Am J Transl Res 2026;18(10):9364-9375
www.ajtr.org /ISSN:1943-8141/AJTR0174163

Original Article

Diagnostic performance of non-contrast MRI combined with magnetic resonance cholangiopancreatography for cholelithiasis and severity of biliary obstruction

Biao Yang[1], Sheng Yu[1], Jing Guo[2], Pinglin Fan[1]

[1]Department of Imaging, Yingshang County People's Hospital, Fuyang, Anhui, China; [2]Department of Imaging, Lixin County People's Hospital, Bozhou, Anhui, China

Received May 19, 2026; Accepted August 19, 2026; Epub October 15, 2026; Published October 30, 2026

Abstract: Objective: To evaluate the diagnostic performance of non-contrast MRI combined with magnetic resonance cholangiopancreatography (MRCP) for cholelithiasis and to assess its utility in grading biliary obstruction severity according to the Bismuth classification. Methods: This retrospective study included 105 patients with suspected cholelithiasis who underwent both non-contrast MRI with MRCP and had surgical confirmation. The surgical outcomes were used as the gold standard to compare the diagnostic accuracy of non-contrast MRI, MRCP, and combined interpretation for gallstone type classification and Bismuth grading. The consistency between each method and the above standard was evaluated using Cohen's kappa coefficient. ROC curve analysis was used to assess the diagnostic efficacy, and the Z-test was employed for pairwise comparison of the area under the curve (AUC). Result: Using surgical findings as the reference, non-contrast MRI showed a stone type classification accuracy of 78.65%, a Bismuth grading accuracy of 80.88%, and an overall accuracy of 76.19% (80/105), with poor agreement ($\kappa = 0.309$, $P < 0.001$). MRCP achieved accuracies of 85.39% for stone type and 89.71% for Bismuth grading, and the overall accuracy was 83.81% (88/105), with a moderate consistency ($\kappa = 0.491$, $P < 0.001$). The combined MRI + MRCP strategy increased these accuracies to 96.63% and 97.06%, respectively, and the overall accuracy reached 96.19% (101/105), with excellent agreement ($\kappa = 0.860$, $P < 0.001$). ROC analysis showed that the AUC of non-contrast MRI was 0.706, that of MRCP was 0.801, and that of combined diagnosis was 0.950. The AUC of combined diagnosis was significantly higher than that of non-contrast MRI ($Z = 3.240$, $P = 0.001$) and MRCP ($Z = 2.252$, $P = 0.024$). Stratified analysis showed that the combined diagnosis maintained a sensitivity ≥ 0.889 and specificity ≥ 0.800 across all clinical subgroups. Notably, in diabetic patients, specificity improved from 0.400 with non-contrast MRI to 0.800 with the combined approach. The ALT level in the incorrectly diagnosed group was significantly lower than that of the correctly diagnosed group ($P = 0.047$). Conclusion: For patients with suspected biliary stones referred for MRI, non-contrast MRI combined with MRCP significantly improved diagnostic accuracy for stone detection and Bismuth grading compared to either method alone. However, the net reclassification improvement over single-method interpretation was not statistically significant (NRI = -0.133 and -0.075, both $P > 0.05$), indicating that the benefit at the individual patient level remains uncertain. The clinical use of this combined strategy warrants further validation in larger prospective studies.

Keywords: Non-contrast MRI, MRCP, cholelithiasis, biliary obstruction, bismuth classification, diagnostic accuracy

Introduction

Cholelithiasis is one of the most common biliary tract disorders in clinical practice and a leading cause of biliary obstruction. Studies have shown that biliary calculi not only directly cause biliary damage but also are closely related to the occurrence of biliary tumors, inflammatory biliary stenosis, and other diseases. If not diagnosed and managed promptly, cholelithiasis can lead to severe complications, including acute cholangitis, pancreatitis, obstructive jaundice, and even liver failure [1, 2]. Therefore, early accurate identification of the presence of calculi, clear determination of the type of calculi, and assessment of the severity of biliary obstruction are of great significance for formulating reasonable treatment plans and improving prognosis.

Currently, the imaging methods used for biliary calculi assessment in clinical practice mainly

https://doi.org/10.62347/XMFD5379

include ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI). Ultrasound examination is simple to operate, non-invasive, and does not involve ionizing radiation and is often used as the first-line screening method for biliary diseases, with high sensitivity for gallbladder stones and intrahepatic bile duct stones. However, the diagnostic efficacy of ultrasound for common bile duct stones is substantially affected by bowel gas interference and operator experience, leading to a higher rate of false-negative or false-positive results [3, 4]. CT offers high spatial and contrast resolution, clearly shows high-density calcified stones, has a fast scan speed and can comprehensively evaluate the biliary system and surrounding tissue structures; however, CT has a relatively high miss rate for stones that are isodense or hypodense [5, 6]. In addition, the ionizing radiation associated with CT and the nephrotoxicity and allergic risks of iodinated contrast agents limit its repeated use in certain patients [7-9].

MRI has excellent soft tissue contrast and multi-planar imaging capabilities, and no ionizing radiation, so it can clearly display the anatomic structure of the biliary system and the pathologic changes of surrounding tissues. Non-contrast MRI plain scan can simultaneously provide multi-parameter information such as T1WI, T2WI, and diffusion-weighted imaging, which is helpful for comprehensive assessment of the lesion range and its relationship to adjacent tissues [10-13]. Magnetic resonance cholangiopancreatography (MRCP) is a non-invasive MRI technique that uses heavily T2-weighted sequences without contrast administration, first described by Waller et al. in 1991, in which static or slow-flowing fluids (e.g., bile and pancreatic juice) appear hyperintense, while solid components such as stones present as hypointense filling defects. MRCP can non-invasively display the entire biliary tree, helping to clearly identify the location, size, and number of calculi, as well as the degree of biliary dilation, providing important imaging evidence for determining the cause and level of obstruction [14, 15].

In current clinical practice, the combined application of non-contrast MRI and MRCP has not yet become a routine strategy for biliary calculi assessment. On the one hand, the sensitivity of non-contrast MRI plain scan may be insufficient in showing less calcified stones, and there are limitations in the assessment of the internal characteristics of the stones [16]; On the other hand, although MRCP can clearly display the morphology of the bile duct lumen, it provides limited information about the bile duct wall and the surrounding soft tissues, and may lead to misjudgment when the bile duct is not dilated. Artifacts related to the imaging and partial volume effects may also affect the detection of small lesions [17]. Theoretically, combining non-contrast MRI with MRCP can leverage the strengths of both techniques: non-contrast MRI provides comprehensive anatomic background and surrounding soft-tissue information, while MRCP details intraluminal morphology. This complementary information is expected to improve the detection rate of bile duct stones and the accuracy of assessing the degree of obstruction, providing a more comprehensive diagnostic basis for clinical practice.

However, systematic retrospective validation is lacking regarding the incremental diagnostic value of non-contrast MRI combined with MRCP compared to a single method, as well as the correlation between the combined diagnosis and the severity of biliary obstruction (Bismuth classification). Based on the above background, this study conducted a retrospective analysis of 105 patients with suspected cholelithiasis confirmed by surgery, aiming to: (1) Evaluate the diagnostic accuracy and classification ability of non-contrast MRI combined with MRCP for bile duct stones; (2) Explore the correlation between the combined method and the Bismuth classification in assessing the severity of obstruction; (3) Compare the differences in diagnostic efficacy between the combined method and the single method, providing evidence-based support for clinical application. The results of the study are presented below.

