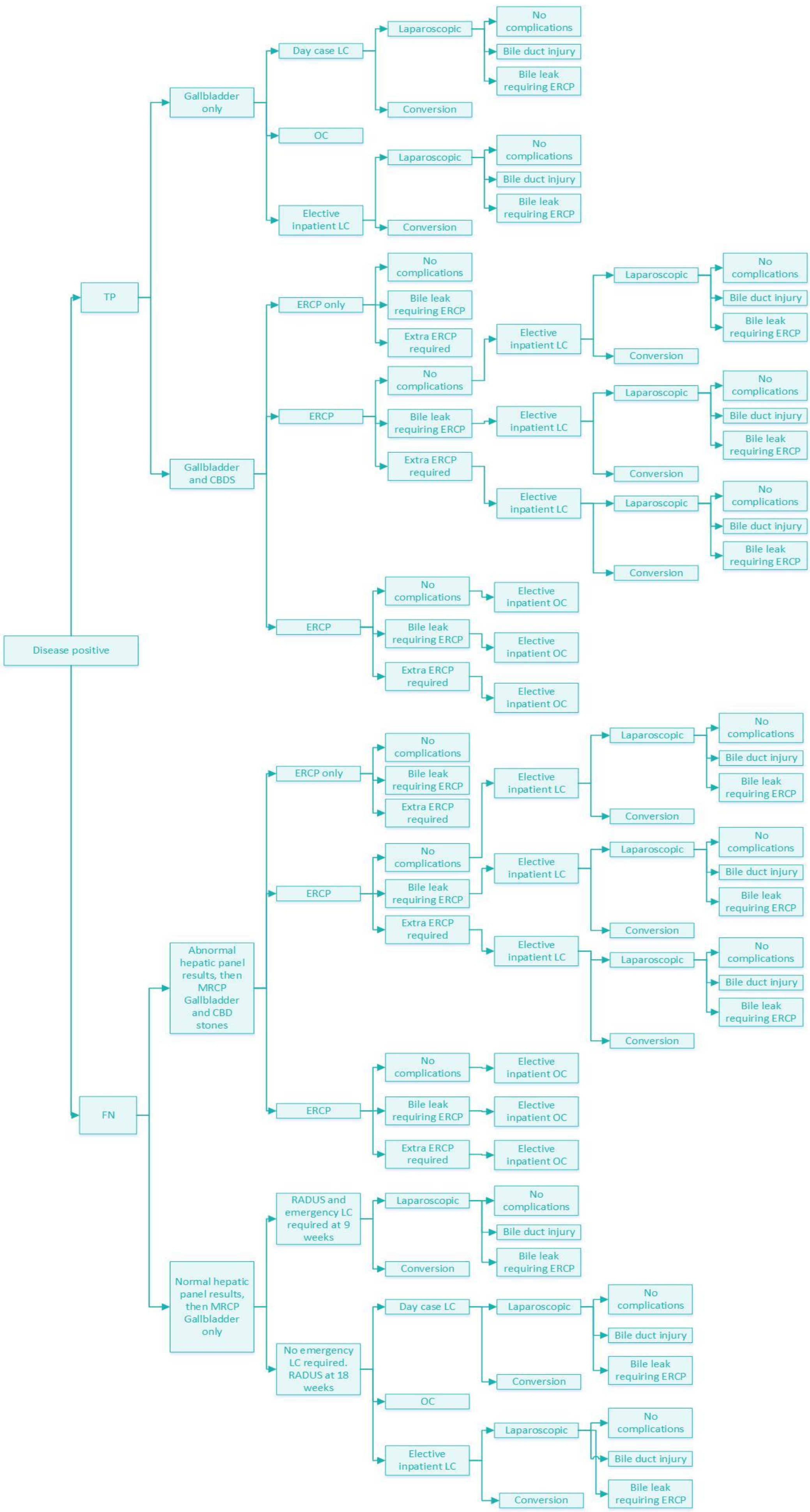

Disease positive
TP
FN
Gallbladder only
Day case LC
Laparoscopic
No complications
Bile duct injury
Bile leak requiring ERCP
Conversion
OC
Elective inpatient LC
Gallbladder and CBDS
ERCP only
ERCP
Extra ERCP required
Elective inpatient OC
Abnormal hepatic panel results, then MRCP Gallbladder and CBD stones
Normal hepatic panel results, then MRCP Gallbladder only
RADUS and emergency LC required at 9 weeks
No emergency LC required. RADUS at 18 weeks

Figure 1b

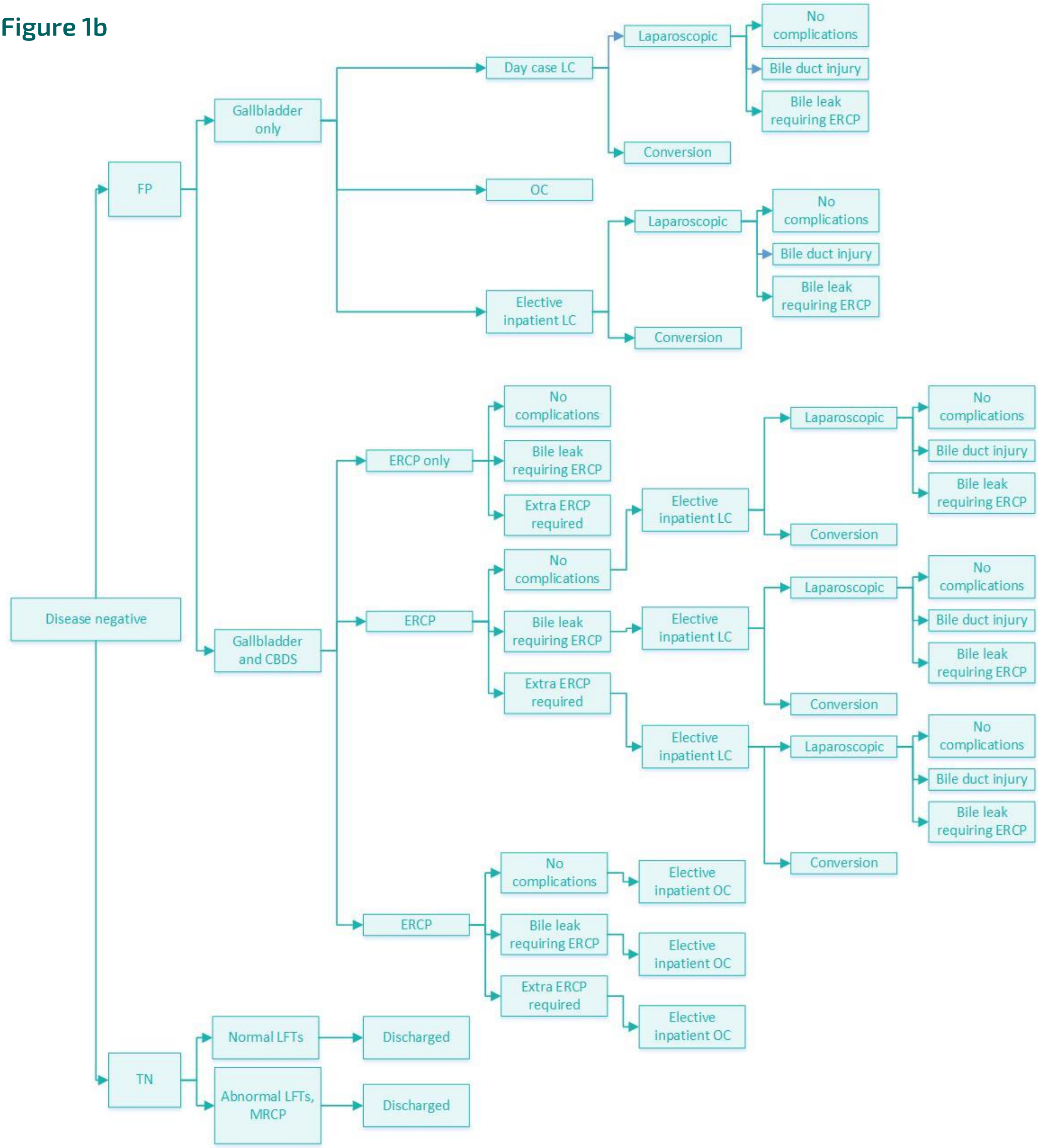


Figure 1a: Decision tree for true positives and false negatives

Figure: 1b: Decision tree for false positives and true negatives

TP: true positive; FN: false negative; FP: false positive; TN: true negative; CBD: common bile duct; ERCP: endoscopic retrograde cholangiopancreatography; LC: laparoscopic cholecystectomy; LFT: liver function test; MRCP: magnetic resonance cholangiopancreatography; OC: open cholecystectomy; POCUS: Hand-held/portable ultrasound

Assumptions

The population entering the model is assumed to have suspected **symptomatic** gallstones, to reflect the emergency department setting. As such, conservative management of gallstones is not included in the model. This is because the NICE Guideline on the diagnosis and management of gallstone disease (CG188) stipulates that laparoscopic cholecystectomy is offered to people with symptomatic gallbladder stones (NICE 2014).

