Surg. Gastroenterol. Oncol. 2026;31(2 Suppl):S5-S14 DOI: 10.21614/sgo-870

Endoscopic Ultrasound as a Triage Test Before ERCP for Suspected Common Bile Duct Stones: An Updated Systematic Review and Diagnostic Test Accuracy Meta-Analysis

Ahmed Abdallah Salman[1]*, Ahmed Elewa[2], Ahmed Abdallah[3], Ahmed Marwan[4]

[1]Department of Internal Medicine, Faculty of Medicine, Cairo University, Egypt
[2]Department of General Surgery, National Hepatology and Tropical Medicine Research Institute, Cairo, Egypt
[3]Department of General Surgery, Faculty of Medicine, Cairo University, Egypt
[4]Department of Internal Medicine, Faculty of Medicine, Mansoura University, Egypt

***Corresponding author:**
Ahmed Abdallah Salman, MD
Department of Internal Medicine
Faculty of Medicine, Cairo University
Egypt
Telephone: +20 100 046 8664
E-mail: awea84@kasralainy.edu.eg
awea844@gmail.com

ABSTRACT

Background: Endoscopic retrograde cholangiopancreatography (ERCP) is therapeutic but exposes patients without persistent common bile duct (CBD) stones to avoidable harm. We assessed the diagnostic accuracy of endoscopic ultrasound (EUS) and the clinical role of an EUS-first triage strategy.

Methods: This updated systematic review used the 13-study EUS dataset from a Cochrane review through September 2012 and reproducible PubMed and Scopus searches from January 2013 to 21 March 2026. Studies of adults with suspected CBD stones were eligible. Diagnostic 2 x 2 data were synthesized with a bivariate random-effects model; strategy outcomes were summarized separately. Risk of bias was assessed with QUADAS-2 and certainty with GRADE. The review was not prospectively registered.

Results: Thirty-nine unique studies were included; 33 studies (3,538 participants) contributed diagnostic data. Pooled sensitivity was 94.4% (95% CI 92.6-95.8%) and specificity was 93.1% (95% CI 89.7-95.4%). The positive likelihood ratio was 13.6 (95% CI 9.1-20.3), the negative likelihood ratio was 0.060 (95% CI 0.045-0.080), and the diagnostic odds ratio was 228 (95% CI 137-380). The HSROC area under the curve was 0.951. Heterogeneity was moderate for sensitivity (I^2=40.6%) and substantial for specificity (I^2=69.8%). Leave-one-out estimates were stable. Strategy studies consistently reduced nontherapeutic ERCP, but populations, thresholds, and resource settings differed.

Conclusions: EUS is an accurate triage test before ERCP, especially when the probability of persistent stones is uncertain. The evidence supports selective rather than universal EUS-first use: direct ERCP remains appropriate when urgent drainage or a definite therapeutic indication exists. The certainty of evidence was moderate for sensitivity, low for specificity, and low for comparative strategy outcomes.

Keywords: choledocholithiasis, endoscopic ultrasound, endoscopic retrograde cholangiopancreatography, diagnostic accuracy, triage, systematic review

Received: 05.04.2026
Accepted: 20.05.2026

INTRODUCTION

ERCP permits immediate extraction of CBD stones, but its value depends on

Copyright © Celsius Publishing House
www.sgo-iasgo.com

a therapeutic target. When no persistent stone is present, the patient accepts the risks of pancreatitis, bleeding, cholangitis, and perforation without receiving therapeutic benefit. Contemporary ASGE and ESGE guidance therefore separates diagnostic confirmation from treatment and recommends EUS or magnetic resonance cholangiopancreatography (MRCP) for many patients whose probability of choledocholithiasis is intermediate or otherwise uncertain (1-3).

EUS is particularly useful when small distal stones or sludge remain plausible after transabdominal imaging. It can also be followed by ERCP during the same sedation when a stone is confirmed. Earlier diagnostic studies reported high accuracy, whereas later studies increasingly evaluated EUS as a gatekeeper intended to reduce nontherapeutic ERCP (4-11). These two questions are related but not interchangeable: test accuracy does not by itself establish that an EUS-first strategy is superior in every risk category or clinical setting.

The previous manuscript combined a broad narrative review with a four-study meta-analysis, which left much of the available diagnostic evidence outside the quantitative synthesis. We therefore updated the search, reconstructed eligible 2 x 2 data, separated diagnostic accuracy from strategy outcomes, and assessed risk of bias and certainty of evidence. The primary objective was to estimate the sensitivity and specificity of EUS for suspected CBD stones. Secondary objectives were to examine heterogeneity, summarize EUS-first effects on ERCP use, complications, and costs, and place the findings alongside ASGE and ESGE pathways and settings where EUS is unavailable.

METHODS

Design, Reporting, and Protocol

This was an updated systematic review and diagnostic test accuracy meta-analysis reported with PRISMA 2020 and PRISMA-DTA (12,13). No protocol was prospectively registered in PROSPERO. The objectives, eligibility criteria, and analytic plan were specified before the update search, but the absence of prospective registration remains a limitation.

Information Sources and Search Strategy

The legacy evidence stream was anchored to the Cochrane diagnostic review by Giljaca et al., which searched MEDLINE, EMBASE, Science Citation Index Expanded, BIOSIS, the Cochrane Library, and ClinicalTrials.gov through September 2012 and contributed 13 EUS studies with 2 x 2 data (4). We reproduced those study-level counts from the published Cochrane forest plot rather than re-extracting unavailable older reports.