Materials and methods

Research design and research subjects

This was a retrospective cross-sectional study. We consecutively included patients with suspected cholelithiasis who underwent both non-contrast MRI and MRCP at the radiology departments of Yingshang County People's Hospital between April 2023 and May 2025. The imaging data, clinical baseline information (demographic characteristics, past medical history,

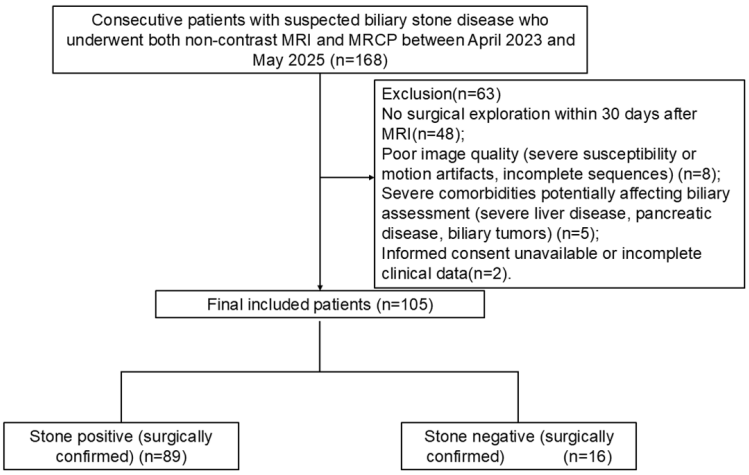


Figure 1. Flow diagram. Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

clinical symptoms, laboratory test results), and surgical records of all patients were retrospectively extracted from the hospital's electronic medical record system and the Picture Archiving and Communication System (PACS). The study protocol was approved by the Ethics Committee of Yingshang County People's Hospital. Due to the retrospective nature of the study, the ethics committee waived the requirement for written informed consent. Patient data were de-identified prior to analysis.

Inclusion and exclusion criteria

Inclusion criteria: (1) Patients with suspected biliary disease presenting with upper abdominal discomfort, fever, jaundice, and/or laboratory abnormalities (ALT > 40 U/L, total bilirubin > 1.2 mg/dL [> 20.5 μmol/L], or WBC > 10 × 10^9/L), who underwent abdominal MRI; (2) Completed non-contrast MRI plain scan and MRCP examination simultaneously; (3) Underwent surgery within 30 days after MRI examination and could obtain clear intraoperative exploration results as the gold standard; (4) Provided informed consent.

Exclusion criteria: (1) Had metal foreign bodies or implanted prostheses that affected MRI quality; (2) Had contraindications for MRI examination, such as cardiac pacemakers, suspected intraocular metal foreign bodies, intracranial or vascular metal clips; (3) Poor MRI image quality (e.g., severe susceptibility or motion artifacts) precluding diagnostic interpretation; (4) Had severe accompanying diseases that might affect the biliary system or interfere with image interpretation, including severe liver diseases, pancreatic lesions, biliary tumors, or other serious comorbidities.

Based on the above criteria, a total of 105 cases were retrospectively screened and reviewed for imaging data [details of the initial screening cases to be supplemented]. After excluding those who did not meet the conditions, 105 patients were finally included for analysis. The patient screening process is shown in **Figure 1**.

MRI equipment and patient preparation

All MRI examinations used a 3.0T superconducting magnetic resonance system (uMR780; United Imaging Healthcare, Shanghai, China). Patients were fasting and dehydrated for 4-6 hours before the examination to reduce the interference of gastrointestinal contents on image quality.

Non-contrast MRI plain scan protocol

Non-contrast MRI plain scan collected coronal and axial images simultaneously. The scanning sequences and parameters were as follows:

T1-weighted imaging (T1WI): Three-dimensional volume interpolated rapid gradient echo sequence; repetition time (TR) = 2.92 ms, echo time (TE) = 1.17 ms, slice thickness = 4.5 mm, field of view (FOV) = 400 × 300 mm, matrix = 288 × 230.

T2-weighted imaging (T2WI): Rapid spin echo sequence; TR = 4190 ms, TE = 94.4 ms, slice thickness = 6 mm, FOV = 380 × 380 mm, matrix = 320 × 320.

Coronal T2WI: Single-shot excited rapid spin echo sequence; slice thickness = 6 mm, TR = 1100 ms, TE = 92.16 ms.

Diffusion-weighted imaging (DWI): Acquired DWI images and generated apparent diffusion coefficient (ADC) maps.

The scanning range was from the diaphragm top to the lower edge of the kidneys, covering

the gallbladder, intrahepatic bile ducts, and extrahepatic bile ducts.

MRCP scanning protocol and post-processing

MRCP used two acquisition techniques: 2D MRCP and 3D MRCP.

2D MRCP: Single-shot excited imaging sequence; slice thickness = 60 mm, TR = 6000 ms, TE = 810 ms, flip angle = 8°.

3D MRCP: Coronal acquisition, instructing the patient to hold their breath (approximately 18 seconds). Using a rapid spin echo sequence, with FOV = 352 × 352 mm, TR = 4000 ms, and TE = 696 ms. A rapid spin-gradient echo sequence (FOV = 352 × 352 mm) was also acquired during an 18-second breath-hold, with TR = 1100 ms and TE = 249 ms.

All the original images cover the gallbladder and the intrahepatic and extrahepatic bile ducts. Maximum intensity projection (MIP) was used for three-dimensional reconstruction to generate a three-dimensional bile duct tree image for the diagnostician to observe from multiple angles.

Image retrospective interpretation process

All image data were independently and retrospectively interpreted by two senior radiologists (each with > 5 years of experience in abdominal imaging) on a PACS workstation. The interpreters were blinded to the surgical findings. The interpretation was conducted at three levels:

The first level - Interpretation of non-contrast MRI alone: Judging the presence of biliary stones, their location and type, as well as the presence of biliary obstruction and assessing the severity of the obstruction based solely on axial and coronal T1WI, T2WI, and DWI/ADC images.

The second level - Interpretation of MRCP alone: Judging the presence of stones, their type, and the classification of biliary obstruction based solely on 2D MRCP static images and 3D MRCP reconstructed images. During interpretation, non-contrast MRI sequences were not to be referenced.

The third level - Combined interpretation: Two weeks after the completion of the previous two levels of interpretation (washout period), the same group of physicians reviewed all sequences of non-contrast MRI and all images of MRCP simultaneously for a comprehensive interpretation. The final diagnosis results were recorded.

In cases of disagreement, a third senior radiologist with > 10 years of experience in abdominal imaging adjudicated to reach a consensus, which was used as the final interpretation for that method.

Inter-observer agreement for biliary stone diagnosis was substantial (κ = 0.737-0.795), and intra-observer agreement was almost perfect (κ = 0.889-0.930), confirming good reproducibility of the interpretation.

Severity assessment of biliary obstruction

The Bismuth classification system was used to assess the severity of biliary obstruction (Plentz and Malek, 2015). The grading criteria are as follows:

Type I: The hepatic duct junction is not affected, and the stenosis is located distally to the common hepatic duct.