It is assumed that clinical suspicion of common bile duct stones is low in this group of people with suspected symptomatic gallstones. Where clinical suspicion is low, the incidence of bile duct stones is lower than typically seen in people with symptomatic gallstones (<5% of cases as opposed to 10-20%) (Williams et al. 2017).

In accordance with NICE Guideline CG188, everyone in the model receives liver function tests. In people with stones diagnosed using ultrasound, liver function test results are assumed to be used in conjunction with ultrasound to 'flag' 100% of common bile duct stones, which can then be managed accordingly. People with false negative results on ultrasound and abnormal liver function test results are assumed to undergo MRCP, in line with NICE recommendations. This is assumed to allow a correct diagnosis to be made. People with true negative results on ultrasound and abnormal liver function test results also undergo MRCP, and are discharged. The probability of having abnormal liver function test results is assumed to be the same regardless of ultrasound result.

Management for common bile duct stones is assumed to be bile duct clearance via ERCP, either alone or with subsequent laparoscopic or open cholecystectomy.

Open cholecystectomy and conversion are assumed to be 100% effective, without further complications.

2.2 Clinical data

2.2.1 Prevalence and accuracy data

In the model base case, the prevalence of gallstones in people with suspected gallstones was estimated to be 43%, based on the prevalence reported in the study by Gustafsson et al. (2018) across both surgeon-performed ultrasound (n=70/170) and radiologist-performed ultrasound (n=73/165). This was varied in probabilistic sensitivity analysis using a beta distribution.

The Gustafsson et al. (2018) study was chosen as the base case as the study provides diagnostic accuracy data for the relevant comparator (radiologist-performed formal ultrasound). Hand-held ultrasound diagnostic accuracy data from other studies were applied in other scenarios (Table 1). In addition, a two-way sensitivity analysis was undertaken to explore the effect of varying sensitivity and specificity on results.

Table 1: Diagnostic accuracy of ultrasound

Scenario	Input	Mean	α, β for Beta distribution	Source
Radiology-performed ultrasound				
All scenarios	Sensitivity	97.3%	71, 2	Gustafsson et al. (2018)
	Specificity	98.6%	2 ,1	
Hand-held/portable ultrasound				
Base case	Sensitivity	87.1%	61, 9	Gustafsson et al. (2018)
	Specificity	96.0%	96, 4	

Scenario	Input	Mean	α, β for Beta distribution	Source
Scenario 1	Sensitivity	92.7%	-	Shekarchi et al. (2018)
	Specificity	97.1%	-	
Scenario 2	Sensitivity	95.2%	-	Kell et al. (2002)
	Specificity	100%	-	
Scenario 3	Sensitivity	100%	-	Ismaeel et al. (2010)
	Specificity	94%	-	
Scenario 4	Sensitivity	95.2%	-	Colli et al. (2015)
	Specificity	84.4%	-	
Scenario 5	Sensitivity	75%	-	Del Medico et al. (2018)
	Specificity	91.3%	-	
Abbreviations: 95% CI: 95% confidence intervals				

2.2.2 Other decision tree probabilities

The probabilities in tables 2-4 do not differ between the strategies being compared.

The probability of false negatives with normal liver function tests requiring emergency laparoscopic cholecystectomy was based on the rate of requiring laparoscopic cholecystectomy in people receiving conservative management or in the delayed laparoscopic cholecystectomy arm of the NICE CG188 health economic model (NICE 2014). Assuming a constant rate, the 9 week (range 0- 18 weeks) probability of requiring an emergency laparoscopic cholecystectomy was calculated. This corresponds to a probability of 6.11%. This probability was varied both in probabilistic sensitivity analysis and in a deterministic sensitivity analysis to the probability of ‘recurrence or non-resolution of acute cholecystitis necessitating emergency laparoscopic cholecystectomy’, cited by Gurusamy et al. (2012) as 18%. This sensitivity analyses was designed to increase the cost and quality of life implications of false positives to test the impact on the results.

Table 2: Probabilities of management strategies following diagnosis of gallstones

Input	Mean	Distribution	Components for PSA	Source/notes
Probability of common bile duct stones	5%	Uniform	Range: 5%-20%	Williams et al. (2017). Common bile duct stones are estimated to be present in 10–20% of individuals with symptomatic gallstones. Where there is no clinical suspicion of ductal stones prior to surgery the incidence is significantly lower and is typically reported to be <5%
Probability of gallstones only	95%	-	-	-
If stones in common bile duct:				
ERCP and laparoscopic cholecystectomy	85%	Dirichlet	19537	Assumption based on ratio of laparoscopic:open FCEs in NHS reference costs 18/19 (NHS England 2020)

Input	Mean	Distribution	Components for PSA	Source/notes
ERCP and open cholecystectomy	10%	Dirichlet	2249	Assumption based on ratio of laparoscopic:open FCEs in NHS reference costs 18/19 (NHS England 2020)
ERCP only	5%	Uniform	Range: 0%-5%	Clinical opinion
If stones in gallbladder only:				
Day case laparoscopic cholecystectomy	73%	Dirichlet	58001	Assumption based on ratio of FCEs in NHS reference costs 18/19 (NHS England 2020)
Elective inpatient laparoscopic cholecystectomy	24%	Dirichlet	2249	
Elective inpatient open cholecystectomy	3%	Dirichlet	19537	
Ln (rate per day) Emergency laparoscopic cholecystectomy required for false negatives with normal liver function tests	-6.91	Normal	SD: 0.971	Used model input 'laparoscopic cholecystectomy required (conservative management/delay only)' from NICE CG188 (NICE 2014)
Abbreviations: 95% CI: 95% confidence intervals; ERCP: endoscopic retrograde cholangiopancreatography; FCEs: finished consultant episodes; PSA: probabilistic sensitivity analysis; SD: standard deviation				

As it was not possible to undertake full literature searches for all model inputs, the health economic model undertaken for NICE CG188 was prioritised as a source of inputs for surgical outcomes such as conversion (Table 3) and outcomes of surgeries (Table 4)

Table 3: Probabilities of conversion from laparoscopic to open cholecystectomy

Input	Mean	Distribution	Standard deviation	Source/notes
Early (day case and elective inpatient) laparoscopic cholecystectomy				
Ln(OddsRatio) conversion, early vs delayed	-0.224	Normal	0.211	NICE CG188 (NICE 2014)
Delayed (false negative) elective laparoscopic cholecystectomy				
Ln(Odds) conversion, delayed	-2.423	Normal	0.356	NICE CG188 (NICE 2014)
Delayed (false negative) non-elective laparoscopic cholecystectomy				
Ln(OddsRatio) conversion, conservative management vs delayed	-0.03	Normal	0.721	The ln(oddsratio) for conservative management from the health economic model for NICE CG188 was used to represent people undergoing emergency laparoscopic cholecystectomy (NICE 2014)

Open cholecystectomy and conversion are assumed to be 100% effective, without further complications.