For the update, PubMed and Scopus were searched from 1 January 2013 through 21 March 2026 without language restriction. The complete queries are provided in Appendix 1. Reference lists of included studies, guidelines, and relevant reviews were checked for additional strategy studies. The searches returned 309 PubMed records and 629 Scopus records. After 263 duplicates were removed by PMID, DOI, or normalized title, 675 unique records remained for screening.

Eligibility Criteria

Diagnostic studies were eligible if they enrolled adults with suspected choledocholithiasis, evaluated EUS as the index test, used ERCP with stone extraction, operative duct exploration, or at least 3-6 months of clinical follow-up as a reference standard, and reported or permitted reconstruction of true-positive, false-positive, false-negative, and true-negative counts. Studies restricted to malignancy, therapeutic EUS, pancreatic disease, or prediction rules without EUS accuracy data were excluded. Strategy studies were eligible for a separate narrative synthesis if they compared or evaluated an EUS-first pathway and reported ERCP use, adverse events, hospital use, or costs.

Study Selection and Data Extraction

Titles and abstracts from the combined PubMed and Scopus update were screened against the eligibility criteria, followed by assessment of potentially relevant reports. One reviewer performed screening and extraction for this revision. Diagnostic counts were checked against reported sensitivity, specificity, predictive values, and total sample size; irreconcilable reports were not pooled. For several non-English Scopus records, reconstructable diagnostic counts were taken from detailed indexed English abstracts when an English full text was unavailable. Data extracted included year, design, clinical risk group, reference standard, 2 x 2 counts, ERCP avoidance, adverse events, and cost findings. Legacy counts were cross-checked against the Cochrane dataset (4).

Risk of Bias and Certainty of Evidence

Risk of bias was assessed with QUADAS-2 across patient selection, index test, reference standard, and flow/timing domains (14). Differential verification-

ERCP for positive EUS and follow-up for negative EUS-was considered a potential source of verification bias. Certainty was rated with GRADE principles for test accuracy, considering risk of bias, inconsistency, indirectness, imprecision, and small-study effects (15).

Statistical Analysis

Sensitivity and specificity were jointly pooled with a bivariate random-effects model on the logit scale (16). A 0.5 continuity correction was used for transformed-model estimation when a study contained a zero cell; observed study estimates remained uncorrected in the forest plot. From the pooled estimates we calculated positive and negative likelihood ratios and the diagnostic odds ratio. A hierarchical summary receiver-operating characteristic (HSROC) curve and area under the curve were derived from the bivariate model.

Heterogeneity was described with between-study variance and univariate I^2 values for logit sensitivity and specificity. Prespecified exploratory analyses compared legacy and post-2012 studies and repeated the model in studies that applied ERCP or operative verification to all participants. Leave-one-out analysis assessed influence. Other study-level meta-regression was not performed because reporting of stone size, operator experience, timing, and patient risk strata was too sparse and inconsistent. Deeks' effective-sample-size regression tested for small-study effects; its result was interpreted cautiously rather than as direct proof of publication bias (17).

RESULTS

Study Selection

The legacy Cochrane review identified 22,789 records, assessed 206 full texts, and contributed 13 EUS diagnostic studies. The database update identified 309 PubMed and 629 Scopus records; after removal of 263 duplicates, 675 unique records were screened. Six hundred thirty were excluded at title/abstract review and 45 reports were assessed for eligibility. Twenty-four update studies were included (20 diagnostic studies with reconstructable 2 x 2 data and four additional strategy-only studies). Two additional pre-2013 strategy studies were found by citation checking. The final review therefore comprised 39 unique studies; 33 entered the diagnostic meta-analysis and 12 informed the strategy synthesis, with overlap between those sets *(fig. 1)*.

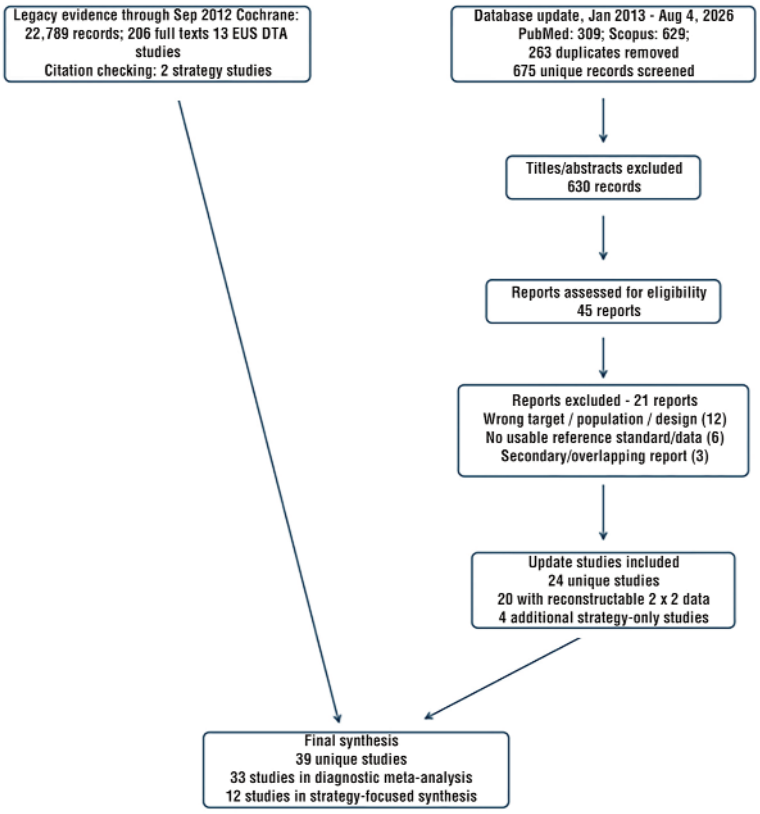


Figure 1 - Final PRISMA 2020 study-flow summary. DTA, diagnostic test accuracy.