Type II: The stenosis extends to the hepatic duct junction but does not affect the left and right hepatic ducts.

Type III: Obstruction affects the left or right hepatic duct.

Type IV: Obstruction simultaneously affects the left and right hepatic ducts.

Type V: The common hepatic duct is narrowed and also affects the branches of the intrahepatic bile ducts.

Of note, Bismuth Type V is not part of the original classification. In this study, we extended the system to include Type V for cases involving both intrahepatic duct branches and common hepatic duct narrowing, which were encountered in our cohort. This adaptation was made for descriptive purposes and should be interpreted with caution.

Reference standards

The surgical records were used as the reference standard. The presence or absence of stones, their location, number, and type (extra-

hepatic bile duct stones, secondary bile duct stones, intrahepatic bile duct stones), and the Bismuth classification determined during surgical exploration served as the reference standard. For cases without stones on surgical exploration, the surgical findings were used as the reference for a negative diagnosis.

Statistical analysis

Data were analyzed using SPSS 26.0. Continuous variables were presented as mean ± standard deviation (SD), and categorical variables as frequencies and percentages. The diagnostic accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of non-contrast MRI, MRCP, and combined diagnosis were calculated based on the surgical results as the gold standard. Paired χ^2 test (McNemar's test) was used to compare the combined diagnosis with each single method in 2 × 2 contingency tables. Cohen's kappa coefficient was used to assess agreement with the gold standard ($\kappa \leq 0.20$, poor; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, good; 0.81-1.00, excellent). The diagnostic efficacy of each method was evaluated using the receiver operating characteristic (ROC) curve analysis. The area under the curve (AUC) and its 95% confidence interval were calculated. The AUC was compared using the Z test, and the standard error was estimated by the bootstrap method (1000 resampling). The optimal diagnostic threshold was determined based on the Youden index. The net reclassification improvement (NRI) index was used to assess the classification improvement of the combined diagnosis compared to a single method.

In addition, stratified analysis was conducted according to smoking status, diabetes, hypertension, jaundice symptoms, and median age. The sensitivity and specificity of different methods in each subgroup were calculated. The differences in laboratory indicators between the correct and incorrect diagnosis groups were compared using the Mann-Whitney U test. The significance level $\alpha = 0.05$ (two-sided).

Results

Baseline characteristics

A total of 105 patients suspected of having cholelithiasis were consecutively included. Surgical confirmation identified stones in 89 patients (84.8%) and no stones in 16 patients (15.2%). The average age of the entire group was 62.43 ± 14.04 years, with 40 males (38.1%) and 65 females (61.9%). The total bilirubin level in the stone-positive group (50.9 ± 55.6 μmol/L) was higher than that in the stone-negative group, consistent with the biochemical profile of obstructive jaundice. The other demographic and clinical characteristics were evenly distributed between the two groups (Table S1). Among the 16 patients with negative surgical findings for stones, the causes of clinical presentation were as follows: acute or chronic cholecystitis without stones (n = 7), cholangitis (n = 4), biliary stricture of unclear etiology (n = 2), pancreatic head lesion (n = 1), and indeterminate cause (n = 2). Notably, 11 of these 16 patients presented with jaundice, with final diagnoses including cholangitis (n = 4), biliary stricture (n = 2), and cholecystitis (n = 5).

Accuracy of three imaging methods for diagnosing biliary calculi

Taking the surgical results as the gold standard, non-contrast MRI correctly diagnosed 80 cases, yielding an overall accuracy of 76.19% (80/105), with poor agreement with the gold standard ($\kappa = 0.309$, $P < 0.001$); MRCP correctly diagnosed 88 cases, with an accuracy rate of 83.81% (88/105), and the consistency was moderate ($\kappa = 0.491$, $P < 0.001$); combined diagnosis correctly diagnosed 101 cases, with an accuracy rate of 96.19% (101/105), and the consistency was good ($\kappa = 0.860$, $P < 0.001$). The accuracy of the combined diagnosis was significantly higher than that of non-contrast MRI ($\chi^2 = 58.594$, $P < 0.001$) and MRCP alone ($\chi^2 = 54.378$, $P < 0.001$). The 2 × 2 cross-tabulation of the three methods is shown in **Table 1**; combined diagnosis increased the sensitivity from 78.7% (non-contrast MRI) and 85.4% (MRCP) to 96.6%, and the specificity from 62.5% and 75.0% to 93.8% (Table S2).

Accuracy of stone type diagnosis

Among the 89 patients with surgically confirmed stones, surgical findings identified 45 extrahepatic, 34 secondary, and 10 intrahepatic bile duct stones. The overall correct classification numbers of stone types by non-con-

Table 1. Cross-tabulation of non-contrast MRI, MRCP, and combined interpretation against surgical findings for diagnosing biliary stones

Method	Diagnosis	Gold Positive	Gold Negative	χ^2	*P* value
Non-contrast MRI	Positive	70	6	9.522	0.002
Non-contrast MRI	Negative	19	10		
MRCP	Positive	76	4	24.040	< 0.001
MRCP	Negative	13	12		
Combined	Positive	86	1	71.758	< 0.001
Combined	Negative	3	15		

Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

Table 2. Diagnostic accuracy of non-contrast MRI, MRCP, and their combination for stone type classification

Stone type	Non-contrast MRI	MRCP	Combined
Extrahepatic	35/45 (77.8%)	37/45 (82.2%)	42/45 (93.3%)
Secondary	27/34 (79.4%)	30/34 (88.2%)	34/34 (100.0%)
Intrahepatic	8/10 (80.0%)	9/10 (90.0%)	10/10 (100.0%)

Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

Table 3. Diagnostic accuracy of the three imaging methods for grading biliary obstruction according to the Bismuth classification

Bismuth type	Non-contrast MRI	MRCP	Combined
I	20/24 (83.3%)	22/24 (91.7%)	23/24 (95.8%)
II	14/16 (87.5%)	15/16 (93.8%)	16/16 (100.0%)
III	10/14 (71.4%)	12/14 (85.7%)	13/14 (92.9%)
IV	8/10 (80.0%)	9/10 (90.0%)	10/10 (100.0%)
V	3/4 (75.0%)	3/4 (75.0%)	4/4 (100.0%)

Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

trast MRI, MRCP, and combined diagnosis were 70 cases (78.65%), 76 cases (85.39%), and 86 cases (96.63%), respectively. The accuracy rates of each subtype are detailed in **Table 2**. Among them, the correct identification rates of secondary and intrahepatic bile duct stones by combined diagnosis reached 100.0%.

Diagnostic accuracy of bismuth grading for severity of obstruction

Among the 89 patients with stones, 68 were found to have biliary obstruction during surgery. The distribution of Bismuth types I through V was 24, 16, 14, 10, and 4 cases, respectively. The number of correct cases for Bismuth grading by non-contrast MRI, MRCP, and combined diagnosis were 55 (80.88%), 61 (89.71%), and 66 (97.06%) respectively. The accuracy rates for each type are shown in **Table 3**. The combined diagnosis achieved 100.0% accuracy for types II, IV, and V, with only 1 misjudgment in type I and 1 in type III. The analysis of misjudgment patterns (Table S3) indicated that non-contrast MRI and MRCP mainly underestimated the obstruction level, while combined diagnosis almost completely eliminated the underestimation error. **Figure 2** visually presents the accuracy rates of the three methods under different Bismuth classifications. The accuracy of the combined diagnosis remained above 92% in all classifications.