Table 4: Probabilities of surgical outcomes

Input	Mean	Distribution	Standard deviation	Source/notes
Early (day case and elective inpatient) laparoscopic cholecystectomy outcomes				
Ln(OddsRatio) Bile duct injury, early laparoscopic cholecystectomy vs delayed laparoscopic cholecystectomy	-0.043	Normal	0.903	NICE CG188 (NICE 2014)
Ln(OddsRatio) Bile leak requiring ERCP, early laparoscopic cholecystectomy vs delayed laparoscopic cholecystectomy	0.91	Normal	0.655	NICE CG188 (NICE 2014)
Delayed (false negative) elective laparoscopic cholecystectomy outcomes				
Ln(Odds) Bile duct injury with delayed laparoscopic cholecystectomy	-4.503	Normal	0.429	NICE CG188 (NICE 2014)
Ln(Odds) Bile leak requiring ERCP with delayed laparoscopic cholecystectomy	-3.635	Normal	0.371	NICE CG188 (NICE 2014)
Delayed (false negative) non-elective laparoscopic cholecystectomy outcomes				
Bile duct injury	As above	As above	As above	Not found in NICE CG188, assume equal to bile duct injury in delayed elective
Ln(OddsRatio) Bile leak requiring ERCP, Conservative management vs delayed laparoscopic cholecystectomy	0.258	Normal	0.778	NICE CG188 (NICE 2014)
ERCP outcomes				
Ln(Odds) Bile leak (requires ERCP)	-3.9	Normal	0.397	NICE CG188 (NICE 2014)
Ln (OddsExtra ERCP required	-2.595	Normal	0.466	NICE CG188 (NICE 2014)
Abbreviations: 95% CI: 5% confidence intervals; ERCP: endoscopic retrograde cholangiopancreatography; FCEs: finished consultant episodes				

2.3 Costs

2.3.1 Hand-held/portable ultrasound costs

Due to uncertainty in whether Healthcare Resource Groups RD44Z-RD46Z (Ultrasound Scan, Mobile or Intraoperative Procedures) represents hand-held ultrasound, the prices of specific hand-held devices (Table 6) were used to calculate a cost per person of hand-held/portable POCUS. It was considered that HRG RD44Z may represent larger mobile ultrasound devices (i.e. not pocket-sized). In addition, it covers intra-operative use which is not relevant to the use of hand-held POCUS to evaluate suspected gallstones in the ED.

The calculations for cost per person of hand-held/portable POCUS are detailed in tables 5, 6 and 7.

To estimate the number of people likely to receive portable POCUS for suspected gallstones in the A&E setting, the numbers of emergency cases in each Local Health Board (LHB) in Wales for the HRG GC17A-K (Non-Malignant, Hepatobiliary or Pancreatic Disorders) were obtained from the Patient Episode Database Wales (PEDW) Statistics 2018/2019 (DCHW 2020). PEDW lists 6,635 emergency cases for GC17A-K across all seven LHBs. To estimate the number of portable POCUS devices required to serve this population, the number of Major A&E Departments and A&E/Minor Injury Units per Local Health Board in Wales was used to calculate the number of emergency cases per A&E department (StatsWales 2020). It was assumed that one additional device would be needed for every 500 people visiting each A&E department. This equates to a total of 21 devices in Wales.

Table 5. Estimating the number of hand-held/portable point-of-care ultrasound devices required in Wales

	Emergency attendances (GC17A-GC17K)[a]	Number of Major A&E Departments[b]	Cases per A&E department	Assumed number of devices required per health board[c]
Betsi Cadwaladr UHB	1,623	3	541	6
Hywel Dda UHB	769	3	256	3
Swansea Bay UHB	1,266	1	1,266	3
Cardiff and Vale UHB	754	1	754	2
Cwm Taf Morgannwg UHB	1,029	3	343	3
Aneurin Bevan UHB	1,193	2	597	4
Powys Teaching HB	1	0	-	0

[a]Sum of 'emergency' attendances for HRGs GC17A, GC17B, GC17C, GC17D, GC17E, GC17F, GC17G, GC17H, GC17J, GC17K (Non-Malignant, Hepatobiliary or Pancreatic Disorders). Source: (Patient Episode Database Wales Statistics 2018/2019 (DCHW 2020))
[b]Source: StatsWales (2020)
[c]Assumption that 1 device is required per 500 emergency attendances in each A&E department
Abbreviations: A&E: accident and emergency; HB: health board; UHB: university health board; PEDW: Patient Episode Database Wales

The studies included in this review (main report, summary of studies Table 1) referred to the following specific devices: SonoSite M Turbo, Sonosite Micromaxx, Samsung HM-70, GE VScan and LOGIQ E. None of these devices were listed in the NHS Supply Chain Online Catalogue (search date 6th October 2020) (NHS Supply Chain 2020). However, the price of one device (Vscan (GE)) was found in an NHS Supply Chain document, though not listed in the Online Catalogue (NHS Supply Chain). This device cost was used in the base case of the model. The impact of other device costs on the results of the health economic analysis were explored in sensitivity analyses. The capital cost for each device is assumed to be spread across a seven year lifetime (Table 6).

Table 6. Cost per person of hand-held/portable point-of-care ultrasound devices

Brand and description	Scenario	Price	Total cost over 7 years[a]	Cost per person (£)[b]	Reference
Vscan (GE) with Premium Service Contract 1 + 3 years' Service	Base case	£6,045	£126,945	£2.73	https://my.supplychain.nhs.uk/catalogue (not found on manufacturer website)

Brand and description	Scenario	Price	Total cost over 7 years[a]	Cost per person (£)[b]	Reference
SonoSite M Turbo	Scenario 6	$5,555 (£4,247.58[b])	£89,199	£1.92	https://cardomedical.com/product/sonosite-m-turbo/ (not found on manufacturer website)
HM-70 Samsung	Scenario 7	$24,000 (£18,372.72[b])	£385,827	£8.31	https://www.medicalpriceonline.com/medical-equipment/samsung/samsung-ugeo-hm70/ (not found on manufacturer website)
LOGIQ e	Scenario 8	$11,978 (£9,166.28[b])	£192,492	£4.14	https://bimedis.com/ge-logiq-e-m357 (not found on manufacturer website)
Sonosite Micromaxx	Scenario 9	$7,975 (£6,047[c])	£126,983	£2.73	http://www.sale-medical.com/products/SonoSite-MicroMaxx-Portable-Ultrasound-Machine.html

[a]Assuming a total number of devices required for emergency cases in Wales of 21.
[b]Assuming total number of emergency cases of 6,635 per year (Patient Episode Database Wales Statistics 2018/2019 (DCHW 2020)) and a device lifetime of 7 years.
[b]Converted using xe.com currency conversion tool. Exchange rate: 1USD = 0.76266 GBP. Date: 14 October 2020
[c]Converted using xe.com. Exchange rate: 1USD = 0.75822GBP. Date 27 October 2020

For the purposes of the model, a week-long training course was assumed, based on the studies included in this report. The Unit Costs of Health and Social Care 2019 (Personal Social Services Research Unit) were used for the cost per 48-hour week. Specialty Registrars and Medical/Surgical Consultants cost £2,245 and £5,227 per 48 hour working week, respectively (PSSRU 2019). In the base case, it was assumed that all 21 individuals trained were Surgical Consultants. The cost of an AMA-accredited American Institute of Ultrasound in Medicine course on 'Right Upper Quadrant Pain: Ultrasound First!' ($60 (£45.72) for a non-member) was applied, assuming that one member of staff requires training per device in Wales. Training costs were spread over 5 years. The cost of training per person undergoing hand-held/portable POCUS is £1.19.