Study Characteristics and Risk of Bias

The 33 diagnostic studies included 3,538 participants and spanned 1996-2026. Clinical settings ranged from unselected suspicion of CBD stones to intermediate-risk, high-risk, computed tomography-negative, and partial-obstruction cohorts. ERCP or operative verification was applied to every participant in 12 studies; the remainder used differential verification, usually therapeutic ERCP after positive EUS and clinical follow-up after negative EUS (*table 1*).

QUADAS-2 identified the reference-standard and flow/timing domains as the main weaknesses. Eighteen studies were rated high risk for the reference standard and 19 for flow/timing, largely because positive and negative index tests received different reference standards or because the timing between tests was incompletely reported. Patient selection was low risk in 25 studies. These limitations reduce confidence in specificity and in comparisons across risk categories (*table 2; fig. 2*).

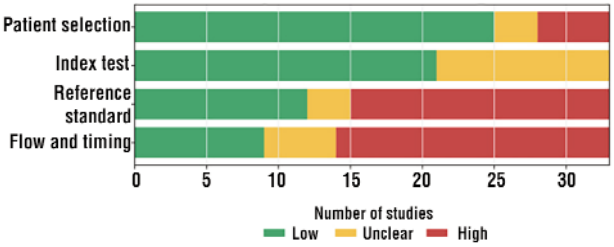


Figure 2 - Aggregate QUADAS-2 risk-of-bias summary across 33 diagnostic studies.

Legacy counts were reproduced from the Cochrane dataset (4). Update-study sources are references (9,18-

Table 1 - Studies included in the diagnostic meta-analysis and reconstructed 2 x 2 data

Study	Year	Era	TP	FP	FN	TN	N	Sensitivity	Specificity
Ang 2012	2012	Legacy	33	0	0	79	112	100.0%	100.0%
Norton 1997	1997	Legacy	21	1	3	25	50	87.5%	96.2%
Prat 1996	1996	Legacy	73	1	5	40	119	93.6%	97.6%
Fernandez-Esparrach 2007	2007	Legacy	59	11	2	63	135	96.7%	85.1%
Ney 2005	2005	Legacy	32	0	1	35	68	97.0%	100.0%
Choo 2012	2012	Legacy	3	0	1	5	9	75.0%	100.0%
Canto 1998	1998	Legacy	16	1	3	44	64	84.2%	97.8%
Montariol 1998	1998	Legacy	35	12	6	162	215	85.4%	93.1%
Fazel 2002	2002	Legacy	15	1	2	22	40	88.2%	95.7%
Kohut 2002	2002	Legacy	85	2	6	41	134	93.4%	95.3%
Liu 2001	2001	Legacy	33	1	1	65	100	97.1%	98.5%
Buscarini 2003	2003	Legacy	237	2	4	216	459	98.3%	99.1%
De Ledinghen 1999	1999	Legacy	10	1	0	21	32	100.0%	95.5%
Anderloni 2014	2014	Update	79	7	0	93	179	100.0%	93.0%
Epstein 2014	2014	Update	18	2	0	51	71	100.0%	96.2%
Sharma 2015	2015	Update	51	1	0	48	100	100.0%	98.0%
Netinatsunton 2016	2016	Update	83	2	2	7	94	97.6%	77.8%
Wang 2016	2016	Update	124	2	0	76	202	100.0%	97.4%
Makmun 2017	2017	Update	23	3	1	4	31	95.8%	57.1%
Jeon 2017	2017	Update	157	8	4	31	200	97.5%	79.5%
Patel 2017	2017	Update	31	1	2	36	70	93.9%	97.3%
Suzuki 2022	2022	Update	9	2	0	15	26	100.0%	88.2%
Jagtap 2022	2022	Update	44	1	2	65	112	95.7%	98.5%
Keczer 2022	2022	Update	53	0	1	41	95	98.1%	100.0%
Eissa 2022	2022	Update	66	9	1	14	90	98.5%	60.9%
Khan 2023	2023	Update	48	0	0	17	65	100.0%	100.0%
Bansal 2017	2017	Update	112	0	0	58	170	100.0%	100.0%
Cui 2019	2019	Update	87	5	3	3	98	96.7%	37.5%
Wang CY 2019	2019	Update	20	2	3	53	78	87.0%	96.4%
Wang XL 2021	2021	Update	83	1	4	44	132	95.4%	97.8%
Kamalaporn 2021	2021	Update	37	1	4	38	80	90.2%	97.4%
Li 2022	2022	Update	43	2	3	12	60	93.5%	85.7%
Sugiura 2026	2026	Update	19	2	0	27	48	100.0%	93.1%

Table 2 - Study-level QUADAS-2 risk-of-bias judgments. Judgments reflect the available reports; legacy ratings were cross-checked against the Cochrane review (4).