ROC curve analysis for diagnostic efficacy

ROC analysis (**Tables 4**, S4; **Figure 3**) showed that the AUC of non-contrast MRI is 0.706 (95% CI 0.574-0.839), that of MRCP is 0.801 (95% CI 0.688-0.914), and the combined diagnosis reaches 0.950 (95% CI 0.871-0.995). The AUC of the combined diagnosis is significantly higher than that of non-contrast MRI (Z = 3.240, P = 0.001) and MRCP (Z = 2.252, P = 0.024), and the Youden index increases from 0.412 (non-contrast MRI) and 0.604 (MRCP) to 0.904.

Re-classification improvement and stratified analysis

Based on the combined diagnosis as the reference, 18 cases of misjudgment (positive for stones) and 5 cases of non-event (negative for stones) by non-contrast MRI were correctly reclassified. MRCP had 11 cases and 3 cases respectively (**Figure 4**), but the net re-classification improvement did not reach statistical significance. After stratification by smoking, diabetes, hypertension, jaundice, and median

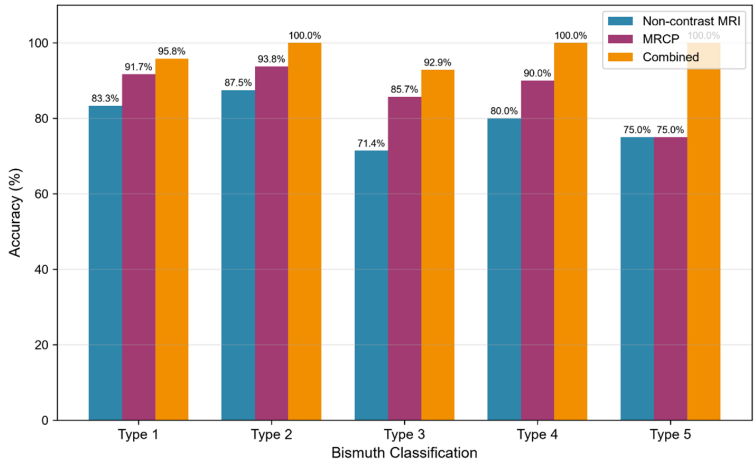


Figure 2. Grouped bar chart comparing diagnostic accuracy of the three methods across Bismuth type I-V. Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

age (Table S5), the combined diagnosis had a sensitivity of ≥ 0.889 and a specificity of ≥ 0.800 in each subgroup. It is notable that the specificity of non-contrast MRI in the diabetes subgroup was only 0.400, and the combined diagnosis increased it to 0.800; in the hypertension subgroup, the specificity of the combined diagnosis reached 1.000; however, as noted in the Discussion, this estimate is based on only 2 negative cases and should be interpreted with caution.

Relationship between laboratory indicators and diagnostic accuracy

Analysis of the three laboratory indicators (Table S6; **Figure 5**) revealed that the alanine aminotransferase (ALT) level in the group with incorrect diagnosis by the combined strategy was significantly lower than that in the group with correct diagnosis (19.76 ± 8.89 U/L vs. 40.26 ± 26.87 U/L, Mann-Whitney U = 321.0, P = 0.047), suggesting that patients with low ALT levels might be more prone to individual missed or incorrect diagnoses in the joint diagnosis due to less severe obstruction and less typical imaging manifestations. The NRI analysis is presented in Table S7. There was no statistically significant difference in total bilirubin and white blood cell count between the correct and incorrect groups (P > 0.05).

Discussion

This study conducted a retrospective analysis on 105 patients with suspected cholelithiasis confirmed by surgery, systematically evaluating the diagnostic performance of non-contrast MRI combined with MRCP for biliary stones and obstruction severity. The results showed that the overall accuracy of the combined diagnosis was 96.19%, significantly superior to non-contrast MRI (76.19%) and MRCP used alone (83.81%), with good consistency (κ = 0.860). ROC analysis confirmed that the AUC of the combined diagnosis reached 0.950, significantly higher than the single method. Additionally, the accuracy of the combined diagnosis for stone type classification and Bismuth grading was also significantly improved. The misjudgment mode analysis revealed that it almost eliminated the common phenomenon of underestimating the obstruction level in the single method. These results indicate that combined interpretation of non-contrast MRI and MRCP provides more comprehensive and reliable diagnostic information.

In this study, the overall accuracy rate of MRCP (83.81%) was superior to that of non-contrast MRI (76.19%). The advantage mainly lies in the diagnosis of extrahepatic bile duct stones and secondary bile duct stones, as well as the identification of Bismuth II-IV type obstructions. This is consistent with previous reports [18-21]. MRCP uses a heavy T2-weighted sequence to make bile and pancreatic fluid show significantly high signals, while stones, as intraluminal filling defects, present as low signals. The excellent liquid-solid contrast provides a natural advantage for judging the presence of stones and the degree of lumen stenosis. In contrast, although non-contrast MRI can provide abundant soft tissue information, in sequences lacking a significant signal difference between bile and stones, small or equal-signal stones are prone to be missed or misjudged as artifacts, which may be the main reason for its relatively lower sensitivity [22-24].

Although MRCP performs well in displaying the lumen, the misjudgment pattern analysis of this study revealed that MRCP underestimated the obstruction level in some cases of Bismuth type III. This might be related to the limited spa-

Table 4. Receiver operating characteristic (ROC) curve analysis and diagnostic performance metrics for each imaging strategy

Method	AUC	SE	95% CI	Sensitivity	Specificity	Youden	Kappa	Z vs Combined	P vs Combined
Non-contrast MRI	0.706	0.068	0.574-0.839	0.787	0.625	0.412	0.309	3.305	< 0.001
MRCP	0.801	0.058	0.688-0.914	0.854	0.750	0.604	0.491	2.243	0.025
Combined	0.950	0.034	0.871-0.995	0.966	0.938	0.904	0.860	Reference	Reference

Abbreviations: AUC, area under the curve; CI, confidence interval; SE, standard error; MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

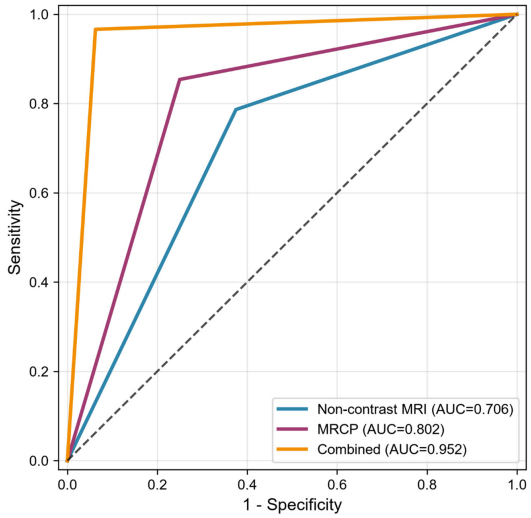


Figure 3. ROC curves of magnetic resonance plain scan, MRCP, and their combined diagnosis for biliary calculi. Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography; ROC, receiver operating characteristic; AUC, area under the curve.

tial resolution of MRCP for evaluating the hepatic portal and secondary bile duct branches. When the bile ducts are not significantly dilated or the stenotic segment is short, the projection overlap effect of the 3D reconstructed MIP images may mask the subtle lumen stenosis, resulting in an underestimation of the obstruction range [25, 26]. Additionally, respiratory motion artifacts and the overlap of high-signal intestinal fluid may also affect the diagnostic confidence of MRCP in the hepatic portal area.