To calculate the costs of staff time for delivering hand-held/portable POCUS, the average duration of scans (15 minutes (10 to 20 minutes)) reported by Gustafsson et al. (2018) was used. No other studies in the gallstones population reported duration of scans. The cost per hour is £47 and £109 (for a 48-hour week) for a Specialty Registrar and Medical/Surgical Consultant, respectively. In the base case, it was assumed that the cost of a Surgical Consultant would apply, in line with Gustafsson et al. (2018). The cost of staff time for delivering the scan is £27.22 per person.

The total cost per person of hand-held/portable POCUS used in the base case of the model is £33.29. This incorporates the costs of staff time for delivering the scan, staff training and device costs.

Table 7. Staff costs for training and delivering hand-held/portable point-of-care ultrasound

Input	Specialty Registrar		Surgical Consultant
Working hours, 48-hour week[a]	£2,245		£5,227
Cost per hour, 48-hour week[a]	£47		£109
Proportion (base case)	0%		100%
Training			
Staff costs for training week (21 staff members; one staff member per device)	£109,761		
Total cost of course for 21 staff members	£960.12		
Cost of training per person scanned[b]	£3.34		
Delivery of hand-held/portable POCUS			
Time (minutes)	Average	Lower	Upper
	15	10	20
Cost per scan	£33.29	£24.22	£42.37

[a]Source: Unit Costs of Health and Social Care 2018/2019(PSSRU 2019)
[b]Assuming training costs spread over 5 years and emergency cases of 6,635 per year (Patient Episode Database Wales Statistics 2018/2019 (DCHW 2020))
Abbreviations: POCUS: point of care ultrasound

2.3.2 Radiologist-performed ultrasound cost

The national average unit costs of ultrasound scans with a duration of less than 20 minutes are from the NHS Reference Costs 2018 to 2019 (Table 8) (NHS England 2020) . This duration was selected as the study by Gustafsson et al. (2018) reported a duration of 10 minutes (5-15 minutes) for radiologist-performed ultrasound. The costs of ultrasound costs without contrast are presented, on the advice of Welsh clinical experts. The direct access costs were used in the model.

The costs of these tests were varied using a gamma distribution in probabilistic sensitivity analysis. The national average, lower quartile and upper quartile unit costs for the equivalent Healthcare Resource Groups (HRGs) were taken from the NHS Reference Costs 2016/2017 in order to vary the 2018/2019 unit costs in a probabilistic sensitivity analysis (as the 2018/2019 NHS Reference Costs do not provide lower and upper quartile unit costs). The difference between the 2016/2017 lower and upper quartiles and the 2016/2017 national average unit cost was then added to the 2018/2019 national average cost to give an estimate of the 2018/2019 lower and upper quartile unit costs.

Table 8. Unit costs of ultrasound scans

Examination	Unit cost (£)	
	Mean	α, β for gamma distribution
HRG: RD40Z Ultrasound Scan, with duration of less than 20 minutes, without contrast	£52.13	3025466, 1.72^E-05

Source: NHS Reference Costs 2018-2019 (NHS England 2020)
HRG- Healthcare Resource Group

In the model base case, the difference waiting time to receive the scan between hand-held/portable ultrasound and radiology-performed ultrasound reported in Kell et al. (2002) is applied (3.1 hours versus 12.0 hours for radiology US). In a sensitivity analysis, the difference in waiting time to receive the scan between hand-held/portable ultrasound and radiology-performed ultrasound was set at 0. To account for the difference in waiting time in the costs, a weighted average cost-per-hour was derived from the national average unit costs for a non-elective short stay for 'Non-Malignant, Hepatobiliary or Pancreatic Disorders, without Interventions', for which the length of stay is 1. This approach was used as the latest iteration of the NHS Reference Costs do not include length of stay data. As with the above cost, the cost per hour was varied using a gamma distribution in probabilistic sensitivity analysis. The proportions of finished consultant episodes were varied using a Dirichlet function.

Table 9. Cost per hour (for time to scan)

	FCEs	Unit cost (£)	α, β for gamma distribution
Non-Malignant, Hepatobiliary or Pancreatic Disorders, without Interventions, with CC Score 8+	5,690	£25.80	38026, 0.0007
Non-Malignant, Hepatobiliary or Pancreatic Disorders, without Interventions, with CC Score 5-7	10,069	£23.27	123707, 0.0002
Non-Malignant, Hepatobiliary or Pancreatic Disorders, without Interventions, with CC Score 2-4	25,524	£21.31	312484, 6.8^E-05
Non-Malignant, Hepatobiliary or Pancreatic Disorders, without Interventions, with CC Score 0-1	41,936	£18.91	605548, 3.12^E-05
Weighted average cost per hour		£20.65	

Source: NHS Reference Costs 2018-2019 (NHS England 2020)
FCEs: Finished consultant episodes
HRG- Healthcare Resource Group; OPCS- OPCS Classification of Interventions and Procedures

2.3.3 Cost of other diagnostic tests

NICE guideline CG188 stipulates that people with symptomatic gallstones are offered radiology-performed ultrasound and liver function tests (NICE 2014). The model therefore assumes that everyone has a liver function test.

In line with NICE recommendations, people who receive negative results on either hand-held or radiologist-performed ultrasound, but in whom liver function tests are abnormal receive MRCP. The costs of liver function tests and MRCP are included in Table 10. The costs of these tests were varied using a gamma distribution in probabilistic sensitivity analysis, with upper and lower values estimated based on NHS Reference Costs 2016-2017; the last iteration of the reference costs to publish these figures.

Table 10. Unit costs of liver function tests and magnetic resonance cholangiopancreatography

Investigation	Unit cost (£)	α, β for gamma distribution
HRG: DAPS04 Clinical Biochemistry (liver function tests)	£1.10	2.4^E+09, 4.65^E-10
HRG: RD01A Magnetic Resonance Imaging Scan of One Area, without Contrast, 19 years and over	£120.83	3501133, 3.45^E-05

Investigation	Unit cost (£)	α, β for gamma distribution
(OPCS: U162, Magnetic resonance cholangiopancreatography)		
Source: NHS Reference Costs 2018-2019 (NHS England 2020) HRG- Healthcare Resource Group; OPCS- OPCS Classification of Interventions and Procedures		

2.3.4 Costs of management of gallstones

In the model, people who are diagnosed with gallstones only (true positives and false positives) are assumed to receive a day-case laparoscopic cholecystectomy, an elective inpatient laparoscopic cholecystectomy or an elective inpatient open cholecystectomy. This was based on the NICE Guideline CG188 recommendation that laparoscopic cholecystectomy is offered to people diagnosed with symptomatic gallbladder stones (NICE 2014). The Guideline further recommends that day-case laparoscopic cholecystectomy is offered for people having an elective planned procedure, unless the circumstances or clinical condition necessitate an inpatient stay. These assumptions were confirmed by expert clinical opinion.

On the basis of clinical opinion, people diagnosed with common bile duct stones receive either ERCP alone, or ERCP with either laparoscopic or open cholecystectomy. Clinical experts advised that ERCP and intraoperative laparoscopic cholecystectomy is not standard practice in NHS Wales, so this option was not included in the model.

The costs of ERCP, laparoscopic and open cholecystectomy are included in Table 11. The costs which are applied depend on whether individuals in the model experience complications. If no complications occur, the 'uncomplicated' costs in Table 11 apply. The management options for people diagnosed with symptomatic gallbladder stones are summarised below:

- People who have a bile duct injury incur the cost of a complicated laparoscopic cholecystectomy
- People who have a bile leak requiring non-elective ERCP during initial laparoscopic cholecystectomy incur the cost of complicated laparoscopic cholecystectomy and an additional non-elective ERCP
- People who have a bile leak requiring non-elective ERCP during initial ERCP incur the cost of both elective inpatient ERCP non-elective ERCP.
- People who require an additional ERCP after initial ERCP incur the cost of both elective inpatient ERCP and non-elective ERCP.