Study	Patient selection	Index test	Reference standard	Flow/timing
Ang 2012	Low	Unclear	High	High
Norton 1997	High	Unclear	High	High
Prat 1996	High	Low	Low	High
Fernandez-Esparrach 2007	Low	Unclear	High	High
Ney 2005	Low	Unclear	High	High
Choo 2012	Low	Low	Low	High
Canto 1998	Low	Low	High	High
Montariol 1998	Low	Unclear	High	High
Fazel 2002	High	Unclear	Low	High
Kohut 2002	High	Unclear	Low	High
Liu 2001	Low	Unclear	High	High
Buscarini 2003	Low	Low	High	High
De Ledinghen 1999	High	Low	High	High
Anderloni 2014	Low	Unclear	High	High
Epstein 2014	Low	Unclear	High	Low
Sharma 2015	Low	Low	Low	Low
Netinatsunton 2016	Low	Low	High	Unclear
Wang 2016	Low	Unclear	High	Unclear
Makmun 2017	Unclear	Low	Low	Low
Jeon 2017	Low	Low	Low	Low
Patel 2017	Low	Low	High	High
Suzuki 2022	Low	Low	Unclear	Low
Jagtap 2022	Low	Low	Unclear	Low
Keczer 2022	Low	Low	Unclear	Low
Eissa 2022	Low	Low	Low	Low
Khan 2023	Low	Low	Low	Low
Bansal 2017	Low	Unclear	High	High
Cui 2019	Unclear	Low	Low	Unclear
Wang CY 2019	Low	Low	High	High
Wang XL 2021	Low	Low	High	High
Kamalaporn 2021	Low	Low	Low	Unclear
Li 2022	Unclear	Low	Low	Unclear
Sugiura 2026	Low	Low	High	High

29,33-39). TP, true positive; FP, false positive; FN, false negative; TN, true negative.

Individual study estimates are shown in *fig. 3*; the HSROC curve is shown in *fig. 4*.

Diagnostic Accuracy

The pooled sensitivity of EUS was 94.4% (95% CI 92.6-95.8%) and pooled specificity was 93.1% (95% CI 89.7-95.4%). The positive likelihood ratio was 13.6 (95% CI 9.1-20.3), the negative likelihood ratio was 0.060 (95% CI 0.045-0.080), and the diagnostic odds ratio was 228 (95% CI 137-380). The HSROC area under the curve was 0.951.

Heterogeneity, Sensitivity Analyses, and Small-Study Effects

Heterogeneity was moderate for sensitivity (I^2=40.6%; Q=53.9, 32 df; p=0.009) and substantial for specificity (I^2=69.8%; Q=106.1, 32 df; p<0.001). Lower specificity in Cui 2019, Makmun 2017, and Eissa 2022 contributed to the dispersion, but no single study determined the summary result. Across leave-one-out analyses, pooled sensitivity ranged from 93.7% to 94.7% and specificity from 92.4% to 93.6%.

In the exploratory era analysis, pooled sensitivity was 92.4% (95% CI 88.3-95.1%) in legacy studies and 95.3% (95% CI 93.6-96.6%) in update studies (p=0.076 for the logit-scale difference). Specificity was 95.2% (95% CI 91.9-97.2%) and 91.0% (95% CI 85.0-94.7%), respectively (p=0.10). The higher sensitivity in later studies may reflect patient spectrum, verification, equipment, or operator factors and should not be interpreted as a temporal causal effect. Restricting analysis to 12 studies with complete ERCP or operative verification yielded sensitivity 93.9% and specificity 86.5%, supporting the concern that differential