The combined diagnosis significantly compensated for the limitations of each individual modality. A key finding of this study was that combined interpretation improved sensitivity from 78.7% (non-contrast MRI) and 85.4% (MRCP) to 96.6%, and specificity from 62.5% and 75.0% to 93.8%. More importantly, misclassification analysis revealed that underestimation of obstruction severity was the predominant error pattern when non-contrast MRI or MRCP was used alone, while the combined diagnosis achieves 100% accurate interpretation in Bismuth types II, IV, and V. This synergistic effect may be attributed to the complementarity of the two imaging modalities - non-contrast MRI provides coronal and axial thin-layer T2WI and DWI sequences that can clearly show indirect obstructive signs such as thickening and edema of the bile duct wall and surrounding soft tissues, while MRCP presents the entire biliary tree in a comprehensive manner and clearly shows the location and morphology of filling defects in the lumen, and the combination of the two helps to more completely restore the true anatomical relationship of biliary obstruction [27, 28]. Therefore, in clinical practice, for cases where it is difficult to clearly determine the stenosis range or the relationship between the stone and the bile duct wall with MRCP, adding the non-contrast MRI sequence may significantly improve the diagnostic accuracy.

ROC curve analysis quantitatively confirmed the efficacy advantage of the combined diagnosis. The AUC of the combined diagnosis reached 0.950 (95% CI 0.871-0.995), significantly higher than that of non-contrast MRI (0.706) and MRCP (0.801), and the differences were statistically significant by Z-test (P = 0.001 and P = 0.024). The Youden index increased from 0.412 of non-contrast MRI and 0.604 of MRCP to 0.904, approaching the theoretical optimal value of 1.0. These results indicate that reviewing non-contrast MRI and MRCP images together can optimize the trade-off between sensitivity and specificity, reducing unnecessary further examinations or misdiagnosis-induced improper treatment.

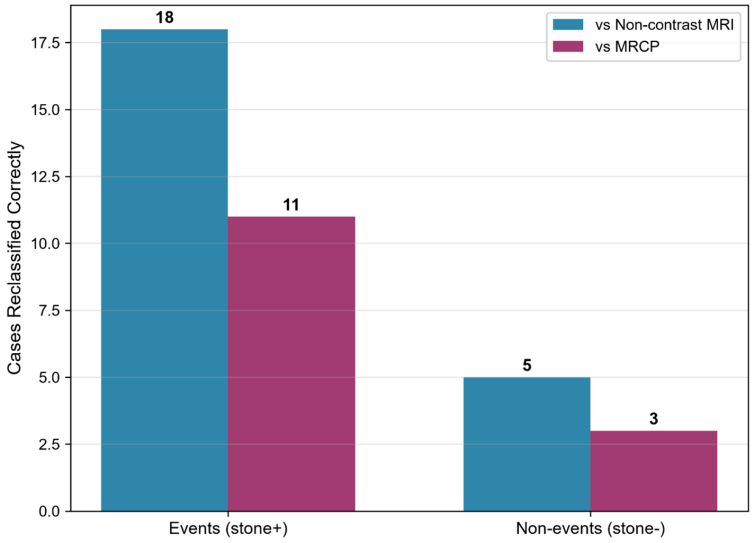


Figure 4. Number of cases correctly reclassified by combined imaging compared to non-contrast MRI and MRCP. Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography; NRI, net reclassification improvement.

This study also explored the stratified influence of clinical characteristics on the diagnostic efficacy of the three methods. The combined diagnosis demonstrated robust diagnostic performance across all subgroups, with sensitivity ≥ 0.889 and specificity ≥ 0.800. It was largely unaffected by factors such as patients' smoking status, diabetes, hypertension, jaundice, or age, indicating its wide clinical applicability. In contrast, non-contrast MRI and MRCP alone had specificity of only 0.400 and 0.600, respectively, in diabetic patients, whereas the combined diagnosis increased specificity to 0.800. Diabetic patients may have abnormal biliary motility and concurrent fatty liver with high signal backgrounds, reducing the stone-background contrast of a single imaging method and increasing false positives [29]. The addition of MRCP lumen imaging in the combined diagnosis can effectively identify such false positives, suggesting that for suspected gallstone patients with diabetes, adopting a combined strategy offers more significant benefits.

In the analysis of the relationship with laboratory indicators, this study found that the ALT level in the group with incorrect joint diagnosis interpretation was significantly lower than that in the group with correct interpretation (P = 0.047). Since ALT is one of the biochemical indicators reflecting liver cell damage and the degree of biliary obstruction [30, 31], a low ALT level may indicate milder biliary obstruction or an early disease stage. At this time, the imaging changes may be more subtle or atypical, increasing the diagnostic difficulty. This finding indirectly suggests that in patients with low ALT levels but high clinical suspicion of cholelithiasis, even if the combined imaging diagnosis does not show a clear positive finding, vigilance should still be maintained and short-term follow-up or other auxiliary examinations should be considered for confirmation. There was no statistically significant difference in total bilirubin and white blood cell count between the correct and incorrect groups, which may be related to the small sample size of this study and the non-specificity of these indicators.

The Net Classification Improvement (NRI) analysis showed that the overall NRI of the combined diagnosis compared to non-contrast MRI and MRCP was -0.133 and -0.075 respectively, and neither reached statistical significance (P > 0.05). This apparent inconsistency may be explained by the fact that NRI and AUC assess different aspects of diagnostic performance [32, 33]. NRI measures the net improvement in reclassification at the individual patient level, while the non-contrast MRI and MRCP in this study already had baseline accuracies of 76%-84%, and the "space" for reclassification improvement is inherently limited; moreover, with a sample size of 105, the number of reclassification events is small, and the limited statistical power may also have prevented NRI from reaching significance. Although the NRI was not statistically significant, **Figure 4** visually demonstrates that the combined diagnosis reclassified 18 false-positive and 5 false-negative cases compared to non-contrast MRI, and 11 event cases and 3 non-event cases compared to MRCP. This suggests the combined technique has clinical value. In future larger sample or multi-center studies, the reclassi-