For the cost of laparoscopic cholecystectomy, the cost of the highest CC score was reserved for the cost of conversion, which was taken from 'total HRGs'. The same cost of conversion is applied for all people in the model who undergo conversion from laparoscopic to open cholecystectomy, regardless of the setting (i.e. day case, elective inpatient). No further complications of conversion are included in the model.

No complications are explicitly included in the model for open cholecystectomy. However, as the unit costs are weighted by number of finished consultant episodes for different complications and comorbidities scores, the costs of adverse are assumed to be included within the national average unit costs. However, these data are not specific to the gallstones population.

In the base case, no long term costs of complications are included in the model. However, in a sensitivity analysis, the approach used by Gurusamy et al was taken, which used an extreme lifetime cost of bile duct injury (£100,000), to account for the structural uncertainty of a limited time horizon (1 year).

In probabilistic sensitivity analysis, the costs in Table 11 (with the exception of conversion, which was fixed) were varied using a gamma distribution. The cost of conversion was fixed as estimates of the upper and lower range could not be found. The proportions of finished consultant episodes were varied using a Dirichlet function.

Table 11. Unit costs of laparoscopic and open cholecystectomy and endoscopic retrograde cholangiopancreatography

	Procedure	FCEs	Unit Cost (£)	α, β for gamma distribution
Complicated day case laparoscopic cholecystectomy	HRG: GA10J Laparoscopic Cholecystectomy, 19 years and over, with CC Score 1-3. Day case	7,138	£2,480	180575, 0.014
Uncomplicated day case laparoscopic cholecystecomy	HRG: GA10K Laparoscopic Cholecystectomy, 19 years and over, with CC Score 0. Day case	20,897	£2,452	533989, 0.005
Complicated elective inpatient laparoscopic cholecystectomy	HRG: GA10J Laparoscopic Cholecystectomy, 19 years and over, with CC Score 1-3. Elective inpatient	7,601	£3,405	170154, 0.02
Uncomplicated elective inpatient laparoscopic cholecystectomy	HRG: GA10K Laparoscopic Cholecystectomy, 19 years and over, with CC Score 0. Elective inpatient	10,699	£3,160	268774, 0.012
Complicated non-elective laparoscopic cholecystectomy	HRG: GA10K Laparoscopic Cholecystectomy, 19 years and over, with CC Score 1-3. Non-elective short stay	10%	£2,601	7004, 0.37
	HRG: GA10K CC Score 1-3. Non-elective long stay	90%	£5,045	128350, 0.039
	Weighted average		£4,793	
Uncomplicated non-elective laparoscopic cholecystectomy	HRG: GA10K Laparoscopic Cholecystectomy, 19 years and over, with CC Score 0. Non-elective short stay	19%	£2,463	5946, 0.414
	HRG: GA10K CC Score 0. Non-elective long stay	81%	£4,512	111447, 0.04
	Weighted average		£4,126	
Conversion	Laparoscopic Cholecystectomy, 19 years and over, with CC Score 4+. Total HRGs	2,896	£4,978	Fixed
Elective inpatient open cholecystectomy	HRG: GA10L Open Cholecystectomy, 19 years and over, with CC Score 3+. Elective inpatient	325	£6,844	3997, 1.7
	HRG: GA10M CC Score 1-2	409	£5,288	4889, 1.08
	HRG: GA10N CC Score 0	436	£4,273	6939, 0.62
	Weighted Average	£5,342		
Elective inpatient ERCP	HRG: GB06E Intermediate Therapeutic Endoscopic Retrograde Cholangiopancreatography with CC Score 6+. Elective inpatient	322	£3,499	1459, 2.40

	Procedure	FCEs	Unit Cost (£)	α, β for gamma distribution
	HRG: GB06E CC Score 4-5	437	£2,126	1151, 1.85
	HRG: GB06E CC Score 2-3	1,135	£1,890	5714, 0.33
	HRG: GB06E CC Score 0-1	1,895	£1,730	8899, 0.19
	Weighted average	£1,974		
Non-elective ERCP	Intermediate Therapeutic Endoscopic Retrograde Cholangiopancreatography with CC Score 6+. Non-elective short stay	1,111	£1,101	4842, 0.23
	CC Score 4-5. Non-elective short stay	749	£966	3702, 0.26
	CC Score 2-3. Non-elective short stay	1,206	£986	3747, 0.26
	CC Score 0-1. Non-elective short stay	1,240	£1,030	5841, 0.18
	Intermediate Therapeutic Endoscopic Retrograde Cholangiopancreatography with CC Score 6+. Non-elective long stay	2,710	£5,620	30274, 0.19
	CC Score 4-5. Non-elective long stay	1,670	£4,196	20622, 0.20
	CC Score 2-3. Non-elective long stay	2,587	£3,552	29953, 0.12
	CC Score 0-1. Non-elective long stay	2,549	£3,205	46512, 0.07
	Weighted average		£3,184	

Source: NHS Reference Costs 2018-2019 (NHS England 2020)
CC- Casemix Companion; HRG- Healthcare Resource Group

2.4 Quality of life inputs

As recommended in the NICE reference case, the model estimates effectiveness in terms of quality adjusted life years (QALYs). These are estimated by combining life year estimates with quality of life (QoL) values associated with being in a particular health state. Mortality is not considered in this analysis because it is not anticipated that there would be survival differences between the two strategies. Therefore, differences in QALYs will be entirely driven by differences in QoL.

A search of the CEA Registry for quality of life weights for 'laparoscopic cholecystectomy' identified a Canadian cost-utility analysis published in 2020 (CEVR 2020, Sutherland et al. 2020). While one UK-based cost-utility analysis in the gallbladder and common bile duct stones population and one in the acute cholecystitis population were also identified, the studies was found to use quality of life data from an Australian study published in 1994 (Wilson et al. 2010, Gurusamy et al. 2012). A second search was undertaken using the search term 'ERCP', through which a third UK-based cost-utility analysis was identified (Morris et al. 2015). The study considers the population of 'suspected common bile duct stones' and it obtained utility data from an Australian study. No utility data for laparoscopic cholecystectomy were found which used the UK value set.

The 'pre-operative for laparoscopic cholecystectomy' utility value from Sutherland et al. (2020) was used in the base case for all individuals in the model prior to receiving surgery (Table 12).

After surgery, either the immediate post-laparoscopic cholecystectomy, immediate post-open cholecystectomy (Wilson et al. 2010), immediate post-ERCP or bile duct injury utilities apply accordingly. Morris et al. (2015) reported utilities for both complicated and uncomplicated ERCP.

Complications of surgeries:

- People who sustain a bile duct injury during laparoscopic cholecystectomy receive the long term bile duct injury utility
- People who experience bile leak during laparoscopic cholecystectomy receive ERCP, and so the post-ERCP utility (complicated) applies instead of the post-laparoscopic cholecystectomy utility.
- People requiring an additional emergency ERCP following ERCP receive the post-ERCP (complicated) utility.

Following the post-surgical period, people in the model receive the post-operative (6 month) utility for laparoscopic cholecystectomy from Sutherland et al. (2020), unless a bile duct injury was sustained.