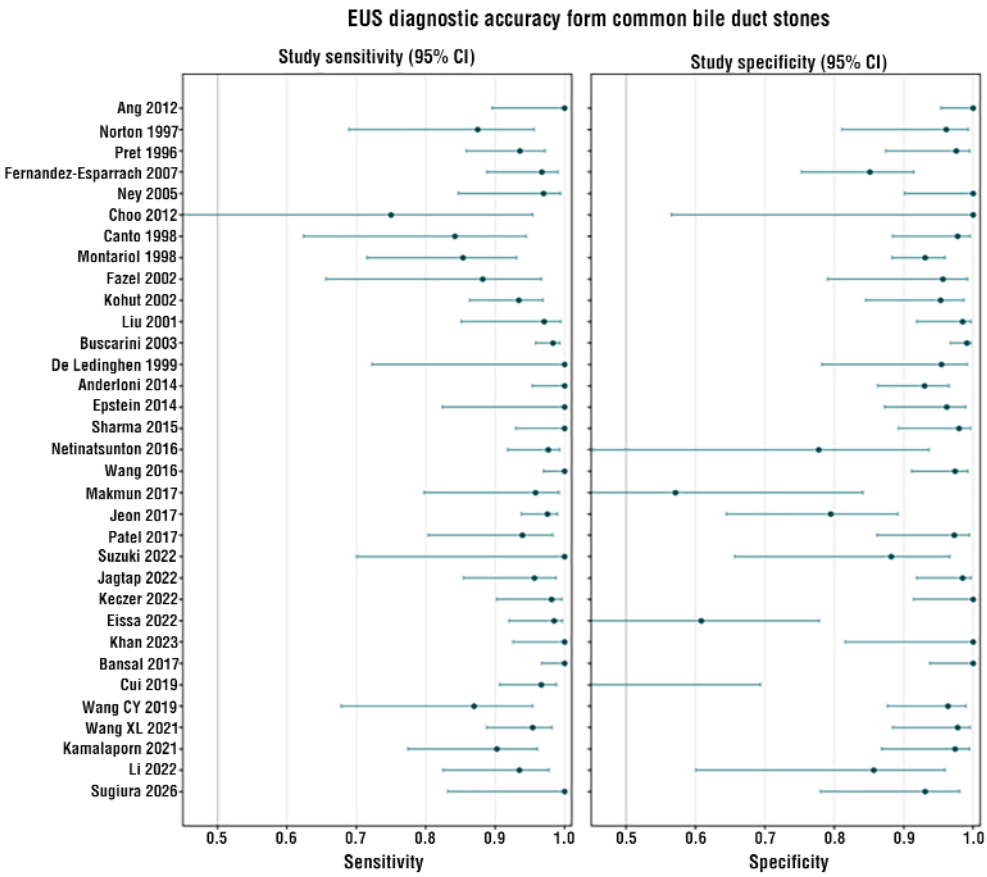


Figure 3 - Study-specific sensitivity and specificity with 95% Wilson confidence intervals.

verification may inflate specificity. Deeks' regression suggested small-study effects (p<0.001). This may reflect publication bias, selective thresholds, or design-related differences between small and large cohorts; it was considered when grading certainty.

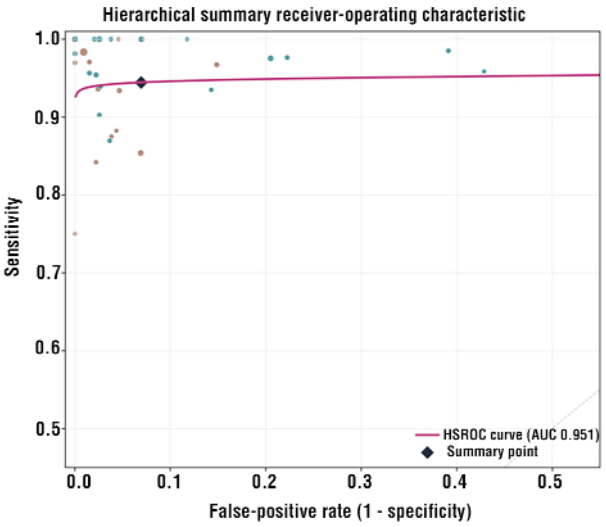


Figure 4 - Hierarchical summary receiver-operating characteristic curve. Point size reflects study sample size; brown points denote legacy studies and teal points the update.

Clinical Strategy Outcomes

Twelve studies addressed EUS as a clinical gatekeeper rather than only as an index test. Across intermediate-risk cohorts, EUS-first pathways reduced diagnostic ERCP without evidence of excess short-term biliary events. Same-session EUS followed by ERCP shortened procedure pathways and reduced hospitalization or costs in selected settings (6-10,33,39). In high-probability but computed tomography-negative patients, Lin et al. prevented ERCP in 27 of 44 patients and reported fewer overall complications with EUS-first than ERCP-first (6.8% vs 21.7%) (10). Patel et al. avoided ERCP in 57.7% of a selected high-likelihood cohort (24). Bansal et al. avoided ERCP in 58 of 170 intermediate- or high-risk patients and reported 6.1% lower pathway costs (33). Sugiura et al. reported that EUS avoided unnecessary ERCP in 79.3% of patients without a final stone diagnosis (39). In the 2026 cohort

by Chiriac et al., only 53.5% of guideline-defined high-risk patients had stones on EUS, although that observational result does not establish that all high-risk patients should avoid direct ERCP (11).

The strategy studies are summarized in *table 3*.

DISCUSSION

This updated analysis supports EUS as an accurate test for deciding whether ERCP is likely to be therapeutic. The 33-study synthesis produced high pooled sensitivity and specificity, strong likelihood ratios, and stable leave-one-out results. It does not, however, prove that EUS-first is superior to ERCP-first in every patient. The strongest clinical case remains the patient with an uncertain or intermediate probability of a persistent CBD stone, in whom a negative EUS can prevent an invasive procedure.

The revised estimates are less extreme than those from the former four-study subset. Expanding the quantitative evidence reduced the risk that pooled results reflected only easily accessible, high-performing reports. It also exposed clinically important heterogeneity. Specificity varied more than sensitivity, particularly in small or highly selected cohorts. Differential verification was common and can favor the index test because EUS-positive patients undergo ERCP while EUS-negative patients are observed. The lower specificity in the complete-verification sensitivity analysis reinforces the need for caution.

The evidence is broadly consistent with ASGE and ESGE guidance but should not be used to erase their risk-based distinctions (1,2). Both guidelines support EUS or MRCP for intermediate or indeterminate probability. Direct ERCP remains appropriate when imaging already shows a CBD stone, when cholangitis requires urgent drainage, or when another definite therapeutic indication is present. In selected high-risk patients without cholangitis or a definite stone-particularly when laboratory values are improving or prior imaging is negative-EUS may still prevent nontherapeutic ERCP, but this is a selective extension rather than a universal rule.