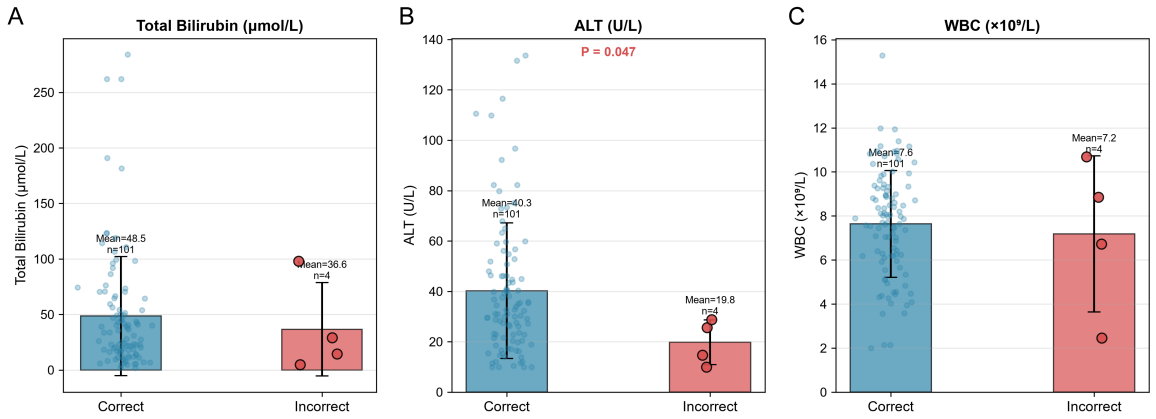


Figure 5. Laboratory parameters in patients correctly (n = 101) vs. incorrectly (n = 4) diagnosed by the combined strategy. (A) Total bilirubin, (B) ALT, and (C) WBC. ALT was significantly lower in the incorrect group (19.76 ± 8.89 vs. 40.26 ± 26.87 U/L, P = 0.047). Data are mean ± SD. Abbreviations: ALT, alanine aminotransferase; WBC, white blood cell; SD, standard deviation.

fication improvement effect of the combined strategy may become more clear.

This study has the following limitations. First, this was a single-center retrospective study with a relatively small sample size (n = 105), which may have introduced selection bias. Future studies should conduct multi-center prospective research to verify the generalizability of the conclusions. Second, the image interpretation was completed by the same group of physicians. Although a 2-week washout period was set to minimize the memory bias, it was still impossible to eliminate completely the residual information interference when interpreting the three levels. Ideally, independent combined interpretation by another group of blinded radiologists would be preferable. Third, surgical findings were used as the reference standard. However, for negative cases with stones, the false-negative rate of intraoperative exploration is difficult to ascertain, and may deviate from the gold standard in rare cases of microlithiasis or sludge. Fourth, in stratified analysis, the number of negative cases in some subgroups (such as the smoking group and the hypertension group) is relatively small, resulting in low precision for specificity estimates. Some results with a specificity of 1.000 should be interpreted with caution. Fifth, this study did not include other imaging methods such as ultrasound and CT for comparison. Therefore, the specific incremental value of the combined MRI protocol compared to these methods needs to be verified.

Conclusion

For patients with suspected biliary stones undergoing MRI, combined interpretation of non-contrast MRI and MRCP significantly improved diagnostic accuracy and Bismuth grading compared to either method alone. However, the net reclassification improvement over single-method interpretation was not statistically significant, suggesting uncertain benefit at the individual patient level. This combined strategy may be considered in patients already undergoing MRI, and its clinical utility compared with standard first-line modalities warrants further validation in larger prospective studies.

Disclosure of conflict of interest

None.

Address correspondence to: Pinglin Fan, Department of Imaging, Yingshang County People's Hospital, No. 566, Ganluo Road, Shencheng Town, Yingshang County, Fuyang 236200, Anhui, China. Tel: +86-0558-2829623; E-mail: fanpl85@163.com

References

[1] Afzalpurkar S, Giri S, Kasturi S, Ingawale S and Sundaram S. Magnetic resonance cholangiopancreatography versus endoscopic ultrasound for diagnosis of choledocholithiasis: an updated systematic review and meta-analysis. Surg Endosc 2023; 37: 2566-2573.

[2] Shabanikia N, Adibi A and Ebrahimian S. Diagnostic accuracy of magnetic resonance

cholangiopancreatography to detect benign and malignant biliary strictures. Adv Biomed Res 2021; 10: 38.
[3] Swaraj S, Mohapatra M, Sathpathy G, Yalamanchi R, Sen K, Menon SM, Madhesia A, Kumaraswamy SM, Radha Krishna K and Bobde DV. Diagnostic performance of ultrasonography versus magnetic resonance cholangiopancreatography in biliary obstruction. Cureus 2023; 15: e33915.
[4] Ali F, Aamir N, Hassan MK, Khan HU and Khan D; Fakhar-E-Alam. Comparison of MRCP and ERCP findings: a retrospective secondary data analysis. J Pak Med Assoc 2022; 72: 284-286.
[5] Udaikumar J, Nimmagadda R, Potluri V, Medarametla R, Garlapati S, Tummala N, Kuppili S, Goodman AJ and Marino D. Comparing endoscopic ultrasound (EUS) vs. magnetic resonance cholangiopancreatography (MRCP) in the etiological evaluation of idiopathic acute pancreatitis (IAP): a systematic review and meta-analysis. Dig Dis Sci 2026; 71: 1108-1118.
[6] Sun K, Li M, Shi Y, He H, Li Y, Sun L, Wang H, Jin C, Chen M and Li L. Convolutional neural network for identifying common bile duct stones based on magnetic resonance cholangiopancreatography. Clin Radiol 2024; 79: 553-558.
[7] Samara O, Azzam MI, Alshrouf MA, Khanfar AN, Mohialdeen RR, Barham OF, Barham MF and AlRyalat SA. Diagnostic accuracy of ultrasonography compared with magnetic resonance cholangiopancreatography in the detection of choledocholithiasis. J Clin Ultrasound 2022; 50: 247-253.
[8] Katariya P, Vaishnani B, Gamit H, Vaghela S and Jasani K. Comparative diagnostic accuracy of ultrasonography and magnetic resonance cholangiopancreatography (MRCP) in the evaluation of obstructive jaundice: a prospective study in Western India. Cureus 2025; 17: e100312.
[9] Chen WC, Lo HY, Tsai YS and Yang YJ. The accuracy of magnetic resonance cholangiopancreatography in the diagnosis of biliary atresia in preterm infants with cholestasis. Pediatr Neonatol 2023; 64: 12-18.
[10] Isram J, Haider E, Khan RSA, Hafeez M, Hinna RE, Baig I, Shahid A and Chaudhry M. Diagnostic accuracy of magnetic resonance cholangiopancreatography in comparison with endoscopic retrograde cholangiopancreatography for detection of the etiology of obstructive jaundice. Cureus 2023; 15: e34484.
[11] Erden A, Kuru Öz D and Erden İ. Ectopic opening of common bile duct into the duodenal bulb: magnetic resonance cholangiopancreatography findings. Diagn Interv Radiol 2022; 28: 286-293.
[12] Al-Dulaimi M, Ibraheem M, Abdulkareem M, Al-Dujaili A and Abdulkareem A. Comparative diagnostic accuracy of clinical assessment, computed tomography (CT), and magnetic resonance imaging with magnetic resonance cholangiopancreatography (MRI/MRCP) in evaluating common bile duct stones. Cureus 2025; 17: e94572.
[13] Kumar A, Mohanty NR, Mohanty M and Dash S. Comparison of MRCP and ERCP in the evaluation of common bile duct and pancreatic duct pathologies. Front Med Technol 2023; 5: 946555.
[14] Karpagam RK, Ravipati C, Ramakrishnan KK, Ramaswami S and Natarajan P. Comparative efficacy of magnetic resonance cholangiopancreatography vs. percutaneous transhepatic cholangiography with percutaneous transhepatic biliary drainage stenting in evaluating obstructive jaundice: a prospective study in South India. Cureus 2024; 16: e65241.
[15] Suzuki M, Sekino Y, Hosono K, Yamamoto K, Kawana K, Nagase H, Kubota K and Nakajima A. 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: 1052-1059.
[16] Khan T, Alam SZ, Al Mamun M, Rahman MS, Islam MU and Bari S. Accuracy of magnetic resonance cholangiopancreatography in case of biliary obstruction comparing post-operative findings: a study of 50 cases. Mymensingh Med J 2021; 30: 1079-1085.
[17] Rațiu I, Lupușoru R, Lungeanu D, Popescu A, Sporea I, Goldiș A, Dănilă M, Miuțescu B, Moga T, Barbulescu A, Tăban S, Dema A and Șirli R. Diagnosis of malignant biliary obstruction: pondering over the ERCP, MRCP and histology. J Int Med Res 2022; 50: 3000605221076924.
[18] Yoo J, Min JH, Lee DH, Hur BY, Kim SW and Kim E. Abbreviated magnetic resonance imaging with breath-hold three-dimensional magnetic resonance cholangiopancreatography: assessment of malignant risk of pancreatic intraductal papillary mucinous neoplasm. J Magn Reson Imaging 2021; 54: 1177-1186.
[19] Mao Y, Ni J, Peng K, Yu Z, Luo S, Xia Y, Fu S, Qu Y, Xu K, Lu L, Gong X, Zhong N and Li B. Superiority of linear-array EUS over MRCP in diagnosing pancreas divisum: evidence from a multicenter retrospective study in Oriental cohorts (with video). Gastrointest Endosc 2025; 102: 233-241, e231.
[20] Mattila A, Helminen O, Pynnönen E, Sironen A, Elomaa E and Nevalainen M. Preoperative MRCP can rule out choledocholithiasis in acute cholecystitis with a high negative predictive