Table 12. Utility estimates in people with gallstones

Description	Utility	Distribution	α, β	Source	Notes
Pre-operative for laparoscopic cholecystectomy	0.8419	Beta	114, 21	Sutherland et al. (2020)	Population: Symptomatic gallstones. EQ-5D-3L, Canadian valuation
Immediate post-laparoscopic cholecystectomy	0.90	Beta	3.48, 0.39	Wilson et al. (2010) Gurusamy et al. (2012)	UK-based cost utility analysis using Australian QoL data from (Cook et al. 1994)
Immediate post- open cholecystectomy	0·80	Beta	255.2, 63.8	Wilson et al. (2010), Gurusamy et al. (2012).	UK-based cost utility analysis using Australian QoL data from (Cook et al. 1994)
Immediate post-Therapeutic Endoscopic Retrograde Cholangiopancreatography (uncomplicated)	0.89	Beta	17.95, 2.22	Morris et al. (2015)	UK based cost-utility analysis using utility data from an Australian study (Howard et al. 2006)
Immediate post-Therapeutic Endoscopic Retrograde Cholangiopancreatography (complicated)	0.7596	Beta	29.24, 9.25	Morris et al. (2015)	UK based cost-utility analysis using utility data from an Australian study (Howard et al. 2006)
Post-operative (6 months) for laparoscopic cholecystectomy	0.9080	Beta	122.6, 12.42	Sutherland et al. (2020)	Population: Symptomatic gallstones. EQ-5D-3L, Canadian valuation

Description	Utility	Distribution	α, β	Source	Notes
Bile duct injury	0.8	Beta	24.8, 6.2	Gurusamy et al. (2012)	Based on Cook et al 1994 (study not found)
[a]Studies identified through search for utility data using search terms 'laparoscopic cholecystectomy' and 'ERCP' (CEVR 2020) EQ-5D-3L: The 3-level version of EuroQoL-5 Dimensions; QoL: quality of life					

In the base case analysis, the difference in waiting time before receiving a scan between hand-held/portable ultrasound and radiology-performed ultrasound (3.1 hours versus 12.0 hours for radiology US, as reported in Kell et al. (2002)) was assumed to have an impact on quality of life. The quality of life impact was quantified by applying the pre-operative utility for longer in RADUS compared with POCUS. In a sensitivity analysis this impact on quality of life was removed (while the impact on costs was retained). In a further sensitivity analysis, the difference in waiting time before receiving a scan between hand-held/portable ultrasound and radiology-performed ultrasound was set at 0.

The pre-operative utility was applied in the model for different durations, depending on the result of the initial (either hand-held or radiology-performed) ultrasound (Table 13). The durations over which the pre-operative utilities were applied (and the ranges for probabilistic sensitivity analysis) were as follows:

- True positives:
    - Day case laparoscopic cholecystectomy is carried out at 2 days (2 days-`1 month). (clinical expert assumption)
    - Elective inpatient laparoscopic cholecystectomy is carried out at 2 months (1 month- 3 months)
    - Elective inpatient open cholecystectomy is carried out at 2 months (1 month-`3 months).
    - Elective inpatient ERCP is at 2 days in the base case (range 2 days to 2 months), with subsequent elective inpatient laparoscopic or open cholecystectomy after recovery from ERCP at 2 months (as per waiting lists) (clinical expert assumption).
- False positives:
    - As above for true positives
- False negatives:
    - In people with abnormal liver function tests, MRCP is carried out at 1 day in the base case (range 1 day to 1 month). After MRCP, people are assumed to have an elective inpatient ERCP after 2 further days in the base case (range 2 days to 2 months). Subsequent elective inpatient open or laparoscopic cholecystectomy is assumed to occur after recovery from ERCP, at 2 months as per waiting lists (clinical expert assumption).
    - People with normal liver function tests are assumed to either require emergency laparoscopic cholecystectomy (and RADUS) at 9 weeks (range 0-18 weeks) or to re-present and undergo radiology-performed ultrasound at 18 weeks (6-40 weeks). This was based on Wilson et al. (2010), which reported the duration 9 weeks a 'time to development of symptoms' and 18 weeks for 'wait time', based on an author estimate. After RADUS to confirm their diagnosis, people in the model receive day case laparoscopic cholecystectomy at 2 days, or elective inpatient open cholecystectomy at 2 months.
- True negatives:
    - People with normal liver function test results receive the utility for long term (6 months) post-laparoscopic cholecystectomy throughout (Table 11). People with

abnormal liver function test results receive the pre-operative utility until MRCP (carried out at 1 day) and then receive the utility for long term (6 months) post-laparoscopic cholecystectomy.

For the durations over which the post-operative utilities were applied after laparoscopic and open cholecystectomy, a model input of 2 weeks was used, based on the Gurusamy et al. (2012) study. The study reported the duration of the post-operative period in the short and medium term based on an author estimate. This was corroborated by an assumption made for the NICE Guideline health economic model, which assumed that quality of life returns to normal following laparoscopic cholecystectomy within between one and four to six weeks.

For the durations over which the post-operative utilities were applied after ERCP in people having subsequent open or laparoscopic cholecystectomy, an assumption was made that people in the model would receive the 'immediate post-operative ERCP' utility value for the entire intervening two months between operations, whereas people having ERCP alone recover at either 1 or 4 weeks for uncomplicated and complicated ERCP respectively, based on assumptions made in a UK cost-utility analysis. Morris et al. (2015) gave the duration of health states for therapeutic ERCP.

Following the immediate post-operative durations, people in the model assume the post-operative (6 months) laparoscopic cholecystectomy utility for the remainder of the time horizon (or the long term utility for bile duct injury if sustained during surgery).

Table 13. Durations for applying utility estimates

Description	Duration	Distribution	Range	Source	Notes
Durations for pre-surgical states					
Time until day case laparoscopic cholecystectomy	2 days	Uniform	2 days-1 month (assumption)	Clinical opinion/ Assumption	
Time until elective inpatient laparoscopic cholecystectomy	2 months	Uniform	1 month to 3 months	Clinical opinion/ Assumption	
Time until elective inpatient open cholecystectomy	2 months	Uniform	1 month to 3 months	Clinical opinion/ Assumption	
Time until elective inpatient ERCP	2 days	Uniform	2 days - 2 months	Clinical opinion	
Time between ERCP and elective inpatient laparoscopic or open cholecystectomy	2 months	Uniform	2 weeks to 2 months	Clinical opinion/ Assumption	
Time to MRCP	1 day	Uniform	1 day-1 month		
Time between ERCP and laparoscopic or open cholecystectomy	2 months	Uniform	2 weeks to 2 months		

Description	Duration	Distribution	Range	Source	Notes
Time to repeat RADUS (false negatives with normal liver function test results)	18 weeks	Uniform	6 weeks to 40 weeks	Gurusamy et al. (2012)	If emergency laparoscopic cholecystectomy not required. Gurusamy et al. (2012) used range 6-52 weeks and a triangular distribution. Uniform was used for consistency
Time to repeat RADUS and emergency laparoscopic cholecystectomy (false negatives with normal liver function test results)	9 weeks	Uniform	0 weeks to 18 weeks	Gurusamy et al. (2012)	Assumption: No delay between RADUS and non-elective laparoscopic cholecystectomy.
Durations for post-surgical states					
Immediate post-laparoscopic cholecystectomy	2 weeks	Uniform	1 week to 3 weeks	Gurusamy et al. (2012)	Author estimate for short term (corresponding to immediate post-procedure period
Immediate post- open cholecystectomy	2 weeks	Uniform	1 week to 3 weeks	Gurusamy et al. (2012)	Author estimate for short term (corresponding to immediate post-procedure period
Immediate post-Therapeutic Endoscopic Retrograde Cholangiopancreatography (uncomplicated)	1 week	Uniform	0-2 weeks	Morris et al. (2015)	UK based cost-utility analysis using utility data from an Australian study (Howard et al)
Immediate post-Therapeutic Endoscopic Retrograde Cholangiopancreatography (complicated)	4 weeks	Uniform	2-6 weeks	Morris et al. (2015)	UK based cost-utility analysis using utility data from an Australian study (Howard et al)
Post-operative (6 months) for laparoscopic cholecystectomy	Remainder of time horizon	-	-		

[a]Studies identified through search for utility data using search terms 'laparoscopic cholecystectomy' and 'ERCP' (CEVR 2020)

3. Results

3.1 Base case results

The results of the base case analysis is presented in Table 14, which shows the total and incremental costs and QALYs over the time horizon (presented on a per patient basis) as well as the incremental cost effectiveness ratio (ICER). In the deterministic analysis, POCUS was found to be less costly but less effective than radiologist-performed ultrasound. The resulting ICER of £58,631 per QALY indicates that POCUS is cost effective at the £20,000 per QALY threshold. (N.B. In the south west quadrant of the cost effectiveness plane, an ICER above £20,000 per QALY is cost effective, as it indicates that the money saved per QALY lost could be better spent elsewhere).