Availability changes the preferred pathway. Where EUS and appropriately trained endosonographers are available, same-session EUS-ERCP can minimize repeated sedation and delay. Where EUS is unavailable, MRCP is a guideline-concordant noninvasive alternative; intraoperative cholangiography may be useful when cholecystectomy is planned. Referral to a center capable of same-session EUS-ERCP is reasonable when local MRCP is nondiagnostic and the

Table 3 - Selected EUS-first strategy studies

Study	Design	Population	Main strategy finding
Canto 1998 (6)	Prospective	Suspected stones	Comparable diagnostic performance; fewer complications and additional diagnoses with EUS.
Karakan 2009 (7)	Randomized	Intermediate probability	EUS detected small stones and did not increase downstream procedures.
Fabbri 2009 (8)	Randomized	Low/moderate risk	Negative EUS avoided ERCP in 33/80; same-session care reduced time and cost.
Netinatsunton 2016 (9)	Prospective economic	Intermediate/high risk	EUS was safer and less costly at intermediate prevalence; direct ERCP study remained economical in high prevalence.
Patel 2017 (24)	Prospective	High likelihood	EUS-first avoided ERCP in 57.7%; two false negatives occurred during follow-up.
Maruta 2019 (30)	Observational	Suspected stones	EUS-first reduced early ERCP-related adverse events without a clear increase in late events.
Lin 2022 (10)	Prospective comparative	High probability, CT negative	ERCP avoided in 27/44; complications 6.8% vs 21.7% with direct ERCP.
Jagtap 2022 (26)	Randomized	Intermediate likelihood	EUS and MRCP had comparable accuracy; choice depended on expertise, access, and preference.
Varone 2025 (31)	Multicenter observational	Intermediate risk	EUS and intraoperative cholangiography outperformed MRCP in the reported cohort.
Chiriac 2026 (11)	Retrospective	Mixed, including high risk	EUS confirmed stones in 53.5% of high-risk patients; 46.4% avoided ERCP.
Bansal 2017 (33)	Retrospective	Intermediate/high risk	EUS-first avoided ERCP in 58/170 and reduced estimated pathway costs by 6.1%.
Sugiura 2026 (39)	Prospective multicenter	High likelihood, initial imaging negative	EUS avoided unnecessary ERCP in 79.3% of patients without a final stone diagnosis.

Table 4 - GRADE certainty profile

Outcome	Evidence	Main concerns	Estimate	Certainty
Sensitivity	33 DTA studies	Serious risk of bias; small-study effects	94.4% (92.6-95.8%)	Moderate
Specificity	33 DTA studies	Serious risk of bias and inconsistency; small-study effects	93.1% (89.7-95.4%)	Low
Likelihood ratios / DOR	33 DTA studies	Same limitations as sensitivity/specificity	LR+ 13.6; LR- 0.060; DOR 228	Low
Avoidance of nontherapeutic ERCP	12 strategy studies	Heterogeneous thresholds; mostly single-center/nonrandomized	Direction consistently favors selective triage	Low
Complications and costs	Subset of strategy studies	Imprecision, local costing, and selective populations	Lower or similar with EUS-first in selected settings	Low

probability of stones remains clinically important. Economic comparisons must be localized because procedure costs, anesthesia models, hospital stay, and test availability differ substantially (9,32).

GRADE certainty was moderate for sensitivity and low for specificity and strategy outcomes. High accuracy estimates should therefore be used to support a clinical decision, not to replace assessment of cholangitis, stone visualization, procedural urgency, local expertise, or patient preference. The significant Deeks test also cautions against assuming that the published evidence is free from small-study or selective-reporting effects.

The GRADE certainty profile is shown in *table 4*.

Limitations

This review was not prospectively registered. The post-2012 update searched PubMed and Scopus but not EMBASE, Web of Science, or regional databases, so coverage remains incomplete. Screening and extraction for the revision were completed by one reviewer rather than two independent reviewers. For several non-English Scopus records, diagnostic counts were reconstructed from detailed indexed English abstracts because an English full text was unavailable, limiting risk-of-bias assessment. Legacy 2 x 2 data were reproduced from the Cochrane review rather than re-extracted from every primary report. Several included studies used differential verification, some update reports permitted reconstruction only after excluding unevaluable participants, and thresholds and follow-up periods varied. Reporting was insufficient for multivariable meta-regression by stone size, operator expertise, same-session practice, or formal guideline risk group. Finally, diagnostic accuracy cannot by itself establish comparative patient benefit.

CONCLUSION

EUS is an accurate triage test before ERCP for suspected CBD stones, with pooled sensitivity of 94.4% and specificity of 93.1%. The evidence supports EUS-first most clearly when a therapeutic target is uncertain and avoiding nontherapeutic ERCP matters. It does not justify delaying urgent drainage or replacing direct therapeutic ERCP when a stone or another intervention-requiring lesion is already established. In centers without EUS, MRCP, intraoperative cholangiography, or referral should be chosen according to clinical urgency, local expertise, and resource availability.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

No specific funding was received for this work.

Data Availability

The reconstructed diagnostic 2 x 2 data are reported in *table 1*. The exact PubMed and Scopus update queries and search counts are reported in *appendix 1*.