value: prospective cohort study with intraoperative cholangiography. J Gastrointest Surg 2023; 27: 2396-2402.
[21] Yang Q, Ding X, Guo Q, Tang Y, Lin J, Huang Y, Liu M and Lei J. Advantages of BioMatrix respiratory gating in free-breathing three-dimensional magnetic resonance cholangiopancreatography: a prospective comparative study. Insights Imaging 2025; 16: 137.
[22] Bush AES, Christopoulos P, Jones RM, Sinha S, Srinivas G and Andrews SN. Safety, quality and efficiency of intra-operative imaging for treatment decisions in patients with suspected choledocholithiasis without pre-operative magnetic resonance cholangiopancreatography. Surg Endosc 2022; 36: 1206-1214.
[23] Wang GX, Ge XD, Zhang D, Chen HL, Zhang QC and Wen L. MRCP combined with CT promotes the differentiation of benign and malignant distal bile duct strictures. Front Oncol 2021; 11: 683869.
[24] Jagtap N, Kumar JK, Chavan R, Basha J, Tandan M, Lakhtakia S, Kalapala R, Nabi Z, Gupta R, Ramchandani M, Talukdar R, Reddy M, Yarlagadda R, Singh J, Memon SF, Venkat Rao G and Reddy DN. EUS versus MRCP to perform ERCP in patients with intermediate likelihood of choledocholithiasis: a randomised controlled trial. Gut 2022; [Epub ahead of print].
[25] Inomata T, Nakaya K, Michimoto K, Kano R, Masuda Y, Suzuki H, Sawaguchi N, Sugawara K and Sugiyama S. Evaluation of the usefulness of cystic duct three-dimensional computed tomography with non-contrast for before laparoscopic cholecystectomy and endoscopic transpapillary gallbladder drainage in comparison to magnetic resonance cholangiopancreatography. J Med Imaging Radiat Sci 2021; 52: 248-256.
[26] Hou JU, Park SW, Park SM, Park DH, Park CH and Min S. Efficacy of an artificial neural network algorithm based on thick-slab magnetic resonance cholangiopancreatography images for the automated diagnosis of common bile duct stones. J Gastroenterol Hepatol 2021; 36: 3532-3540.
[27] Eurboonyanun K, Promsorn J, Sa-Ngiamwibool P, Eurboonyanun C, Finnegan S, Ferreira C, Herlihy A, Shumbayawonda E, Lahoud RM, Atre I, O'Shea A and Harisinghani M. Quantitative MRCP metrics as imaging biomarkers to differentiate benign from malignant bile duct obstructions. Front Oncol 2025; 15: 1576163.
[28] Patel VB, Musa RK, Patel N and Patel SD. Role of MRCP to determine the etiological spectrum, level and degree of biliary obstruction in obstructive jaundice. J Family Med Prim Care 2022; 11: 3436-3441.
[29] Zhao H, Xu Q, Gao R, Yin B, Sun G, Xue K, Yang Y, Li E, Zhu L, Feng F and Wu W. Clinical feasibility of 5.0 T MRI/MRCP in characterizing pancreatic cystic lesions: comparison with 3.0 T and MDCT. Diagnostics (Basel) 2024; 14: 2457.
[30] Nakagawa Y, Uchida H, Hinoki A, Shirota C, Tainaka T, Sumida W, Makita S, Yasui A, Takimoto A and Guo Y. Laparoscopic bile duct plasty for hilar bile duct stenosis (HBDS) in patients with congenital biliary dilatation: Diagnosis of HBDS by preoperative MRCP and laparoscopic strategy to relieve HBDS. J Hepatobiliary Pancreat Sci 2023; 30: 473-481.
[31] Khan RSA, Alam L, Khan ZA and 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: 1275-1279.
[32] Thoi DKT, Lim JH, Park JS and Park S. Deep learning-assisted diagnosis of extrahepatic common bile duct obstruction using MRCP imaging and clinical parameters. Curr Med Imaging 2025; 21: e15734056363648.
[33] Liu S, Zhang R, He Y, Liu Y, Wang R, Zhou Z, Ma H, He X, Yan S, Huang L, Jiang K and Liu H. Adjunctive value of 3D MRCP in biliary atresia: a retrospective two-center analysis of cholestatic infants. Insights Imaging 2025; 16: 277.