Table 14. Base case results (deterministic)

Treatment strategy	Radiologist-performed ultrasound (1)			Hand-held/portable ultrasound (2)		
Ultrasound test result	n	Costs	QALYs	n	Costs	QALYs
True positives	415	£1,417,543	373	372	£1,193,612	333.9
False negatives	12	£40,095	11.7	55	£180,409	48.4
False positives	6	£21,516	5.7	23	£73,599	20.6
True negatives	567	£137,763	514.6	550	£22,247	499.6
Total	1000	1,616,917	905.0	1000	£1,469,866	902.5
Difference in costs (2-1) per 1000 people			Saves £147,051			
Difference in QALYs (2-1) per 1000 people			2.51 fewer			
ICER			£58,631 per QALY			

ICER: incremental cost effectiveness ratio; QALYs: quality-adjusted life years
N.B. In the South West quadrant of the cost-effectiveness plane, an ICER above £20,000 per QALY indicates that hand-held/portable ultrasound is more cost effective, as more money is saved per QALY lost.

3.2 Deterministic sensitivity analysis results

A series of deterministic sensitivity analyses were conducted, whereby an input parameter is changed, the model is re-run and the new cost-effectiveness result is recorded. This is a useful way of estimating uncertainty and determining the key drivers of the model result.

3.2.1 Scenario analyses

Table 15 shows that the model results are sensitive to some key assumptions. In particular, in scenario 12, POCUS becomes dominated (less effective and more costly) if there is no difference in time to scan between POCUS and RADUS. The results also change if different model inputs are used for the diagnostic accuracy of POCUS, based on the Colli et al. (2015) and Del Medico et al. (2018) studies (scenarios 4 and 5).

Table 15. Results of scenario analyses

Scenario	Modelled scenario	Result
	Base case	£58,631 per QALY
1	Hand-held/portable US diagnostic accuracy data from Shekarchi et al. (2018)	£84,995 per QALY

Scenario	Modelled scenario	Result
2	Hand-held/portable US diagnostic accuracy data from Kell et al. (2002)	£137,660 per QALY
3	Hand-held/portable US diagnostic accuracy data from Ismaeel et al. (2010)	£73,381 per QALY
4	Hand-held/portable US diagnostic accuracy data from Colli et al. (2015)	Dominated
5	Hand-held/portable US diagnostic accuracy data from Del Medico et al. (2018)	£15,441 per QALY
6	Device costs Sonosite MTurbo	£58,955 per QALY
7	Device costs HM-70	£56,409 per QALY
8	Device costs LOGIQe	£58,068 per QALY
9	Device costs Sonosite MicroMaxx	£58,631 per QALY
10	Number of devices required in Wales equalling number of major A&E departments (13 devices)	£59,553 per QALY
11	Number of devices required in Wales equalling five per major A&E department (65 devices)	£53,560 per QALY
12	No difference in waiting time to receive the scan between POCUS and RADUS	Dominated
13	No impact of difference in waiting time to receive the scan on utilities	£57,112 per QALY
14	No impact of difference in waiting time to receive the scan on costs	Dominated
15	Extreme long term cost of bile duct injury: £100,000	£51,921 per QALY
16	Emergency laparoscopic cholecystectomy probability 18%	£62,568 per QALY

RADUS: Radiology-performed ultrasound; US: ultrasound
N.B. In the South West quadrant of the cost-effectiveness plane, an ICER above £20,000 per QALY indicates that hand-held/portable ultrasound is more cost effective, as more money is saved per QALY lost.

3.2.2 Two-way sensitivity analysis

Table 16. Two-way deterministic sensitivity analysis of sensitivity and specificity of hand-held/portable ultrasound: effect on ICER

	S	E	N	S	I	T	I	V	I	T	Y
S		60%	70%	80%	90%	95%	96%	97%	98%	99%	100%
P	60%	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated
E	70%	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated
C	80%	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated
I	90%	£6,056	£7,902	£10,655	£15,209	£18,821	£19,719	£20,694	£21,755	£22,914	£24,186
F	95%	£24,806	£30,668	£39,622	£54,992	£67,709	£70,949	£74,499	£78,406	£82,726	£87,528
I	96%	£28,766	£35,530	£45,914	£63,884	£78,894	£82,739	£86,961	£91,619	£96,784	£102,544
C	97%	£32,801	£40,504	£52,389	£73,132	£90,628	£95,135	£100,096	£105,584	£111,688	£118,516
I	98%	£36,913	£45,593	£59,056	£82,757	£102,954	£108,187	£113,962	£120,369	£127,516	£135,541
T	99%	£41,104	£50,802	£65,923	£92,783	£115,916	£121,947	£128,621	£136,046	£144,358	£153,724
Y	100%	£45,377	£56,134	£72,999	£103,236	£129,566	£136,475	£144,143	£152,701	£162,213	£173,188

3.2.3 Threshold analysis

A threshold analysis was conducted to assess the uncertainty around 'time to scan'. It showed that the difference in 'time to scan' needs to be 4.2 hours or greater for POCUS to be cost effective.

3.3 Probabilistic sensitivity analysis results

Probabilistic sensitivity analysis (PSA) was conducted to assess the combined parameter uncertainty in the model. In this analysis, the mean values that were utilised in the base case were replaced with values drawn from distributions around the mean values. The results of 10,000 runs of the PSA are shown using ICER scatterplots and cost-effectiveness acceptability curves (CEAC). The ICER scatter plots show the incremental costs and QALYs associated with each of the 10,000 runs of the PSA along with the mean result. It can be seen that there are points across all four quadrants, with most results in the South West quadrant (indicating that POCUS is cheaper, but less effective than RADUS). There are also a substantial number of points in the North West quadrant, where POCUS is dominated (more expensive and less effective than RADUS).

The CEAC graphs show the probability of each strategy being considered cost effective at the various cost-effectiveness thresholds on the x axis. It can be seen that, at a threshold of £20,000 per QALY, POCUS was found to have a 49% probability of being cost effective, while there was a 51% probability that RADUS was cost effective.