REFERENCES

1. Buxbaum JL, Abbas Fehmi SM, Sultan S, Fishman DS, Qumseya BJ, Cortessis VK, et al. ASGE guideline on the role of endoscopy in the evaluation and management of choledocholithiasis. Gastrointest Endosc. 2019;89(6):1075-1105.e15.
2. Manes G, Paspatis G, Aabakken L, Anderloni A, Arvanitakis M, Ah-Soune P, et al. Endoscopic management of common bile duct stones: European Society of Gastrointestinal Endoscopy (ESGE) guideline. Endoscopy. 2019;51(5):472-491.
3. Anderson MA, Fisher L, Jain R, Evans JA, Appalaneni V, Ben-Menachem T, et al. Complications of ERCP. Gastrointest Endosc. 2012;75(3):467-73.
4. Giljaca V, Gurusamy KS, Takwoingi Y, Higgie D, Poropat G, Štimac

D, et al. Endoscopic ultrasound versus magnetic resonance cholangiopancreatography for common bile duct stones. Cochrane Database Syst Rev. 2015;2015(2): CD011549.
5. Meeralam Y, Al-Shammari K, Yaghoobi M. Diagnostic accuracy of EUS compared with MRCP in detecting choledocholithiasis: a meta-analysis of diagnostic test accuracy in head-to-head studies. Gastrointest Endosc. 2017;86(6):986-993.
6. Canto MI, Chak A, Stellato T, Sivak MV Jr. Endoscopic ultrasonography versus cholangiography for the diagnosis of choledocholithiasis. Gastrointest Endosc. 1998;47(6):439-48.
7. Karakan T, Cindoruk M, Alagozlu H, Ergun M, Dumlu S, Unal S. EUS versus endoscopic retrograde cholangiography for patients with intermediate probability of bile duct stones: a prospective randomized trial. Gastrointest Endosc. 2009;69(2):244-52.
8. Fabbri C, Polifemo AM, Luigiano C, Cennamo V, Fuccio L, Billi P, et al. Single session versus separate session endoscopic ultrasonography plus endoscopic retrograde cholangiography in patients with low to moderate risk for choledocholithiasis. J Gastroenterol Hepatol. 2009;24(6):1107-12.
9. Netinatsunton N, Attasaranya S, Sottisuporn J, Witeerungrot T, Jongboonyanuparp T, Piratvisuth T, et al. Comparing cost-effectiveness between endoscopic ultrasound and endoscopic retrograde cholangiopancreatography in diagnosis of common bile duct stone in patients with predefined risks. Endosc Ultrasound. 2016;5(3): 165-72.
10. Lin MY, Lee CT, Hsieh MT, Ou MC, Wang YS, Lee MC, et al. Endoscopic ultrasound avoids adverse events in high probability choledocholithiasis patients with a negative computed tomography. BMC Gastroenterol. 2022;22(1):94.
11. Chiriac S, Sfarti C, Minea H, Zenovia S, Girleanu I, Huiban L, et al. Endoscopic ultrasound and endoscopic retrograde cholangiopancreatography for bile duct stones-avoiding the avoidable. Biomedicines. 2026;14(1):91.
12. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
13. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM; and the PRISMA-DTA Group; et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: the PRISMA-DTA statement. JAMA. 2018;319(4):388-396.
14. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155(8):529-36.
15. Hultcrantz M, Mustafa RA, Leeflang MMG, Lavergne V, Estrada-Orozco K, Ansari MT, et al. Defining ranges for certainty ratings of diagnostic accuracy: a GRADE concept paper. J Clin Epidemiol. 2020;117:138-148.
16. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982-90.
17. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005; 58(9):882-93.
18. Anderloni A, Ballare M, Pagliarulo M, Conte D, Galeazzi M, Orsello M, et al. Prospective evaluation of early endoscopic ultrasonography for triage in suspected choledocholithiasis: results from a large single centre series. Dig Liver Dis. 2014;46(4):335-9.
19. Epstein AM, Duberman BL, Dynkov SM, Pozdeev VN. Endosonography in diagnosis of choledocholithiasis. Eksp Klin Gastroenterol. 2014;(10):33-37. Russian
20. Sharma R, Menachery J, Choudhary NS, Kumar M, Puri R, Sud R. Routine endoscopic ultrasound in moderate and indeterminate risk patients of suspected choledocholithiasis to avoid unwarranted ERCP: a prospective randomized blinded study. Indian J Gastroenterol. 2015;34(4):300-4.
21. Wang M, He X, Tian C, Li J, Min F, Li HY. The diagnostic accuracy of linear endoscopic ultrasound for evaluating symptoms suggestive of common bile duct stones. Gastroenterol Res Pract. 2016;2016: 6957235.
22. Makmun D, Fauzi A, Shatri H. Sensitivity and specificity of magnetic resonance cholangiopancreatography versus endoscopic ultrasonography against ERCP in diagnosing choledocholithiasis: the indonesian experience. Clin Endosc. 2017;50(5):486-490.
23. Jeon TJ, Cho JH, Kim YS, Song SY, Park JY. Diagnostic value of endoscopic ultrasonography in symptomatic patients with high and intermediate probabilities of common bile duct stones and a negative computed tomography scan. Gut Liver. 2017;11(2):290-297.
24. Patel R, Ingle M, Choksi D, Poddar P, Pandey V, Sawant P. Endoscopic ultrasonography can prevent unnecessary diagnostic ERCP even in patients with high likelihood of choledocholithiasis and inconclusive ultrasonography: results of a prospective study. Clin Endosc. 2017;50(6):592-597.
25. Suzuki M, Sekino Y, Hosono K, Yamamoto K, Kawana K, Nagase H, et al. Endoscopic ultrasound versus magnetic resonance cholangiopancreatography for the diagnosis of computed tomography-negative common bile duct stone: prospective randomized controlled trial. Dig Endosc. 2022;34(5):1052-1059.
26. Jagtap N, Kumar JK, Chavan R, Basha J, Tandan M, Lakhtakia S, et al. EUS versus MRCP to perform ERCP in patients with intermediate likelihood of choledocholithiasis: a randomised controlled trial. Gut. 2022;71(10):2005-2010.
27. Keczer B, Dubravcsik Z, Szepes A, Madacsy L, Szijarto A, Hritz I. Diagnostic sensitivity of endoscopic ultrasonography in patients with suspected choledocholithiasis. Orv Hetil. 2022;163(10):400-406. Hungarian
28. Eissa M, Okasha HH, Abbasy M, Khamis AK, Abdellatef A, Rady MA. Role of endoscopic ultrasound in evaluation of patients with missed common bile duct stones. World J Gastrointest Endosc. 2022; 14(9):564-574.
29. Khan RSA, Alam L, Khan ZA, Khan UA. Comparing the efficacy of EUS versus MRCP with ERCP as gold standard in patients presenting with partial biliary obstruction-finding a better diagnostic tool. Pak J Med Sci. 2023;39(5):1275-1279.
30. Maruta A, Iwashita T, Uemura S, et al. Efficacy of the endoscopic ultrasound-first approach in patients with suspected common bile duct stone to avoid unnecessary Endoscopic Retrograde Cholangiopancreatography. Intern Med. 2019;58(12):1673-1679.
31. Varone A, Mendoza A, Gelbard R, Tracy B, Poulose B. Suspected common bile duct stones: which test is best? Surg Endosc. 2025; 39(10):6525-6529.
32. Morris S, Gurusamy KS, Sheringham J, Davidson BR. Cost-effectiveness analysis of endoscopic ultrasound versus magnetic resonance cholangiopancreatography in patients with suspected common bile duct stones. PLoS One. 2015;10(3):e0121699.
33. Bansal NK, Panda N, Alagammai PL, Narsimhan M, Ardhanari R. Benefits of endoscopic ultrasonography-first strategy in patients with intermediate and high risk for choledocholithiasis. Acta Gastroenterol Latinoam 2017;47(4):252-258
34. Cui Y, Bao J, Song Y, Zhang L, Mei Q. Clinical effect of endoscopic ultrasonography in the diagnosis and evaluation of common bile duct stones. J Clin Hepatol. 2019;35:2741-2743.
35. Wang C, Xu F, Dai C. Value of endoscopic ultrasonography combined with magnetic resonance cholangiopancreatography in patients suspected of common bile duct stones. J Clin Hepatol. 2019; 35(1):127-130.
36. Wang X, Wang Z. Clinical value of endoscopic ultrasonography before endoscopic retrograde cholangiopancreatography for undetermined common bile duct stones. Chin J Dig Endosc. 2021; 38:556-559.
37. Kamalaporn P, Chirnaksorn S, Rattanasiri S, Kitiyakara T. Comparison of radial echoendoscopy and predictive factors in the evaluation of patients with suspected choledocholithiasis. Siriraj Med J. 2021;73:380-385.
38. Li X, Yang T, Xu T, Sun R. Diagnostic value of endoscopic ultrasonography for small common bile duct stones. Chin J Dig Endosc. 2022;39:1018-1021.
39. Choi YH, Lee YS, Lee SH, Son JH, Ryu JK, Kim YT, et al. Role of EUS at high risk for choledocholithiasis without severe cholangitis and visible stone on crosssectional imaging: a multicenter randomized clinical trial. Endosc Ultrasound. 2021;10(6):455462.