Table S1. Baseline demographic and clinical characteristics of the study population

Variable	All (n = 105)	Stone+ (n = 89)	Stone- (n = 16)	Z/χ^2	
Age (years)	62.43 ± 14.04	61.89 ± 13.92	65.44 ± 14.79	-1.070	0.286
Weight (kg)	56.46 ± 4.37	56.68 ± 4.34	55.21 ± 4.51	1.106	0.271
Total bilirubin (μmol/L)	48.1 ± 53.0	50.9 ± 55.6	32.5 ± 32.8	1.556	0.121
ALT (U/L)	39.48 ± 26.69	39.33 ± 26.52	40.30 ± 28.48	0.067	0.950
WBC ($\times 10^9$/L)	7.62 ± 2.45	7.64 ± 2.47	7.49 ± 2.45	0.116	0.911
Gender				0.111	0.739
Male	40 (38.1%)	35 (39.3%)	5 (31.2%)		
Female	65 (61.9%)	54 (60.7%)	11 (68.8%)		
Smoking				*	0.740
No	83 (79.0%)	71 (79.8%)	12 (75.0%)		
Yes	22 (21.0%)	18 (20.2%)	4 (25.0%)		
Diabetes				*	0.525
No	80 (76.2%)	69 (77.5%)	11 (68.8%)		
Yes	25 (23.8%)	20 (22.5%)	5 (31.2%)		
Hypertension				2.420	0.120
No	71 (67.6%)	57 (64.0%)	14 (87.5%)		
Yes	34 (32.4%)	32 (36.0%)	2 (12.5%)		
Chronic liver disease				*	0.351
No	97 (92.4%)	83 (93.3%)	14 (87.5%)		
Yes	8 (7.6%)	6 (6.7%)	2 (12.5%)		
Pancreatic disease				*	0.999
No	96 (91.4%)	81 (91.0%)	15 (93.8%)		
Yes	9 (8.6%)	8 (9.0%)	1 (6.2%)		
Abdominal pain				*	0.592
No	7 (6.7%)	7 (7.9%)	0 (0.0%)		
Yes	98 (93.3%)	82 (92.1%)	16 (100.0%)		
Fever				0.978	0.323
No	57 (54.3%)	46 (51.7%)	11 (68.8%)		
Yes	48 (45.7%)	43 (48.3%)	5 (31.2%)		
Jaundice				1.146	0.284
No	49 (46.7%)	44 (49.4%)	5 (31.2%)		
Yes	56 (53.3%)	45 (50.6%)	11 (68.8%)		
Leukocytosis				0.004	0.948
No	55 (52.4%)	46 (51.7%)	9 (56.2%)		
Yes	50 (47.6%)	43 (48.3%)	7 (43.8%)		

Abbreviations: ALT, alanine aminotransferase; WBC, white blood cell; SD, standard deviation. Note: Continuous variables are presented as mean ± SD; categorical variables as n (%). *P* values were calculated using Mann-Whitney U test (Z statistic) for continuous variables, and χ^2 test or Fisher's exact test for categorical variables as appropriate. *Fisher's exact.

Table S2. Overall diagnostic performance of the three imaging methods

Method	Sensitivity	Specificity	PPV	NPV	Accuracy	Youden	Kappa
Non-contrast MRI	0.787	0.625	0.921	0.345	0.762	0.412	0.309
MRCP	0.854	0.750	0.950	0.480	0.838	0.604	0.491
Combined	0.966	0.938	0.989	0.833	0.962	0.904	0.860

Abbreviations: PPV, positive predictive value; NPV, negative predictive value; MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

Table S3. Misclassification patterns of Bismuth grading by the three imaging methods

Bismuth Gold	N	Non-contrast MRI_Under	Non-contrast MRI_Over	Non-contrast MRI_Exact	MRCP Under	MRCP Over	MRCP Exact	Combined Under	Combined Over	Combined Exact
I	24	2	2	20	1	1	22	0	1	23
II	16	1	1	14	1	0	15	0	0	16
III	14	3	1	10	2	0	12	1	0	13
IV	10	2	0	8	1	0	9	0	0	10
V	4	1	0	3	1	0	3	0	0	4

Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

Table S4. Effect size analysis of AUC differences between combined and single modalities

Comparison	AUC Difference	95% CI of Difference	Z	P	Cohen's h (approx)
Combined vs Non-contrast MRI	0.247	0.101-0.394	3.305	< 0.001	0.591
Combined vs MRCP	0.150	0.019-0.282	2.243	0.025	0.415

Abbreviations: AUC, area under the curve; CI, confidence interval; MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

Table S5. Stratified diagnostic performance of each modality by clinical features (sensitivity and specificity)

Stratum	n_positive	n_negative	Non-contrast MRI_Sens	Non-contrast MRI_Spec	MRCP_ Sens	MRCP_ Spec	Combined_ Sens	Combined_ Spec
Smoking: No (n = 83)	71	12	0.817	0.583	0.831	0.667	0.986	0.917
Smoking: Yes (n = 22)	18	4	0.667	0.750	0.944	1.000	0.889	1.000
Diabetes: No (n = 80)	69	11	0.739	0.727	0.870	0.818	0.957	1.000
Diabetes: Yes (n = 25)	20	5	0.950	0.400	0.800	0.600	1.000	0.800
Hypertension: No (n = 71)	57	14	0.772	0.571	0.877	0.786	1.000	0.929
Hypertension: Yes (n = 34)	32	2	0.812	1.000	0.812	0.500	0.906	1.000
Jaundice: No (n = 49)	44	5	0.886	0.800	0.955	0.800	0.977	1.000
Jaundice: Yes (n = 56)	45	11	0.689	0.545	0.756	0.727	0.956	0.909
Age > median: No (n = 54)	48	6	0.833	0.667	0.875	1.000	0.958	1.000
Age > median: Yes (n = 51)	41	10	0.732	0.600	0.829	0.600	0.976	0.900

Abbreviations: MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography; Sens, sensitivity; Spec, specificity.

Table S6. Association between laboratory values and diagnostic correctness for each modality

Value	Method	Correct (Mean ± SD)	Incorrect (Mean ± SD)	U	Z	P	Effect Size (r)
Total Bilirubin (μmol/L)	Non-contrast MRI	47.07 ± 46.73	51.35 ± 70.62	1048.000	0.361	0.721	0.035
Total Bilirubin (μmol/L)	MRCP	49.87 ± 56.10	38.89 ± 32.67	806.500	0.509	0.614	0.050
Total Bilirubin (μmol/L)	Combined	48.54 ± 53.55	36.64 ± 42.08	240.500	0.644	0.525	0.063
ALT (U/L)	Non-contrast MRI	38.08 ± 24.65	43.98 ± 32.52	985.500	-0.109	0.916	0.011
ALT (U/L)	MRCP	39.45 ± 27.72	39.65 ± 21.21	690.000	-0.505	0.617	0.049
ALT (U/L)	Combined	40.26 ± 26.87	19.76 ± 8.89	321.000	1.992	0.047	0.194
WBC (× 10^9/L)	Non-contrast MRI	7.64 ± 2.54	7.54 ± 2.20	1035.500	0.267	0.792	0.026
WBC (× 10^9/L)	MRCP	7.78 ± 2.39	6.79 ± 2.67	911.000	1.418	0.157	0.138
WBC (× 10^9/L)	Combined	7.64 ± 2.43	7.18 ± 3.54	208.000	0.100	0.927	0.010

Abbreviations: ALT, alanine aminotransferase; WBC, white blood cell; SD, standard deviation; MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.

Table S7. Net reclassification improvement (NRI) of combined diagnosis compared with single methods

Comparison	Events: Corrected (Up/Down)	Non-events: Corrected (Up/Down)	NRI Event	NRI Non-event	NRI Overall	Z	P
Combined vs Non-contrast MRI	18/2	5/0	0.180	0.312	-0.133	-0.894	0.371
Combined vs MRCP	11/1	3/0	0.112	0.188	-0.075	-0.653	0.514

Abbreviations: NRI, net reclassification improvement; MRI, magnetic resonance imaging; MRCP, magnetic resonance cholangiopancreatography.