Figure 2. ICER scatterplot for POCUS compared with RADUS

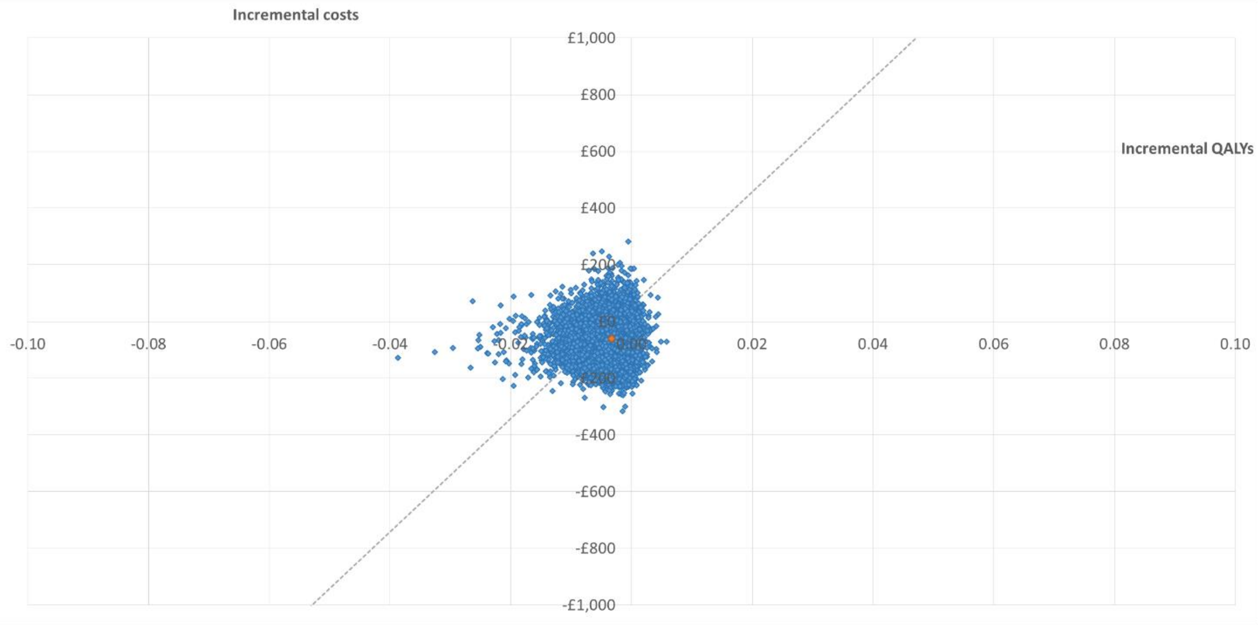
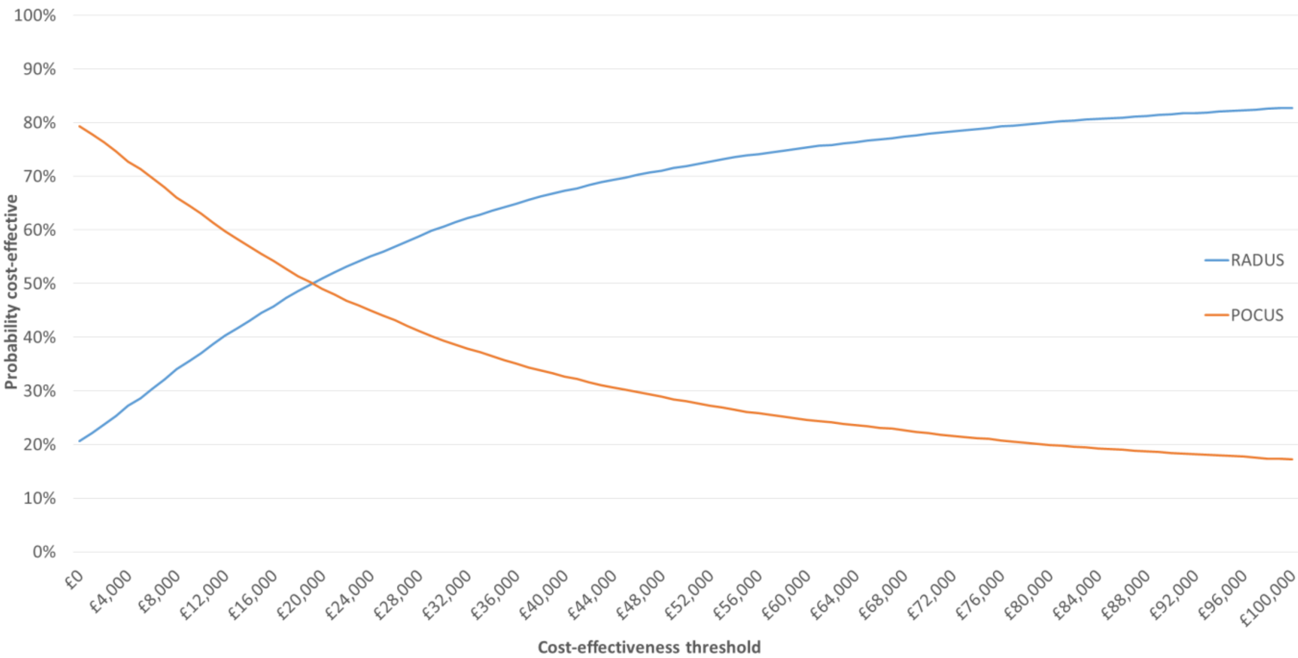


Figure 3. CEAC for POCUS compared with RADUS



4. Discussion

4.1 Discussion

The results of the base case (deterministic analysis) show that hand-held/portable US was cheaper, though less effective in comparison with radiologist-performed ultrasound. The ICER of £58,631 per QALY indicates that POCUS is cost effective at the £20,000 per QALY threshold. (N.B. In the south west quadrant of the cost effectiveness plane, an ICER above £20,000 per QALY is cost effective, as it indicates that the money saved per QALY lost could be better spent elsewhere).

The result was found to be driven primarily by an assumption made in the base case of the model that there is a reduced waiting time to receive the scan in POCUS compared with RADUS, which was based on an outcome reported by a single study (Kell et al. (2002), Table 2). In a scenario analysis where no difference in 'time to scan' was applied, POCUS is dominated (more costly and less effective) than RADUS. In a threshold analysis, it was found that the difference in 'time to scan' needs to be 4.2 hours or greater for POCUS to be cost effective. In two further sensitivity analyses, the impact of the difference in waiting time to receive the scan on costs and utilities were explored. When it was assumed that there are no cost implications of a difference in waiting times for scans on costs, the results of the analysis changed and POCUS is dominated (more costly and less effective than RADUS). However, it was assumed that there are no quality of life implications of a difference in waiting time for scans, POCUS remains cost effective (though slightly less so than in the base case).

In two other scenario analyses, which applied the diagnostic accuracy data from the Del Medico et al. (2018) and Colli et al. (2015) studies, POCUS was found to be not cost effective (ICER: £15,441 per QALY gained) and dominated, respectively. POCUS was found to be cost effective in the remaining scenario analyses.

In probabilistic sensitivity analysis, at a threshold of £20,000 per QALY, POCUS was found to have a 49% probability of being cost effective, while there was a 51% probability that RADUS was cost effective. The ICER scatterplot shows that there are points across all four quadrants, with most results in the South West quadrant (indicating that POCUS is cheaper, but less effective than RADUS).

4.2 Limitations

- It is beyond the scope of a rapid review to undertake a systematic literature search for all model inputs. This limitation underpins all model inputs apart from diagnostic accuracy.
- Conservative treatment was not included in the model as NICE Guidelines recommend that people with symptomatic gallstones receive cholecystectomy. Conservative treatment is used not uncommonly in symptomatic gallstones (if minor, short-lived and non-recurrent symptoms especially if comorbidity etc.).
- No utility data for laparoscopic cholecystectomy were found which used the UK value set.
- No one in the model receives ERCP with intraoperative laparoscopic cholecystectomy
- Some individuals in the model receive ERCP only, whereas some clinical experts felt that ERCP alone is not recommended unless the patient is not fit to undergo major surgery.
- People who have ERCP only (and true negatives) do not have their gallbladder removed in the model. Therefore there is a probability that these individuals may experience gallbladder cancer or develop gallstones in the future. This is not captured in the model, which has a short time horizon of 6 months.
- Open cholecystectomy and conversion from laparoscopic to open cholecystectomy are assumed to be 100% effective, and complications are not included.
- MRCP and (repeat) radiology-performed ultrasound is assumed to be 100% effective in diagnosing gallstones in false negatives.

- Liver function tests are assumed to flag all individuals with common bile duct stones requiring MRCP.
- Endoscopic ultrasound not included in model, as NICE Guideline recommends MRCP where liver function tests are abnormal. A Clinical expert noted that endoscopic ultrasound is in the gall stone pathway when patients have symptoms with abnormal LFT and ultrasound, with negative MRCP. They felt that endoscopic ultrasound is the most sensitive test for small stones in common bile duct, which is missed in the other two investigations
- People having laparoscopic or open cholecystectomy following ERCP receive the postoperative ERCP utility until the cholecystectomy for the entire intervening time between operations, whereas people having ERCP alone recover at 4 weeks.
- The model assumes there are no remaining unseen CBD or gallstones
- The model does not include emergency open cholecystectomy
- The model does not consider laparoscopic cholecystectomy with intraoperative ERCP. When people having laparoscopic cholecystectomy experience a complication which necessitates ERCP, the cost of a separate emergency ERCP episode is applied, whereas in the utilities this event is assumed to happen concurrently, with the 'complicated ERCP' utility applied for the post-operative period.
- The limitations of the clinical studies from which inputs were derived also apply to the model (see Table 1 in main report)