Supplementary Material

Appendix 1 - Search strategy and audit

PubMed and Scopus search date: 21 March 2026

Exact PubMed query: ("Choledocholithiasis"(Mesh) OR choledocholithiasis(tiab) OR "common bile duct stone"(tiab) OR "common bile duct stones"(tiab)) AND ("Endosonography"(Mesh) OR "endoscopic ultrasound"(tiab) OR "endoscopic ultrasonography"(tiab) OR EUS(tiab)) AND (diagnos*(tiab) OR sensitiv*(tiab) OR specific*(tiab) OR accuracy(tiab) OR ERCP(tiab) OR "Cholangiopancreatography, Endoscopic Retrograde"(Mesh)) AND ("2013/01/01"(Date - Publication) : "2026/03/21" (Date - Publication))

Exact Scopus query: TITLE-ABS-KEY((choledocholithiasis OR "common bile duct stone" OR "common bile duct stones") AND ("endoscopic ultrasound" OR "endoscopic ultrasonography" OR EUS) AND (diagnos* OR sensitiv* OR specific* OR accuracy OR ERCP)) AND PUBYEAR > 2012 AND PUBYEAR < 2027

Source	Coverage	Records	Reports assessed	Included
Legacy Cochrane review (4)	Through Sep 2012	22,789	206	13 EUS DTA studies
PubMed update	1 Jan 2013-21 Mar 2026	309	33	18 unique studies
Scopus-only after deduplication	1 Jan 2013-21 Mar 2026	366	12	6 additional unique studies
Combined database update	After 263 duplicates removed	675	45	24 unique studies
Citation checking	Reference lists/guidelines	Not separately counted	2	2 pre-2013 strategy studies